

Lyn Marie Batty. Self-Archiving of Articles Published in High-Impact Journals in the Social Sciences. A Master's paper for the M.S. in L.S. degree. April, 2009. 72 pages. Advisor: Robert M. Losee

Self-archiving of peer-reviewed articles represents one path toward increasing the public's exposure to relevant social science research. This paper discusses self-archiving behavior of scholars who have published articles in high-impact journals in the fields of anthropology, sociology, economics and political science. A total of 1,686 articles published during 2007 and 2008 were searched to determine which of the articles had been self-archived. The resulting data illuminates the self-archiving practices of 3,440 individual authors within the four social science fields. The results indicate that self-archiving behavior varies depending on the author's discipline, region, and author order. Self-archiving policies of journal publishers are also examined to compare restrictions on self-archiving with authors' self-archiving practices.

Headings:

Open access publishing

Scholarly publishing

Electronic data archives

SELF-ARCHIVING OF ARTICLES PUBLISHED  
IN HIGH-IMPACT JOURNALS IN THE SOCIAL SCIENCES:

by  
Lyn Marie Batty

A Master's paper submitted to the faculty  
of the School of Information and Library Science  
of the University of North Carolina at Chapel Hill  
in partial fulfillment of the requirements  
for the degree of Master of Science in  
Library Science.

Chapel Hill, North Carolina

April 2009

Approved by:

---

Advisor: Robert M. Losee, Ph.D.

## Table of Contents

Introduction.....	1
Literature Review.....	3
Open Access.....	3
Opening Access to Social Science Research .....	6
Self-Archiving.....	11
Institutional and Subject Repositories .....	13
Institutional and Funder Mandates.....	17
Scholars' Web Pages.....	20
Publisher Self-Archiving Policies.....	24
Methodology and Results .....	29
Methods.....	29
Results.....	32
Discussion and Conclusion .....	51
References.....	65
Appendix.....	70

## **Introduction**

Advances in information technology have produced new avenues for the communication of scholarly research. As print-only scholarly journals have given way to journals, articles, and other forms of scholarly communication in digital formats, the Internet has become the preferred medium for providing access to the scholarly journal article. Scholars, libraries, research institutions, funding agencies, and governments are in the process of trying to conceptualize the full extent to which the internet may be used to further the dissemination of scholarly research results, not merely within the circles of academia, but to any person, in any country and in any circumstance, who might benefit from the knowledge.

The open access movement, by demanding fewer barriers between the public and scholarly knowledge, has caused upheaval in a scholarly publishing tradition that can trace its roots back to the first publication of the *Philosophical Transactions* in 1665. Open access proponents have advocated for new publishing models that enable scholars to distribute their scholarly publications freely. Academic libraries have promoted open access policies as an economic counterbalance to the rapid inflation of the costs of scholarly journals, a phenomenon known as the “serials crisis.” Various academic disciplines have developed their own methods for opening access to their publications, by supporting subject repositories or by supporting the creation of open access journals within their fields.

The president of the Social Science Research Council has publicly called for deliberate steps toward opening access to the research and scholarship created within the social sciences (Calhoun, 2008). Calhoun discussed the opportunities lost by allowing social science debate to become enclosed in the tight circles of academia, noting that, “certainly a social science turned in on itself fails to achieve much public significance. But more important than the desire to promulgate the knowledge of social scientists is the effort to bring their knowledge to bear on pressing social issues.”

Scholars who produce research and seek to disseminate the results of their work face a range of decisions in determining how best to circulate their research results. Before the advent of the Internet, scholars sought to publish articles in the top-ranking journals in their field in order to maximize the dissemination of their work. Scholars have a range of additional options today, including publishing their research in open-access journals, or paying author fees to their publisher to support immediate access to their article online. They may seek to negotiate with publishers to procure additional author rights to post their paper in an e-print archive or on their own homepage.

Large research funding organizations, such as the National Institutes of Health (NIH) and The Wellcome Trust have led the way in mandating greater access to the research their funding supports. Universities in the United States, including Harvard, Stanford and MIT, and many more universities in the UK, Europe, and Australia, adopted policies that mandate archiving of faculty scholarship in institutional repositories. As academia continues to assess its critical role in the dissemination of research and scholarship, more attention is being paid to the university’s responsibility to provide

support to faculty and researchers who have largely been left to their own devices in attempting to facilitate wider access to their published research.

As a response to the open access movement, faculty and researchers have invested time and effort in fashioning their own methods for opening access to their published works, largely through self-archiving. Even though acting individually, the combined efforts of self-archiving scholars have had an impact. Self-archived articles may be located by searching for titles or author's names with commercial search engines such as Google, Google Scholar, and Yahoo!, allowing many peer-reviewed articles to be easily found and read for free on the Internet. Studies documenting the behavior of self-archiving authors in various disciplines have been instructive for institutions and funding bodies considering adopting mandated open access policies, and may provide insights to guide efforts to expand upon the current self-archiving practices of scholars.

This study was undertaken in order to uncover potential patterns in self-archiving behavior in a group of published scholars in the social sciences. The data gathered by this study is intended to help illuminate the self-archiving practices and behaviors of authors who publish in high-impact journals in the social sciences, and to guide suggestions for increasing open access to articles published in respected disciplinary journals.

## **Literature Review**

### **Open Access**

Open access proponents are concerned with removing barriers to scholarly literature in order to disseminate research and promote sharing of knowledge. Open

access literature has been defined as scholarly literature that is, “digital, online, free of charge, and free of most copyright and licensing restrictions” (Suber, 2005). While the public’s access to knowledge is universally accepted to be a “public good,” the early interest in the open access movement appears to have arisen less from philanthropic ideals and more from a pragmatic response to a number of disconcerting developments in scholarly communication. It had become apparent over time that the cost of buying subscriptions to the range of scholarly journals required by large research universities had inflated beyond the capability of even the most affluent institutions.

Originally, academic libraries sounded the alarm regarding the rapid acceleration in the cost of scholarly journals to academic libraries. The cost increases predicted the potential loss of access to large amounts of scholarship if the trend continued. In the years between 1986 and 2005, the Association of Research Libraries (ARL), consisting of the 120 largest research libraries in North America, collectively increased their serials expenditures by 302% (ARL, 2005). The crisis caused by escalating serial costs stirred a debate about the future of scholarly communication. At the same time, new technologies had created possibilities to vastly increasing the reach of scholarly communication and while at the same time reduce costs and eliminating barriers to access.

Before the advent of the Internet, contracting with large publishing companies was the most efficient way to disseminate a scholar’s work to a broad audience, and scholars benefited from trading their copyrights for the amount of access only a large print publisher could provide. Today, the incentives are not as clear for the scholar or for the university. Considering the reach of the Internet, it is possible that the traditional

method of transferring ownership of scholarly production to commercial publishers may, in fact, significantly limit access to research and scholarship.

Early open access advocates promoted opening access to previously inaccessible peer-reviewed scholarly articles through the use of new publishing technologies and a reexamination of the economics of publishing (Harnad, 1995; Guedon, 2001). Despite concern that open access ideas might signal the end of the scholarly journal, traditional scholarly publishing modes are still thriving. However, there are now a number of additional avenues available to authors to increase the possibilities of free public access to their scholarship. These avenues include author self-archiving of works published originally in subscription-only peer reviewed journals, author-pays open access, funder-subsidized open access, and a number of others (Willinsky, 2006, p. 212).

Universities and large funding agencies are in the process of moving toward more concrete efforts designed to make scholarship freely available. A policy statement recently released by four leading associations serving research universities<sup>1</sup> re-asserted the critical function of the university in the dissemination of knowledge. The document, titled, *A Call to Action*, stated the need for Universities to take a renewed leadership position in ensuring that research created within the university was widely and publicly accessible. The document suggested that the tacit acceptance of the transfer of a majority of university research output had, over time, lead to the frustration of certain critical functions of academia:

*Past norms and practical requirements for dissemination have led to practices of transferring control of access to and use of faculty work outside the academy, limiting the university and faculty members' ability to ensure broad dissemination and wide use. Where the academy has*

---

<sup>1</sup> Association of American Universities, Association of Research Libraries, The Coalition for Networked Information, and the National Association of State Universities and Land Grant Colleges.

*relinquished the ability to manage its intellectual capital to best serve its needs and priorities it should act to regain this capability.* (Association of American Universities *et al.*, 2009).

### **Opening Access to Social Science Research**

*People expect from the social sciences – anthropology, sociology, economics, and political science – the knowledge to understand their lives and control their future.* Edward O. Wilson

Open access to scholarly research produced within the social sciences provides opportunities to enrich public understanding of critical social issues. Social scientists produce relevant studies squarely addressing timely and universal themes, such as social inequality, the functioning of markets, electoral processes, family structures, human evolution, and conflict resolution.

Voices within the social sciences have acknowledged the need for greater access to the scholarly output of their fields. As was noted by the president of the Social Science Research Council (SSRC) in his paper, *Social Science for Public Knowledge*, “certainly a social science turned in on itself fails to achieve much public significance. But more important than the desire to promulgate the knowledge of social scientists is the effort to bring their knowledge to bear on pressing social issues” (Calhoun, 2008).

Social scientists from all disciplines are producing relevant studies on important social issues. As noted by a pair of anthropologists describing the universal relevance of their field, “in the era of globalization and cyberspace, [anthropologists] are reporting on conversations with war refugees in the Congo, Islamic militants in the slums of Egypt, illegal immigrants who clean your local Wal-Mart and can barely make the rent, and young women who lose their eyesight assembling computers in sweatshops . . .”

(Gusterson & Besteman, 2005, p.5). Health, poverty, employment, marriage, communities, voting, politics, immigration, war – all are issues to which engaged social scientists could lend their expertise by sharing the results of their research and knowledge.

The question is not the relevance of high-quality social science research to everyday concerns, but rather how to make the research accessible to the general public so that they begin to seek it out, as they have done with medical research. The goal of increasing access to medical research was the catalyst that drove some of early strides towards opening research to the general public. Funders of medical research, such as the National Institutes of Health (NIH) and The Wellcome Trust, actively promoted open access for the largely medical science research that their funding supported. These agencies have used their considerable leverage to mandate archiving of the articles supported by their funds within publicly accessible archives such as PubMed Central, often within 6 months after publication.

Additionally, the medical and biological sciences have produced a larger number of gold open access journals than the social sciences, including the significant open access publishers, Public Library of Science (<http://www.plos.org>) and BioMed Central (<http://www.biomedcentral.com>). However, the number of open access journals in the social sciences is slowly increasing. The Directory of Open Access Journals (DOAJ) currently identifies 73 Sociology journals, 54 Anthropology Journals, 110 Political Science journals, and 82 Economics journals. In order for a journal to be counted in the DOAJ, the journal must be fully open access, and must exercise quality control in the form of peer review or an editorial board.

Self-archiving practices are less prevalent in the social sciences than in a number of other sciences. Pre-prints are not circulated within the social sciences in the way that they are in the so-called hard sciences. In many sciences, preprints serve to rapidly disseminate answers to shared problems in the discipline and to establish the primacy of results. The humanities and the social sciences have traditionally placed greater emphasis upon the monograph as the primary scholarly publishing vehicle. As subject repositories such as arXiv started to develop, the physical and medical sciences had works ready-made for e-print repositories (pre-prints) as part of their disciplinary habits. Social scientists have lagged somewhat in taking up new self-archiving habits needed to provide open access to their articles.

The Social Science Research Network (<http://www.ssrn.com>) is a subject repository that has been successful in providing a centralized approach to a large corpus of social science research. Authors from all social science fields, including economics, business, law, political science and sociology can easily upload copies of their articles and papers into the archive. SSRN maintains a digital archive of over 186,000 downloadable full text documents, representing the publications of over 110,656 authors. SSRN statistics show high download activity. Paper downloads average over 7 million per year. It is primarily scholars from economics, law and business who have been most successful in using SSRN. However, greater numbers of political science authors and sociologists are now beginning to post preprints and postprints in SSRN.

Open access journals and digital repositories have the potential to increase the public's understanding of scholarship and scientific processes, yet mediation is also an important aspect of connecting the layperson and open access research (Zuccala, 2009).

In addition to posting digital copies of peer-reviewed research on the web, Zuccala also advocates for better understanding of “where and how laypeople look for peer-reviewed literature” (p. 359) and how laypeople interpret and understand scholarly literature in order to make it useful for their own purposes.

Scholarly treatments of social science issues, such as one might find in a political science journal discussing the reliability of polling data in presidential elections, or a peer-reviewed article measuring the correlation between military occupation and the incidence of terrorism within a country, are often presented in ways that are unfamiliar to the public. Scientists do not generally have a conversational or even journalistic writing style. This is not to say that the findings of scientists are incomprehensible or even disinteresting to the public, for stories abound of patients or their family members who have analyzed reams of studies explaining complicated medical trials in order to find answers to health questions. However, members of the public, along with practitioners working in fields interested in social science research, such as policy-makers, educators, social workers, community activists, and professional and citizen journalists, could potentially benefit from discussion and analysis of stand-alone research articles by those more familiar with the writing and methods of scholarly researchers. (Zuccala, 2009; Willinsky, 2004).

Two UK organizations, the Academy of Social Sciences and the Economic and Social Research Council (ESRC) prepared a joint paper titled, “*Learned Societies in the Social Sciences: Developing Knowledge Transfer and Public Engagement.*” The paper discussed the important role of learned societies in the social sciences in facilitating dialogue between researchers and the public. The report also stressed the value of

mediators between the public and peer-reviewed research. Policy-makers and members of the public reported frustration with the lack of ability to find available social science experts from academia, to assist them in understanding and applying scholarly research outputs (ESRC & Academy of Social Sciences 2008, p. 23).

Mediation between the scholarly research system and the public can take a number of forms, including informal means like Web pages, blogs and other online sources of mediation. A number of websites and blogs have been created within the social sciences to provide a forum for discussion of current research within the social sciences. For example, the Social Science Research Council has created a website (<http://www.ssrc.org>) which serves to highlight important research developments in the social sciences. The website incorporates some web 2.0 features, including “The Immanent Frame,” a blog for the discussion of religion and the public sphere, another blog to discuss issues related to ongoing developments in Darfur, and a podcast that focuses on current political and social issues.

Additional efforts are being made open access to research in the social science through a number of learned societies and associations. For example, the American Sociological Association selects a number of articles in each issue of the publication *American Sociological Review* to make immediately open access online. These articles, available both at the journal website and on the American Sociological Association (ASA) website (<http://www.asanet.org>), are selected based upon an assessment of their potential contribution to a wider public discussion of important issues of general interest to members of the public.

Social scientists are actively engaged in the debate about open access, discussing the role of learned societies, exploring the possibilities of alternative publishing models, and promoting robust self-archiving schemes and new subject repositories, such as the recently launched anthropology repository, Mana'o (<http://manaoproject.org>). The current discussions surrounding open access within the social sciences are not merely concerned with the familiar issues of peer-reviewed journals, repositories and citation counts. Instead social scientists are thinking more broadly about the use of wikis, multimedia, open datasets, RSS feeds, blogs and other alternative forms of dissemination and debate as a counterbalance to the more static and unidirectional approach of the traditional scholarly journal (Kelty *et al.*, 2008).

### **Self-Archiving**

Often called the “green road” to open access, author self-archiving of peer reviewed articles on the free web has long been seen as a viable avenue towards increasing the percentage of peer-reviewed scholarship available to the general public. Self-archiving has been defined as the act of placing a digital copy of a scholarly article on the internet in order to make it freely available, preferably in a repository which is OAI-compliant (Budapest Open Access Initiative, 2002). Self-archiving has represented an alternate route to open access for authors who do not have the option publishing in a fully open-access peer-reviewed journal (the “gold road” to open access).

As self-archiving practice has evolved, it has become clear that attitudes and behaviors of authors with respect to self-archiving vary between disciplines. For example, some disciplines have established habits of self-archiving in dedicated subject

repositories such as arXiv (physics, mathematics and computer science) and RePEc (economics). The disparity in self-archiving rates between disciplines has been shown in a number of studies of self-archiving practices (Hajjem, Harnad & Gingras, 2005; Bergstrom & Levaty, 2007; Antelman, 2006).

Antelman (2006) reviewed self-archiving practices within six social science disciplines. In the study, economics articles showed a 60% rate of archiving, while other social science disciplines (political science, psychology, sociology, anthropology and geography) had self-archiving rates lower than 30%.

Bergstrom and Lavaty (2007) compared the self-archiving habits of economists and political scientists. The study determined that the self-archiving percentage of articles in economics journals was approximately 70%, while the self-archiving percentage of articles in political science journals was closer to 30%. Further analysis of possible variables within the sample determined that the strongest predictor of whether an economics article was self-archived was the citation influence per article of the journal in which the article was published. The conclusion was that articles published in higher impact journals were more likely to be self-archived.

Wren (2005) also found that an article was more likely to be self-archived if published in a high-impact journal. Wren's large-scale study looked at all articles published within a set of 13 biomedical journals over a 10-year period. The 13 journals were divided into roughly equal groups of journals with high, medium and low impact factors. The study determined that self-archived articles were strongly correlated with having been published in journals with the highest impact factors.

Wren suggested four possible reasons for the correlation between self-archiving and high-impact journals: 1) important papers published in high-impact journals create a demand for access which is met by the response of providing the paper online; 2) authors who publish in prestigious, highly cited journals are simultaneously trying to supplement citations by making their article openly accessible; or 3) self-archiving may arise from what the author deems the “trophy effect,” the posting of high-prestige publications by the author to demonstrate scholarly accomplishments.

The higher rate of self-archiving of articles published within high-impact journals may be explained in part by the fact that prolific authors within a field have developed consistent self-archiving habits in response to an expressed interest in their work by a wide network of associated researchers. Instead of responding individually to requests for reprints via telephone or email contact, these authors may have developed a habit of posting their papers on their webpages. The study discussed below seeks to explore the self-archiving behavior of individual authors who published in the highest impact-factor journals in their field, to determine if there are distinctions to be made in the self-archiving practices of this group.

### **Institutional and Subject Repositories**

Institutional or subject repositories are archives of digital scholarship, either comprised of scholarly output of a particular institution, or of scholarship from a particular discipline contributed by authors from a range of institutions. Proponents of the development of institutional and subject repositories have argued that author self-archiving within e-print repositories (green OA) will provide a swifter path to open

access than endeavoring to establish a large number of open access journals (gold OA) in every discipline (Harnad *et al.*, 2008). Gold OA proponents are working towards an OA ideal, where scholarly research is published in peer-reviewed open access journals supported by libraries, funding bodies, scholarly societies, and the cost savings of greater publishing efficiencies. Until that day, green OA proponents suggest working within the current system of publication, and trying to increase access to published works through self-archiving.

There are currently over 100 institutional repositories worldwide, and the number is growing (<http://roar.eprints.org>). However, despite their ease of use, and their potential to greatly increase the percentage of open access scholarship within an institution or a disciplinary field, institutional repositories have been slow to populate. Researchers who have studied factors contributing to the successful population of institutional repositories have constructed instructive studies of the self-archiving habits of scholarly authors. A number of relevant studies of IRs have determined that depositing rates in different disciplines show variation, indicating that disciplinary culture or habits are a factor in the self-archiving behavior of scholars (Xia, 2007). Others have disputed the disciplinary explanation, noting that success depends upon mandate policies and extensive librarian support to self-archiving authors (Haddow, 2008).

IR studies have been useful in investigating common barriers to author self-archiving, including the deterrents of time (Carr & Harnad, 2005), fear of violating publisher copyright policies (Coleman, 2006), and lack of awareness (Swan & Brown, 2005). The conclusion of many IR proponents is that the most promising way to

overcome all of the barriers to self-archiving with one action is through an institutional mandate, requiring deposit by all authors (Harnad *et al.* 2008; Xia, 2007).

Habits related to research and scholarly communication are created over time, and become ingrained and difficult to alter without significant disruption. Understanding scholarly habits is an important factor in determining how to get research into repositories. Foster and Gibbons (2005), an anthropologist/librarian team at the University of Rochester, undertook a work-practice study to try to understand why institutional repositories were failing to gather content, despite significant efforts. The researchers focused on how to get institutional repositories to fit into the current work of scholars, rather than the more common approach of asking how to get scholars to add their works to institutional repositories.

Foster and Gibbons used a range of data-gathering techniques, including videotaping sessions of faculty in their offices, watching faculty find and use articles and other research materials, and interviewing faculty through personal and telephone interviews. The study determined that faculty were uninspired by terms such as “institutional repository” and “metadata”. The authors of the study found that the objectives of faculty were clear and consistent. What scholars primarily wanted was to “do their research, read and write about it, share it with others, and keep up in their fields” (Foster & Gibbons, 2005).

The need to focus on the individual scholar, rather than on the institution, was also suggested by Gandel, Katz & Metros (2004). The authors speculated that the current failings in populating institutional repositories could be explained by a mismatch between the rigid schemes and unfamiliarity of institutional repositories and the more comfortable

individualized approach to self-archiving created by authors to address their own scholarly purposes. The authors proposed allowing scholars to continue to create “personal digital repositories” with meaning and context created by the scholars themselves. The authors suggested that meta-tools could be developed after the fact, to link content and loosely describe an institutional or organizational collection as a whole, while providing pointers or indexes for more specific areas.

A large number of the IR’s that have been created within universities have not created meaningful links between a scholar’s homepage and the repository. Creating a separation between a scholar’s works in an IR and additional information about the research found on the homepage is counterproductive. The separation of IR and scholar webpage is unmindful of the possibility that the scholar will be required to self-archive in two places, or give up one or the other.

The scholar’s web page primarily serves the individual practices and intentions of the scholar, while an institutional repository largely serves the needs of the institution in showcasing its collective scholarly output. Until there are additional incentives for scholars to self-archive both on personal homepages and in IR’s, scholars will most likely continue with their established habits of self-archiving - usually on personal Web pages or within popular subject repositories. Therefore, without mandates, or at least a lot more support for authors, the majority of IRs may continue to exist as little more than article indexes, largely bereft of full-text deposits.

## **Institutional and Funder Mandates**

*The quality of the information made to the public... is dependent upon the proportion of peer-reviewed research to which there is public access compared to the vast amount of other sources of information. (Willinsky, 2006, p. 114)*

Meaningful public access to peer-reviewed research is a game of percentages. The goal is to increase the amount of peer-reviewed research that is open access until it grows to the point where it can be found by the public as easily as the toll access databases, the pay-per-view sites, and the great mass of free “science” available on the web.

Simply encouraging authors to contribute their articles to open access journals or to self-archive has not yet created the hoped-for critical mass of peer-reviewed research available to the public. NIH discovered this predicament when their first open access policy for funded researchers was optional. The result, after a short test period, was that an extremely low percentage of authors voluntarily deposited their research. As a result, greater attention has now turned to the potential impact of mandatory OA policies on authors’ self-archiving behavior (*see Harnad et al.*, 2008).

In March, 2009, the National Institute of Health’s open access policy was made permanent. Previously, the NIH policy that mandating self-archiving author manuscripts in PubMed Central had been subject to annual renewals after it was signed in late 2007.

The current NIH policy regarding self-archiving reads as follows:

*All investigators funded by the NIH submit or have submitted for them to the National Library of Medicine’s PubMed Central an electronic version of their final, peer-reviewed manuscripts upon acceptance for publication to be made publicly available no later than 12 months after the official date of publication: Provided, That the NIH shall implement the public*

*access policy in a manner consistent with copyright law.* (NIH Open Access Policy, online)

Without mandates, individual author self-archiving behavior has tended to be unreliable. Having a decentralized array of papers located outside of OAI-compliant repositories results in a set of widely scattered deposits that prove difficult to index or harvest effectively. Funding organizations have not only the ability to mandate author self-archiving, but also the ability to dictate archiving in a centralized and well-indexed repository. By joining together and pointing all article depositors to one centralized location – PubMed Central, funders of medical research have determined a path of action which not only increases the percentage of open-access peer reviewed research, but also increases the findability and searchability of that research.

Publishers, to a certain extent, have deferred to the new OA policies of the large medical funders and have generally accommodated the funders' attempts to recapture research articles after short embargo periods. One publishing industry spokesperson, in responding to the tremendous leverage possessed by the funding organizations stated: "We must not underestimate the power of the funders. They...are excited by their new-found power to dominate the scholarly communication system, imposing terms and conditions on the allocation of research grants that predate any author contracts with publishers. The "pay-to-publish" model, supported by research funders, is here to stay" (Campbell, 2007). While publishers have been accommodating, they have not been silent in voicing their concern that the new funder OA policies may eventually disrupt subscription revenues and interfere with pay-per-view business models (Brussels Declaration, 2007).

Universities also play a central role in the creation of research; however, they have traditionally played a more limited role in the dissemination of research. Despite the fact that institutes of higher education collectively provide enormous sums of money to support research and scholarship, the university has long maintained a *laissez-faire* attitude toward the business of transferring ownership to the scholarship created by their faculty. Recently, however, much like the large research funders, a number of universities have decided to become more active participants in how knowledge, specifically research and scholarship created within their sphere of influence, is disseminated to the world.

In February 2008, Harvard University's Faculty of Arts & Sciences voted to create the first mandated open access policy in the United States. Following the Arts & Sciences faculty, Harvard's Law School passed a similar mandate in May 2008, followed by Stanford University's School of Education unanimously adopting an open access mandate in June, 2008.

The most recent OA mandate adopted within the United States is also the most significant. On March 20, 2009, the full faculty of MIT voted unanimously to approve a resolution to give MIT permission to disseminate the work of the faculty through the repository DSpace. MIT scholars may opt out of the archiving policy on a paper-by-paper basis if they are not able to negotiate publisher permission for archiving within DSpace. MIT is the first university to commit the whole faculty to an open access mandate. Until this point, other universities in the U.S. have committed departments to their mandates, but not the entire faculty. The growth in the number of institutional repositories, combined with the recently adopted institutional and funder OA mandates,

has the potential to greatly increase access to scholarly research that has long been unavailable to the public.

### **Scholars' Web Pages**

Academic scholars, working towards tenure and promotion in higher education, feel pressure to publish in peer-reviewed journals with high impact factors. Within tenure and promotion committees publication and citation are key. Less attention is paid to the rising cost of licensing the highly esteemed journals by the academic library, or to the lack of access to research published in high-impact subscription-only journals.

Almost all of the incentives for academic scholars point toward the goal of maximizing a list of published works, and maximizing the citations to those works. There is no surprise, then, that the typical webpage of an academic scholar includes name and email address, a small photograph, and a CV compiled of lists of publications, dutifully citing to the journals in which the works were published. Some pages have far more content; some have less. If the audience for a scholar's webpage is intended to be a hiring or tenure committee, a limited range of material should suffice. However, if a scholar envisions a wider audience of peers, local and global students, or the intellectually curious layperson, does the scholar create a different sort of web presence?

Willinsky (2006) places a good deal of the responsibility for opening access to research squarely upon scholars themselves. Willinsky states that the public's right to know carries with it, "the responsibility of researchers and scholars to ensure that there are no unwarranted impediments to the widest possible circulation of the ideas and the information with which they work." (Willinsky, p. 146.) The Wellcome Trust, as a

condition of funding, “expects authors of research papers to maximize the opportunities to make their results available for free” (Wellcome Trust, 2008, online).

In determining whether or how to self-archive their work, individual authors can easily find themselves at the intersection of several powerful forces: the traditional pressure to publish in important journals, the established machinery of scholarly publishing and copyright transfer, and the increasing, often influential, voices calling for free and open access to their scholarly work. Their behavior in response is, not surprisingly, unpredictable. The variety of content, or lack of content, posted on scholars’ homepages may attest to a lack of guidance and support for scholars trying to determine whether to have a public web presence at all, and if so, how to create it without running afoul of publishers and tenure committees.

In the midst of progress in the open access movement, the tenure process has remained fairly traditional. Few OA advocates have suggested that authors be expected to stop submitting their research to established, high-impact journals in lieu of newer, open-access journals whose position in the discipline has yet to be determined. Professional reputation, tenure and promotion are unavoidable concerns to scholars. One way that certain successful scholars have attempted to reconcile the conflicting incentives is to publish articles in prestigious, high-impact journals, then, with that goal accomplished, attempt to make the work more open access by self-archiving.

In the act of self-archiving their published articles, scholars have acted largely on their own, creating their own methods and practices. A number of studies have indicated that, as a general rule, scholars who do self-archive are most likely to self-archive their papers on their own webpage, rather than in an institutional or disciplinary repository

(Swan & Brown, 2005; Antelman, 2006). Many published scholars self-archive in multiple locations: on their own web page, as well as in subject repositories such as arXiv or RePEc. Scholars are additionally being asked or mandated to self-archive in Institutional Repositories, and in funder-determined locations, such as PubMed.

It is uncertain what effect mandatory policies that direct article deposit in IRs or funder designated repositories will have on maintenance of academic Web pages. The academic's personal homepage is the traditional location for article drafts, working papers, presentations, datasets, and other types of communication that explains and elaborates on the scholar's research. There have been a number of studies conducted that examined the potential of the scholar's personal homepage as a destination site for broader audiences. Barjak, Li and Thelwall (2007), performed a link analysis of the homepages of 456 scientists, analyzing country, academic discipline, and a number of other variables in order to determine which factors would predict a greater number of inlinks to a scholar's webpage. One finding of their study was that the most linked-to content on a scholar's home page is full-text, i.e. articles, papers, manuscripts, and conference presentations. The authors advise scholars who are interested in raising their online presence, and their visibility in result lists, is to include full-text content on their homepages to increase the number of inlinks to their pages. From an open access perspective, the visibility of a scholar's website in a Google results-list ranking can mean the difference between a searcher finding the author's self-archived article for free, or instead finding a list of pay-per-view sites offering the opportunity to access the article for an average of \$30.00.

There is great variety among web pages of individual scholars. In order to examine how scholars as individuals are represented on academic web pages, Thoms and Thelwall (2005) examined a sample of academic websites from countries within the European Union. The study identified the differences between academic websites that are university created versus those that are created by the scholar. The authors concluded that wholly university-created websites have the potential to render a faculty member a “nonentity” (p. 4), depriving the scholar of individuality and focusing primary attention on academic publication lists and little other content. “Non-entity” websites provided no opportunity for individualization or regular updating, and it can be surmised, no opportunity for self-archiving. Thoms and Thelwell found that even among academics who constructed their own web pages, the content of the self-constructed web pages tended to echo the priorities of the university. The study seemed to indicate that the structure and content of an academic’s webpage was influenced by university or department norms. If the construction of personal web pages is influenced by what is considered to be “typical” within the scholar’s university or department, then the creation of uniform web pages with no obvious page, tab, or location for a scholar’s self-archived papers could be a potent disincentive for the scholar to self-archive.

The variety of new institutional and funder mandates may potentially direct authors to archive in a number of separate locations, splitting the author’s work into locations lacking context personal to the scholar. However, the scholar’s established homepage, containing articles, working papers, datasets, field photographs, notes, blogs, RSS feeds, recent media coverage, and links to peers’ websites, still has the potential to operate as a hub, linking to the author’s articles in a variety of repositories. In this

manner, the author can maintain a centralized web presence to showcase all of his or her works, rather than having a career's worth of scholarship exist in parts scattered about different repositories and journals.

### **Publisher Self-Archiving Policies**

Authors who submit papers to a range of journals face an array of complex policies governing which rights they may retain, and how these rights might impact their ability to self-archive their published works. The SHERPA/RoMEO database is an online service that attempts to document the current self-archiving policies. The database (<http://www.sherpa.ac.uk/romeo.php>) is searchable by publication name and also by publisher to help authors quickly determine the policies of the journals prior to submission of articles. SHERPA encourages authors to use the database as a starting point, but to confirm the exact details of a particular journal's copyright transfer agreements at the publisher's website.

Among self-archiving authors, there are a variety of choices in determining which version of a paper to archive. Some authors consistently self-archive all papers upon completion. Other authors submit a paper to a journal, and then self-archive the submitted paper with a notation that the paper has been submitted for publication at a particular journal. These early versions of subsequently published papers are often called "preprints".

Other authors self-archive the version of the paper that has been peer-reviewed and "accepted" for publication, but has not undergone copyediting, pagination and typesetting by the publisher. This version is often referred to as the postprint, or

sometimes the “accepted manuscript”. Identifying the difference between a pre-print and a post-print is difficult just by looking at the paper. While the author will know whether or not the version has undergone peer-review and edits, it is often difficult to discern simply from the paper. Indications of the “peer-reviewed” postprint include a line in the first footnote thanking one or more referee(s), or a line of text inserted under the title of paper which says, “final version” or “accepted for publication in journal X” or “forthcoming in Journal Y.”

The definitions of journal article versions are in constant evolution. An effort to being to standardize the definitions of journal article versions was undertaken by the National Information Standards Organization (NISO) together with the Association of Learned and Professional Society Publishers (ALPSP). The joint working group recognized the following distinct versions of journal articles: Author’s original, Submitted Manuscript under Review, Accepted Manuscript, Proof, Version of Record, Corrected Version of Record, Enhanced Version of Record (NISO, 2008).

As a response to the funder policies mandating self-archiving within 6 or 12 months after publication, publishers have made certain accommodations and changes to their self-archiving policies. Some publishers have developed separate agreements for authors who are the recipients of funding. Other publishers have modified their embargo periods to match the terms of the funder mandate.

Many of the larger publishers have created author-pays programs that allow authors to make their article immediately open access. In some cases, using an “author pays” system is the only way that the author can comply with both a publisher’s self-archiving policy and a funder mandate. When funding is specifically conditioned upon

self-archiving the peer-reviewed version of research papers open access, the author can not enter into a copyright agreement with a publisher that does not allow the funding condition to be met. In securing the funds to pay the publishing costs, the author can look to either the grant funds, or to the author's institution. Author-pays involves the transfer of fees to the journal's publisher to accommodate the publisher's interests, yet does not eliminate the need for the author's institution to also pay for a subscription to the full journal. Some publishers give slight discounts to any author whose library subscribes to the journal. Table 1 shows examples of some of the author-pays fee arrangements.

Table 1: Publishers and Author-Pays Programs

Publisher	Acceptance Fee Subscriber	Acceptance Fee Non-Subscriber
<i>Cambridge Press</i>	\$2,700	\$2,700
<i>Elsevier</i>	\$3,000	\$3,000
<i>Oxford Journals</i>	\$1,500	\$2,800
<i>Routledge</i>	\$3,250	\$3,250
<i>Sage Publications</i>	\$3,000	\$3,000
<i>Springer</i>	\$3,000	\$3,000
<i>Wiley-Blackwell</i>	\$3,000	\$3,000
Average	\$3,045	\$3,175

Source: Sherpa/RoMEO, <http://www.sherpa.ac.uk/romeo/PaidOA.html>

Mergers and acquisitions within the publishing industry also bring about changes in self-archiving policies. In 2007, John Wiley & Co. and Blackwell Publishing merged. The merger created a company that publishes the journals of over 500 learned societies. The self-archiving policies of a publisher the size of Wiley-Blackwell have the potential to significantly impact scholars in a multitude of fields.

As institutional repositories and authors proceed with archiving versions of articles, publisher policies are a factor to be considered. Publisher policies that restrict

any deposit of the refereed version of the work draws a sharp distinction between the work of the author (the submitted manuscript), and the work of the journal's referees and editors (the peer-review process). Under a "preprint only" policy, the refereed article (postprint), having benefited from the insights and suggestions of the referees and editors, will not be available for self-archiving within an institutional repository or on an author's personal Web page.

Publishers may be comfortable with the current levels of self-archiving, because self-archiving has not yet achieved the critical mass of freely available articles that might affect subscription levels. Academic librarians will not be comfortable cancelling subscriptions with current self-archiving rates at an estimated 10-20% (*see Harnad et al.*, 2008). However, surveys have indicated that librarians are keeping an eye on the amount of content in open access repositories (Pinfield, 2007). Studies of the business models of publishers as they relate to open access repositories have shown that even well populated subject repositories such as arXiv have not produced significant decreases in subscriptions in relevant journals (Pinfield, 2007; Swan & Brown, 2005). However, it is prudent to assume that the largesse of publishers with regard to permissive self-archiving policies will begin to evaporate if journal or database cancellations begin to impact their business models.

From a self-archiving author's perspective, the variety of different policies can complicate the attempt to establish a coherent self-archiving practice for an individual scholar or for an institutional repository. There are stark differences in self-archiving policies contained within the universe of publisher agreements. The publisher agreements contain a range of permissions, prohibitions, embargo periods, and a rather

confusing array of article “version” definitions. It is possible that many authors simply ignore them.

Antelman (2006) determined that in a set of approximately 2,000 articles self-archived by authors within the social sciences, there was no discernable relationship between publisher self-archiving policy and self-archiving practice. In fact, Antelman’s study found some counterintuitive results. In a number of cases, the self-archiving rate for articles published within journals with more restrictive policies (SHERPA/White) showed a greater percentage than articles published within journals with more lenient self-archiving policies (SHERPA/Green). Ultimately, self-archiving practices appear to be more consistent within disciplines, and also tend to reflect university or department norms (Davis & Connolly, 2007), indicating that self-archiving practices are developed independently from publisher policies, and do not adjust to accommodate different levels of restrictions on self-archiving.

Scholars may publish in a number of journals, representing a range of publishers. For the scholar who wishes to develop a consistent practice of self-archiving his or her papers, and who simultaneously wishes to create a coherent look and feel to his or her Web page, the conflicting directives from publishers about how to self-archive can be daunting. For the younger, or untenured scholar, the issues involved in self-archiving may be even more intimidating. A Ph.D. candidate who has published an article and wishes to self-archive may lack a university-supported homepage or access to an institutional repository. Early career scholars may also fear violating publishing policies in a way that might deter future publishing opportunities.

If confusion and complexity surround the practice of self-archiving published articles, the likelihood of a “better safe than sorry” attitude towards self-archiving is more likely to result. In deciding whether, and how to self-archive a published work, the scholar is expected to walk a precarious line between the pressure to self-archive and eliminate barriers to open access, and the pressure to not self-archive if it could jeopardize his own career interests and the economic interests of the journal in which he published. Such conflicting incentives may explain much of the randomness in self-archiving habits among published authors, as discerned in the results of the study discussed below.

## **Methodology and Results**

### **Methods**

In determining the open access availability of a set of articles published in high-impact journals in the social sciences, the study follows two inquiries: one focusing on the OA status of the articles themselves, and the second focusing on the self-archiving behavior of the individual authors of the articles. The inquiry involves first searching for open access versions of peer-reviewed articles to determine how many may be located with common search engines, and second, determining the characteristics of the articles’ individual authors with regard to self-archiving practices. The present study addresses the following research questions.

- 1) What percent of articles published in a group of high-impact peer reviewed journals in the fields of anthropology, sociology, economics, and political science have been self-archived? What versions of articles were self-archived? Where do social scientists self-archive?

2) Are there common attributes among the authors in the study who were found to have self-archived, either by discipline, or by country or region affiliation?

The data consists of a group of scholarly articles published in 2007 and 2008 among a set of high-impact journals in the fields of Anthropology, Economics, Political Science, and Sociology. The journals in the study are the top 10 journals within each of the four social science fields as ranked by ISI impact-factor. The group of the top ranked journals was chosen specifically for two reasons: 1) to identify journals considered of general importance and broad interest in their field, and 2) to study the behavior of authors who have published articles in high-impact journals, a group that has been identified as more likely to self-archive than other authors (Wren, 2005; Bergstrom & Levaty, 2007). Selecting the 10 highest impact journals from the fields of anthropology, economics, political science, and sociology produced a group of 38 journals<sup>2</sup>

All articles published in regular issues of the journals during January-June 2007 and January-June 2008<sup>3</sup> produced the set of articles included in the study (n=1,686). Only full-length articles were included in the study. Book reviews, editorials, and other supplemental materials in the journals were not counted for the study, as these materials are less likely to be self-archived by authors.

The article title and the name of all of the article's authors were located using the Tables of Contents found at the various websites of the journals or the journals' publishers. Each article was searched to determine if it had been self-archived by the

---

<sup>2</sup> Two journals, *Global Networks* and *Social Networks*, were among the top 10 journals in two fields, so the number of unique journals reviewed was 38. Duplication was omitted in aggregations of total articles and total authors.

<sup>3</sup> Includes one July 2008 issue, as it was published 2 months late.

following authors: 1<sup>st</sup> author, 2<sup>nd</sup> author, 3<sup>rd</sup> author, and 4<sup>th</sup> or last author. For the articles with more than 4 authors, the last author's name was searched in lieu of the 4<sup>th</sup> author.

The initial search for an open access version of each article was by searching for the title and one or more authors' names in *Google Scholar*, using the Advanced Scholar Search functions. If the initial search did not find an open access version of the article, then the full title entered as a phrase in *Google*. Each author's institution was determined using the Table of Contents and abstracts found at the publisher or journal website.

By visiting an author's personal or departmental website, it was determined whether the article was self-archived there, or if there any were links to the article in a subject repository. If applicable, the 2<sup>nd</sup>, 3<sup>rd</sup>, and 4<sup>th</sup> or last authors were searched in the same manner as the first author to determine if any of the co-authors had archived the article on their personal webpage or within an IR or SR. If all previous searches had not produced a self-archived copy of an article, then a final search of the author's name was conducted using OAIster and OpenDOAR. OAIster is a searchable database that harvests article metadata from OAI-compliant open access repositories ([www.oaister.org](http://www.oaister.org)). OpenDOAR is a directory of academic open access repositories that uses a customized search engine from Google to index the contents of open access repositories ([www.opendoar.org](http://www.opendoar.org)). The various search methods used in the study were based on the findings and recommendations of Norris, Oppenheim, and Rowland who systematically compared performances for popular search engines in searching for open access articles (Norris, Oppenheim & Rowland, 2008a).

Additional data regarding the institutional affiliation of the author and the location by country and region of the author's institution, the journal title, the publication date of

the article, and the publisher's self-archiving policies were recorded. Adopting a method created for use in a series of studies of open access vs. toll access articles, each author was identified by country and region (North America, Europe, UK and Rest of World) in order to determine possible patterns of self-archiving behavior (Norris, Oppenheim, & Rowland, 2008b).

In contrast with a number of previous self-archiving studies, which used specialized web crawlers to search for self-archived versions of articles published in high impact journals, searches within this study were conducted using common search engines, using the types of searches which might be conducted by members of the general public trying to find a social science article which interested them. In addition, searching individual author's Web pages for each of the articles provided an opportunity to visit a multitude of sites, either created individually by scholars, or by a university or a department for their faculty, as well as the chance to review the usability of features of a large number of institutional repositories.

## **Results**

The results of searching for self-archived copies of articles in anthropology, economics, political science and sociology journals produced the following count: of a total of 1,636 articles in the study, 40.68% (n=686), were found to have been self-archived by at least one author, and 59.31% (n=1,000) were not found to have been self-archived.

Out of a set of 3,440 individual authors within the fields of anthropology, economics, political science and sociology, 28.8% of individual authors were found to

have self archived a copy of their article either on their personal web page or in an institutional or subject repository.

The results indicated similar rates of self-archiving between articles published in 2007 and articles published in 2008. Of the 831 articles published in 2007, 41.9% had been self-archived. Of the 855 articles published in 2008, 39.2% had been self-archived. Results for the four social science disciplines are discussed below.

## **Anthropology Results**

### **Journals and Impact Factor**

The 10 anthropology journals cover a broad range of topics including evolution, physical and cultural anthropology, local and global networks, and antiquities. The anthropology journals had an impact factor range of 3.138 to 1.533 (Table 2).

Table 2: Anthropology Journals and 2007 ISI Impact Factors

Title	Impact Factor
<i>Evolutionary Anthropology</i>	3.138
<i>Journal of Human Evolution</i>	2.712
<i>Current Anthropology</i>	2.312
<i>American Journal of Physical Anthropology</i>	2.273
<i>Global Networks</i>	1.886
<i>Annual Review of Anthropology</i>	1.883
<i>American Journal of Human Biology</i>	1.805
<i>Social Networks</i>	1.644
<i>Cultural Anthropology</i>	1.636
<i>American Antiquity</i>	1.533

### **Article Self-Archiving Rate**

The total number of anthropology articles included in the study was 529, with 260 articles published in the first half of 2007, and 269 articles were published in the first half of 2008. Out of the 529 total articles, 136 articles (25.52%) were found to have been self-archived by one or more of the contributing authors, while for 393 (74.48%) of the articles, no self-archived copy could be located.

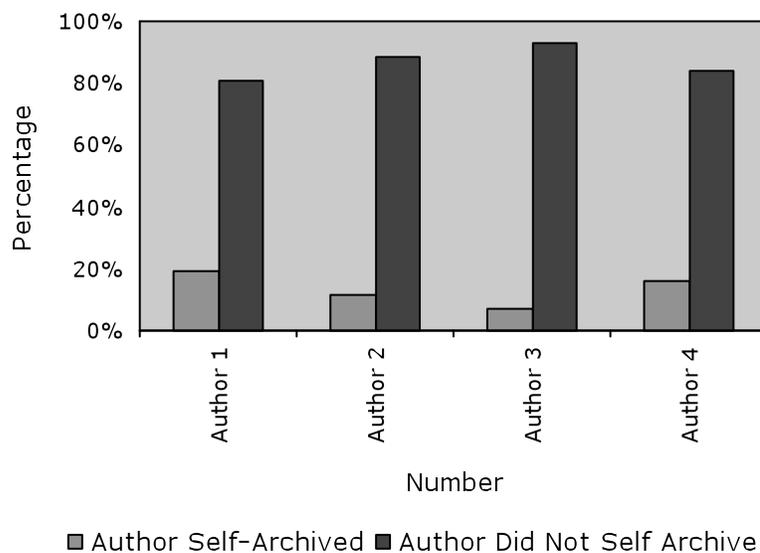
### **Anthropology Authors**

The 529 anthropology articles were written by a total of 1,457 authors. The average number of authors per article was 2.79, the mode was 1, and the median was 7. The study selected the 1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup>, and 4<sup>th</sup> or last author of each article as a subset, to determine whether each of the individual authors had self-archived their article on a personal webpage or in an IR or subject repository. Identifying the subset of all 1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup>, and last authors produced a total number of 1,202 authors who were included in the study. The set of anthropology articles included a large set of articles (27%) having more than 4 authors, with a group of over 200 authors who were listed as 5<sup>th</sup> author or higher.

Within the group of 1202 anthropology authors, 176 authors (14.64%) had self-archived a version of their published article, while the remaining 1,026 (85.36%) authors had not self-archived a version of their article that could be located online.

Within the group of 1<sup>st</sup> authors, 19.1% (n=101) had self-archived their article; of 2<sup>nd</sup> authors, 11.3% (n=38) had self-archived. Within the group of 3<sup>rd</sup> authors, 7.3% (n=15) had self-archived their article, and of 4<sup>th</sup> or last authors, 16.3% (n=21) had self-archived a version of their article (Figure 1).

Figure 1: Self-Archiving Percentage by Author Number



The authors included anthropologists from over 50 countries, including many local anthropologists working along with North American and European anthropologists at internationally important sites.<sup>4</sup>

Examining anthropology authors on a regional basis shows that majority (45.28%) of authors come from institutions in North America. The next largest group of authors came from Europe (27.61%). Authors from the United Kingdom comprised 10.74 % of total authors, and 16.36% of the authors were from the combined group of countries from the rest of the world (Table 3).

<sup>4</sup> Local anthropologists included in the set of authors came from Bolivia, Botswana, Costa Rica, Guyana, India, Iran, Pakistan, Peru, Myanmar, South Africa, and Tanzania, for example.

Table 3: Percent of Anthropology Authors by Region

North America 45.28%	Europe 27.61%	United Kingdom 10.74%	Rest of World 16.36%
-------------------------	------------------	--------------------------	-------------------------

### Article Versions Self-Archived

Of the authors who self-archived, anthropology authors posted articles most frequently in the publisher's pdf version. Self-archiving of earlier versions, preprints or postprints, was far less common (Table 4). Of the anthropology articles found to be self-archived, 87.5% were in publisher's pdf format, 9.7% were a preprint version of the published article, and only 2.84% were a postprint or accepted manuscript version of the published article.

Table 4: Anthropology Self-Archiving by Version of Article

Preprint 2.84%	Postprint 9.7%	Publisher's PDF 87.5%
-------------------	-------------------	--------------------------

The anthropologists in the study self-archived most frequently on their personal or departmental Web page. Fewer of the anthropology articles were self-archived in institutional repositories. There were a number of articles within anthropology that were self-archived in PubMed Central. Archiving within PubMed is a natural fit for a number of anthropology journals with a strong focus on biology and health, including the *American Journal of Physical Anthropology*, and the *American Journal of Human Biology*.

## Economics Results

### Journals and Impact Factor

The group of economics journals included coverage of a wide range of issues related to economic theory and practice, including market theory, fiscal policy, finance, accounting, statistics, international trade and development, and political economics.

Table 5 shows the titles and the impact factors for the 10 economics journals.

Table 5: Economics Journals and 2007 ISI Impact Factors

Title	Impact Factor
<i>Journal of Political Economics</i>	4.190
<i>Journal of Economic Literature</i>	3.973
<i>Quarterly Journal of Economics</i>	3.688
<i>Journal of Accounting &amp; Economics</i>	3.034
<i>Journal of Financial Economics</i>	2.988
<i>Econometrica</i>	2.972
<i>Journal of Economic Perspectives</i>	2.831
<i>Journal of Economic Geography</i>	2.679
<i>Review of Economics Studies</i>	2.539
<i>Journal of Economic Growth</i>	2.292

### Article Self-Archiving Rate

There were 410 economics articles included in the study, with 202 articles published in the first half of 2007, and 208 articles published in the first half of 2008. Of the articles published in 2007, 144 (71.29%) had been self-archived. Of the articles published in 2008, 159 (76.44%) of the articles had been self-archived. For the combined 2007 and 2008 articles, 73.9% were found to be self-archived by one or more of the authors.

### **Economics Authors**

Within economics, there were 410 articles written by 846 authors. The average number of authors per economics article was 2.06, with a mode of 2 and a median of 4. Selecting the 1<sup>st</sup>-4<sup>th</sup> authors for each article produced a subset of 839 authors who were included in the study.

As indicated in previous studies, economists show higher rates of self-archiving relative to other disciplines within the social sciences. Of the 839 economics authors in the study, 58% (n=487) of all authors included in the study were found to have self-archived a version of their published article, many in more than one location - while the remaining 352 authors (41.95%) had not self-archived their article.

Within the group of 410 authors listed as first or sole author on each of the economics articles, 60% (n=246) had self-archived their article, either on a personal Web page, or within a subject repository, primarily within RePEc and SSRN. Of the group of second authors, 57.49% had self-archived (n=167) their article. Of 3<sup>rd</sup> authors, 52.89% (n=64) had self-archived, and of the group of 4<sup>th</sup> or last authors, 50% (n=10) had self-archived their article.

Examining the numbers of economics authors from each region showed that North American authors comprised a large majority of the total authors (Table 6). North American authors comprised 74.38% of the total economics authors. Europe had the second highest group of authors with 12.6%, authors from the United Kingdom comprised 7.9% of all authors, and authors from countries in the rest of the world made up only 5.15%. The disparity is largely due to the fact that the many of the high-impact

economics journals selected for the study are published in the U.S. and concerned with U.S. economics.

Table 6: Percent of Economics Authors by Region

North America	Europe	United Kingdom	Rest of World
74.38%	12.6%	7.9%	5.15%

### Article Versions Self-Archived

Economics authors made far more extensive use of preprints and postprints than authors from the other three disciplines. Of economics authors who were found to have self-archived their article, 26.3% percent had self-archived a preprint of the article, 41.1% had self-archived a postprint, and 32.65% had self-archived the publisher's pdf version of their article (Table 7).

Table 7: Economics Self-Archiving by Version of Article

Preprint	Postprint	Publisher's PDF
26.3%	41.1%	32.65%

Economists self-archived more frequently than other authors. Because of the greater number of examples of self-archiving, it was interesting to see the variety of different habits of the economists. Articles with multiple authors showed a variety of versions of the same paper. For example, one author would self-archive a preprint, another a postprint and a third a publisher's pdf. Most authors posted their articles on their home pages, often in hyperlinked lists of working papers. Many articles were found

archived on SSRN with links from the author's web page. Economists tended to have a helpful practice of hyper linking to their co-author's web page.

## **Political Science Results**

### **Journals and Impact Factor**

The 10 political science journals cover a broad range of issues including the American and European Union political systems, public opinion, voting behavior, confliction resolution, and geopolitical issues. Table 8 shows the political science journals by title and impact factor.

Table 8: Political Science Journals and 2007 ISI Impact Factors

Title	Impact Factor
<i>Political Analysis</i>	2.535
<i>American Political Science Review</i>	2.317
<i>American Journal of Political Science</i>	2.032
<i>Public Opinion Quarterly</i>	2.030
<i>Journal of Conflict Resolution</i>	2.030
<i>Political Geography</i>	1.922
<i>European Journal of Political Research</i>	1.679
<i>European Union Politics</i>	1.651
<i>Quarterly Journal of Political Science</i>	1.636
<i>Journal of Politics</i>	1.456

### **Article Self-Archiving Rate**

The total number of political science articles in the study was 385. Of the total articles, 165 articles (42.86%) were found to have been self-archived by one or more of the contributing authors, and 220 (57.14%) were not found to have been self-archived.

### Political Science Authors

Within political science, there were 385 articles written by 717 authors. The average number of authors per political science article was 1.86, the mode was 1, and the median was 3.5. Selecting the authors who were listed in the 1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup>, and 4<sup>th</sup> or last position for each article produced a set of 705 authors who were searched individually to determine whether the author had self-archived.

A total of 227 political science authors (32.19%) had self-archived a version of their published article, while the remaining 478 authors (67.8%) had not self-archived a version of their article.

Of the political science authors who were identified as first or only author, 34.28% (n=132) had self-archived their article. Of the group of 2<sup>nd</sup> authors, 31.1% (n=69) had self-archived, of 3<sup>rd</sup> authors 26% (n=19) had self-archived their article, and of 4<sup>th</sup> or last authors, 28% (n=7) had self-archived. There was not enough difference between the self-archiving rates of 1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup> and 4<sup>th</sup> authors to show significance.

Table 9 reflects the percentages of political science authors by region. The majority of political science authors were from North America (72.5%), again reflecting a large percentage of American authors. European authors made up 13.3% of the total, reflecting the presence of two journals focused on European politics. United Kingdom authors comprised 9.8% of the total, and authors from countries in the rest of the world made up 4.4% of the total authors.

Table 9: Percent of Political Science Authors by Region

North America	Europe	United Kingdom	Rest of World
72.5%	13.3%	9.8%	4.4%

### Article Versions Self-Archived

Of the political science authors who were found to have self archived, 30.8% had self archived a pre-print of the article, 11.61% had self-archived a postprint, and 57.59% had self-archived the publisher's pdf version of the article (Table 10).

Table 10: Political Science Self-Archiving by Version of Article

Preprint	Postprint	Publisher's PDF
30.8%	11.61%	57.59%

The majority of the self-archived articles were located on the author's personal or departmental Web page, although within Europe and Australia, a fair number of articles had been self-archived in institutional repositories. Within the group of political science authors the use of subject repositories was sparse, although a handful of papers were located in SSRN and arXiv.

### Sociology Results

#### Journals and Impact Factor

Table 11 shows the sociology journals and their 2007 impact factors. Many of the journals dealt broadly with the topic of sociology, two focused primarily on populations within the U.S. and one focused largely on populations in the UK. Other sociology journals were more narrowly focused on health issues, family issues, and social networks. The 10 sociology journals included in the study had an impact factor range of 3.338 to 1.577.

Table 11: Sociology Journals and 2007 ISI Impact Factors

Title	Impact Factor
<i>American Journal of Sociology</i>	3.338
<i>American Sociology Review</i>	3.277
<i>British Journal of Sociology</i>	2.449
<i>Annual Review of Sociology</i>	2.400
<i>Global Networks</i>	1.886
<i>Sociology of Health &amp; Illness</i>	1.759
<i>Journal of Marriage and Family</i>	1.756
<i>Economy and Society</i>	1.678
<i>Social Networks</i>	1.644
<i>Social Problems</i>	1.577

### Article Self-Archiving Rate

Of the 418 sociology articles within the study, 95 articles (22.73%) were found to have been self-archived by one or more of authors, while 323 (77.27%) of the articles there was no self-archived copy that could be located with any of the search methods. Of the total articles, 210 were published in the first half of 2007, and 208 articles were published in the first half of 2008.

### Sociology Authors

Within sociology, there were 418 articles written by 822 authors. The average number of authors per sociology article was 1.96, the mode was 1, and the median was 4. After selecting the 1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup> and 4<sup>th</sup> or last authors for each article, there were 799 authors who were included in the search of individual authors.

The sociology authors in the study represented institutions in 28 separate countries.<sup>5</sup> Among the group of sociology authors, 54% (n=432) were from the United States. On a regional basis, the majority of sociology authors came from North America (60.7%). Authors from the UK represented 21.8% of all authors. Authors from Europe represented 11.3% of the total, and authors from all other countries represented 6.16% of the total sociology authors. Table 12 shows the percentage of sociology authors by region.

Table 12: Percent of Sociology Authors by Region

North America	Europe	United Kingdom	Rest of World
60.7%	11.3%	21.8%	6.16%

As reported by other studies (Norris, Oppenheim & Rowland, 2008b; Antelman 2006), sociology authors showed a relatively low rate of self-archiving. Of the total number of authors, 129 sociology authors (16.15%) had self-archived a version of their published article, while the remaining 670 authors (83.85%) had not self-archived.

Of sociology authors who were listed as 1<sup>st</sup> author, 19.14% (n=80) had self-archived their article. For 2<sup>nd</sup> authors, 3<sup>rd</sup> authors, and 4<sup>th</sup> (or last) authors, the self-archiving numbers were 15.52% (n=36), 9.62% (n=10), and 6.7% (n=3), respectively.

Although the differences in these self-archiving percentages are small, they are

<sup>5</sup> Sociology authors in the study were from the following countries: Africa, Argentina, Australia, Canada, Denmark, Egypt, Estonia, Finland, France, Germany, Ireland, Israel, Italy, Korea, Mexico, The Netherlands, Norway, New Zealand, Pakistan, Singapore, Slovenia, Spain, Sweden, Switzerland, Turkey, UK, U.S., Vietnam.

nonetheless significant ( $p < 0.028\%$ ) and show an association between author position and self-archiving rates within sociology.

### Article Versions Self-Archived

Sociologists most frequently self-archived the publisher's pdf version of their articles (57.36%). Sociologists used fewer preprints (25.58%) and postprints (17.05%). Table 13 shows the percentage of articles self-archived by version. Preprint versions were more commonly self-archived by authors from Europe or the United Kingdom. Many of the UK and European authors had self-archived their preprint in an institutional repository. When postprints versions of articles were found, they were often found in PubMed Central as a deposited author's manuscript, after being published in a journal with medical or health related issues, such as *Sociology of Health & Illness*.

Table 13: Sociology Self-Archiving by Version of Article

Preprint	Postprint	Publisher's PDF
25.58%	17.05%	57.36%

The results reveal distinct disciplinary differences in self-archiving patterns, both in the total percentage of articles self-archived, and in the individual self-archiving patterns of authors within each discipline. In addition, disciplines show distinct habits in the version of the article the authors tend to self-archive the most frequently, and whether authors are most likely to self-archive on their own Web pages, or in subject or institutional repositories.

The study also analyzed authors by institution and department to determine if self-archiving habits were associated with groups of scholars in the same setting. Most institutions had small numbers of authors in the study due to the short window of publication dates in the study (12 months) combined with the limited number of journals per field. Therefore, statistically significant results showing reliable differences in self-archiving rates for individual institutions or departments could not be achieved. However, a few institutions in a number of fields produced far higher than average results in self-archiving rate. Those results are qualitatively analyzed in the discussion section.

### **Publisher Self-Archiving Policies**

To demonstrate the range of self-archiving policies among the group of journals in the study, the tables in the Appendix combine the publisher self-archiving policies of the 38 journals in the study. The policies are divided into in three tables, showing how different policies handle the self-archiving of preprints, postprints, and the publisher's pdf. Each self-archiving policy was located on the publisher or journal websites, or was located within the SHERPA/RoMEO database.

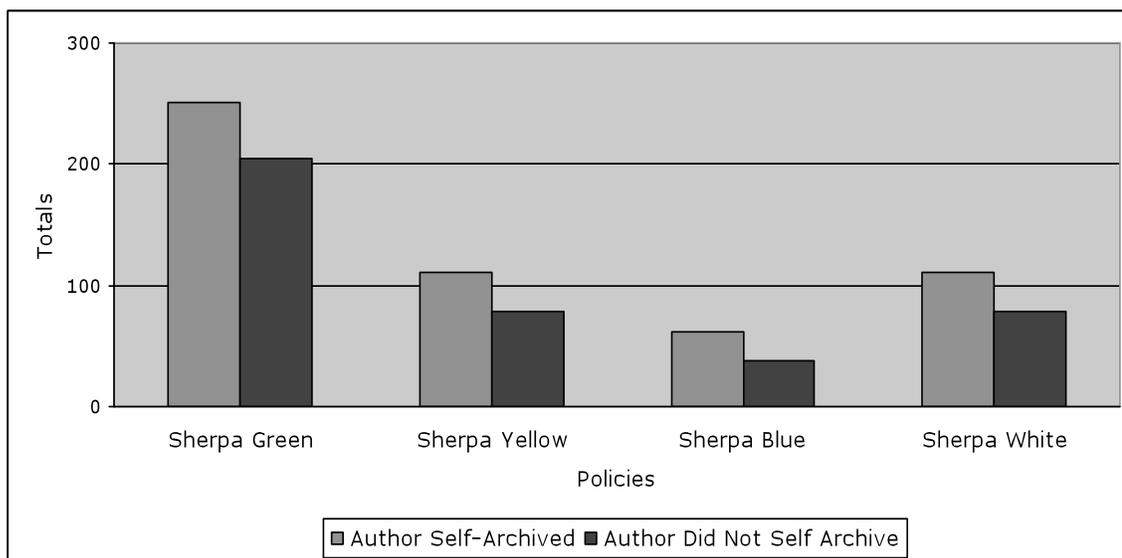
The SHERPA/RoMEO database functions by drawing distinctions between publishers and journals in terms of what level of self-archiving permission they allow. The "color" designations are general categories, as two journals of the same color may handle the precise permissions or instructions regarding preprints, postprints or publisher's pdfs in different ways. The SHERPA colors roughly place journals in a hierarchy from the most permissive (SHERPA/green), to the most restrictive

(SHERPA/white). Publisher policies within each of the four disciplines are discussed individually below.

### Economics

Five of the 10 journals are SHERPA/Green (preprint allowed/postprint allowed), three are SHERPA/Yellow (preprint allowed/postprint allowed after embargo), one is SHERPA/Blue (preprint not allowed/postprint allowed); and one was SHERPA/White (preprint restricted/postprint not allowed). Five of the economics journals expressly disallow use of the publisher's pdf version of the article; five journals either allow the publisher pdf use or are silent. Figure 2 indicates that regardless of the publisher policy, the self-archiving rates of the individual economics authors remained fairly consistent.

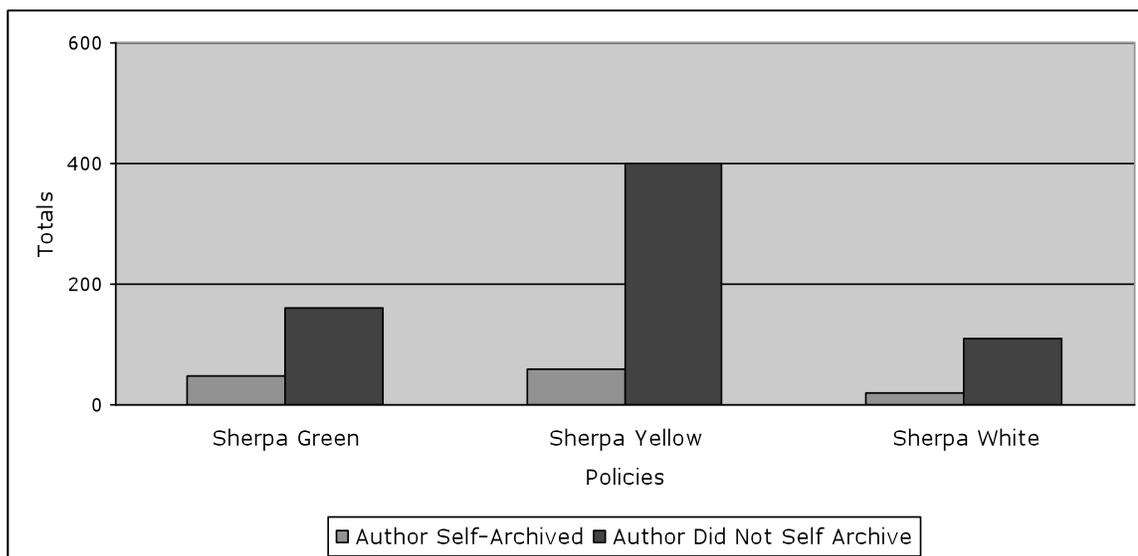
Figure 2: Self-Archiving within Different Publisher Policies (Economics)



## Sociology

Three of the 10 sociology journals have policies described as SHERPA/Green (preprint allowed/postprint allowed), five are SHERPA/Yellow (preprint allowed/postprint allowed after embargo), and 2 are SHERPA/White (preprint restricted/postprint not allowed). Of the 10 sociology journals, 8 journals expressly disallow use of the publisher's pdf version, and 2 journals allow use of the publisher's pdf after embargo. Figure 3 shows self-archiving rates of sociology authors within different publisher's policies. The data indicates that there are a greater percentage of non self-archiving authors who published articles in SHERPA/Yellow journals than who published in journals with SHERPA/Green policies. This does not necessarily indicate that the policies are the reason for the difference in rates, only that the difference exists.

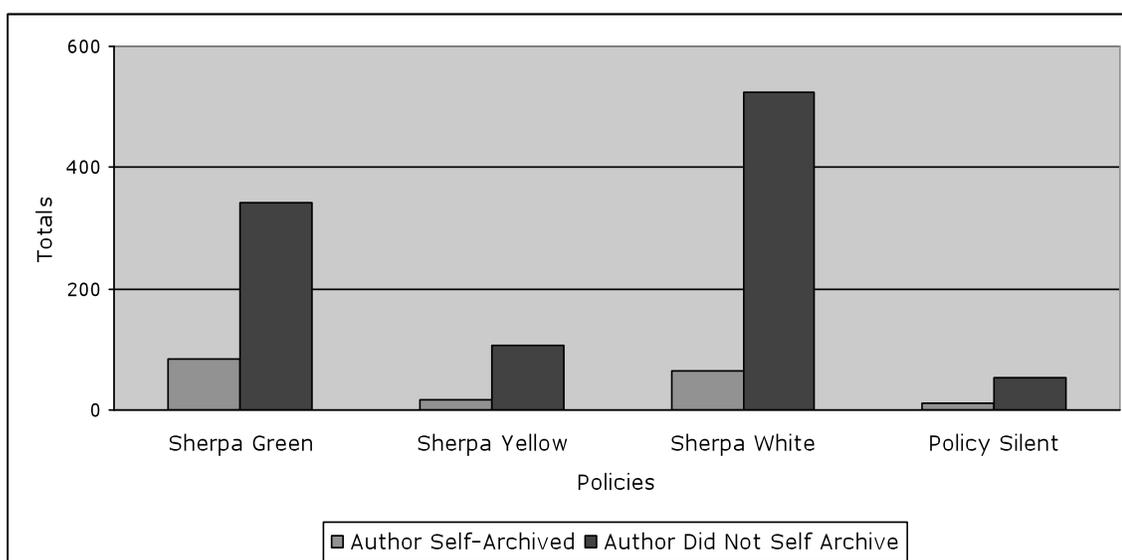
Figure 3: Self-Archiving within Different Publisher Policies (Sociology)



## Anthropology

Four of the 10 anthropology journals have policies described as SHERPA/Green, 2 are SHERPA/Yellow, and three are SHERPA/White. One journal's policy could not be located on the journal site or on the SHERPA/RoMEO database. Seven of the 10 anthropology journals expressly disallow any use of the pdf; 2 journals allow use of the publisher's pdf (one after embargo; one immediately). One journal, *American Antiquity*, did not have a policy on pdf use that could be located; however, the journal is not an electronic journal, so no electronic versions of the published articles are widely available. Figure 4 indicates that anthropologists who did not self-archive were more likely to have published their articles in SHERPA/White journals (the most restrictive SA policies). However, the self-archiving rate of authors within SHERPA/Green journals (80.32% did not self archive) and SHERPA/White Journals (88.9% did not self archive) is not significantly large enough to suggest that the policies have a marked effect on self-archiving behavior.

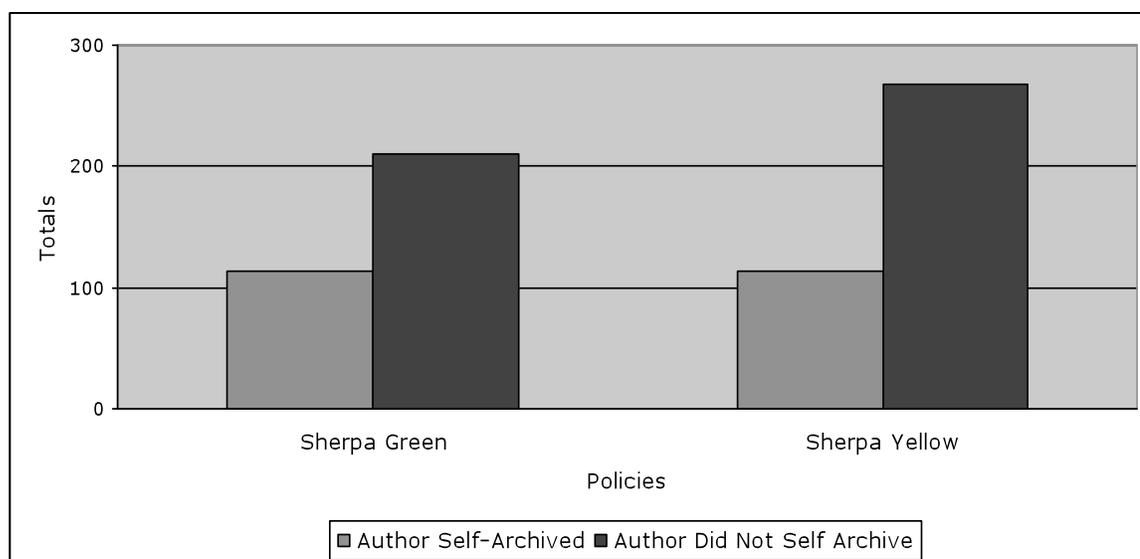
Figure 4: Self-Archiving within Different Publisher Policies (Anthropology)



## Political Science

Four of the political science journals have SHERPA/Green self-archiving policies and six journals have SHERPA/Yellow policies. The six SHERPA/Yellow journals all contain embargo periods: three for 12 months, three for 24 months. No political science journals in the study were identified as SHERPA/White, so that there are no complete prohibitions against self-archiving in this group of journals. Three of the journals allow self-archiving of the publisher's pdf version of the final paper, while seven journals specifically prohibit posting of the publisher's pdf version. Figure 5 shows that political science authors had only two different policies: SHERPA/Green and SHERPA/Yellow. There was a slightly greater tendency for authors in SHERPA/Green journals to self-archive (35.2% did self-archive) than authors in SHERPA/Yellow journals (29.7% did self-archive).

Figure 5: Self-Archiving within Different Publisher Policies (Political Science)



Of the four disciplines, the publisher's policies in economics and political science were more permissive. Political science had 10 journals that were either SHERPA/Green or Yellow. Economics had 8 policies that were either SHERPA/Green or Yellow. The majority of economics authors (54.3%) published in SHERPA/Green journals. Half of the economics journals allowed the use of the publisher's pdf, a higher percentage than any of the other disciplines, also suggesting more permissive policies among high-impact economics journals.

Overall, journals in sociology and anthropology had more restrictive policies. The largest group of authors in anthropology (49.0%) published in journals with a SHERPA/white policy. The majority of sociology authors (57.5%) published in journals with SHERPA/yellow policies, and 16.4% of the authors published in SHERPA/white journals.

## **Discussion and Conclusion**

Social scientists are considering ways to increase interaction with the public and share the results of scholarship they have created. One way for social scientists to share the results of their research and insights with the public is by self-archiving of peer-reviewed articles to the fullest extent allowed by copyright policies. With that goal in mind, it is helpful to assess the current state of self-archiving practices of social scientists in order to determine how self-archiving rates might be improved.

The overall rates of self-archiving found were consistent with similar studies of self-archiving within various social science disciplines. Specifically the self-archiving rates discovered for articles from anthropology, economics, political science, and

sociology were consistent with those found by Antelman (2006), Bergstrom and Levaty (2007), and Norris, Oppenheim & Rowland (2008b).

Previous studies of self-archiving rates have typically focused on the rate of self-archiving per article, as this was an efficient way of determining what percent of the peer-reviewed scholarship within a discipline was open access. This study was interested in determining how individual authors had self-archived their article, even if there were multiple authors.

The data from this study shows that the percentage of open access articles is higher than the percent of individual authors that self-archive. Within the various disciplines, this disparity is shown in Table 14. The reason for the disparity is that it only requires self-archiving by one author make an article open access. Articles published by more than one author designate a corresponding author. At a minimum, if all groups of co-authors designated a “self-archiving” author, then all articles would be open access to the fullest extent allowed by publisher policy.

Table 14: Percent of Self-Archived Articles vs. % of Authors Who Self-Archive

Discipline	Articles self-archived (% of total articles)	Authors who self-archived (% of total authors)
Anthropology	25.52%	14.6%
Economics	73.41%	58.0%
Political Science	57.14%	32.2%
Sociology	22.25%	16.2%

## **Regional Characteristics**

When examined by regional categories, the majority of authors in the study were from North America (61.24%). Authors from Europe comprised 17.38% of the total authors. Authors from the United Kingdom comprised 12.43% of total authors, and the authors from all other countries (Rest of the World) represented 8.94% of total authors.

The highest percentage of self-archiving authors came from North America (35.3%). Within each discipline, having a large percentage of North American authors seems to produce a higher overall self-archiving rate. The result can be seen in Figure 6 within the smaller economics figure. In this case, the large percentage of self-archiving authors from North America produced a high percentage of OA articles overall, counterbalancing the fact that the rest of the regions (Europe, UK and Rest of the World) self-archived at far lower percentages. The presence of a majority of North American authors had the effect of raising the percent of overall self-archiving within all disciplines; however, even among North American authors, strong disciplinary differences were evident in self-archiving rates, as seen in Figure 6.

Figure 6: Self-Archiving by Region and Subject

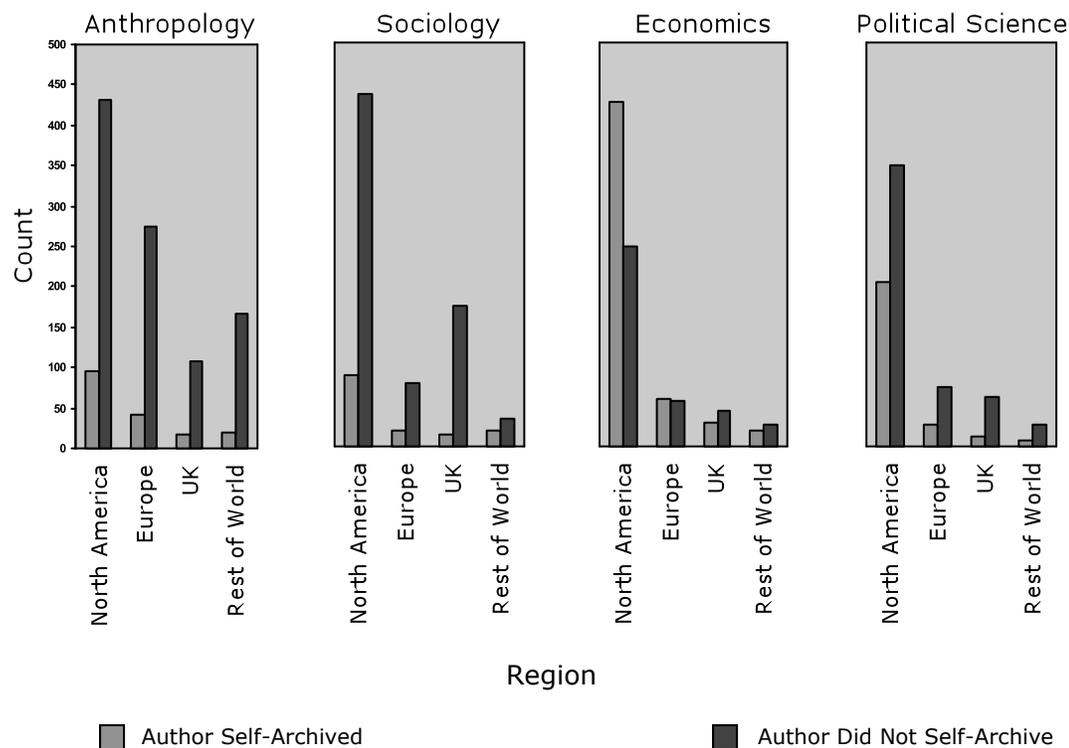
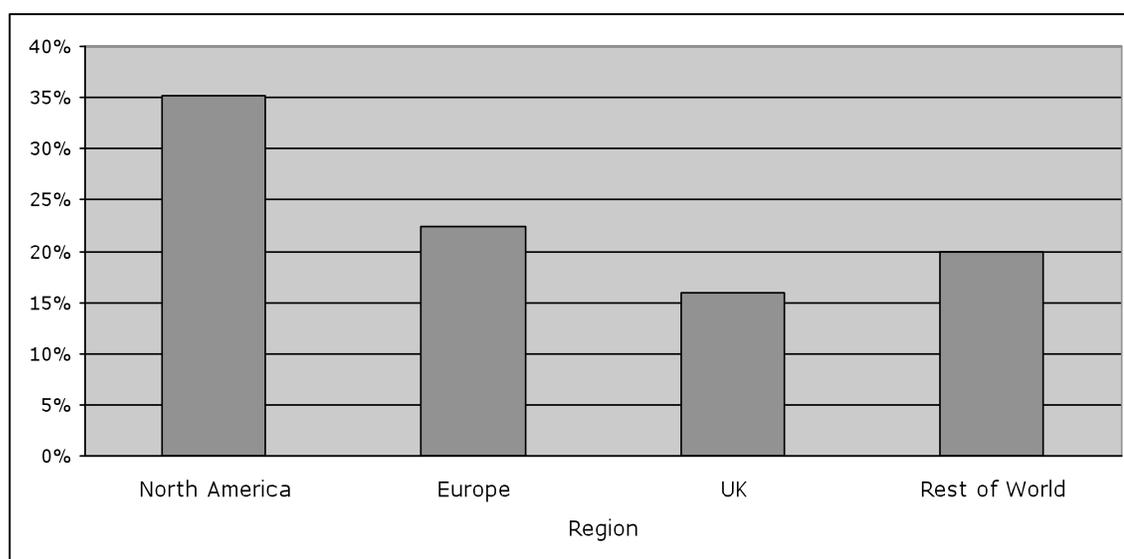


Figure 7 compares the self-archiving rates of authors by region, without reference to discipline. North American has the highest percentage of self-archiving authors (35.3%), followed by Europe (22.5%) and the countries of the Rest of the World (19.4%). The United Kingdom authors had the lowest percentage of self-archiving (15.96%), a result also found by Norris, Oppenheim & Rowland (2008b). However, there are a large number of fully functional institutional repositories at universities in the UK. While many articles are already fully indexed in the UK repositories, they currently point to the DOI at the publisher's website, and no open access version of the article has been uploaded. Self-archiving mandates are proliferating in Europe and the UK, and as many publisher self-archiving embargos expire (usually 12-24 months after publication for

Social Science journals), institutional repositories in the UK and Europe are poised to rapidly increase the self-archiving rates of these indexed articles.

Figure 7: Percentage of Authors Self-Archiving by Region



### Author Order

The association between author number and self-archiving rate showed mixed results. For economics and political science, two fields with the highest percentage of author self-archiving, there was not a statistically significant difference in self-archiving rates based on author order, in that all authors self-archived at fairly comparable rates.

Within sociology and anthropology, the difference in self-archiving rates and the author's order (1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup> or 4<sup>th</sup>/last) did show differences that were significant (Sociology:  $p < 0.028$ ; Anthropology:  $p < .0001$ ). For anthropology, the role of the 4<sup>th</sup> or last author was unique within the study, in that it was the only author position that increased in the likelihood to self-archive from the previous author position (3<sup>rd</sup>). Author

last position as it relates to determining primary credit for an article or which author served in a supervisory position is difficult to determine by looking at the byline, and author position as it relates to credit or importance may be influenced by disciplinary culture (Wren *et al.*, 2007).

For many anthropology articles, the last author (especially at the end of a large group) may have acted in a supervisory position. Many of the anthropology articles consisted of long lists of “co-authors” who worked on locations and sites in remote areas. There may be an understanding among these authors that the supervisory or corresponding author is in a better position to self-archive than other co-authors. This may explain the higher rates of self-archiving for first and last authors in anthropology. Within sociology, the rate of self-archiving decreased steadily between subsequent authors, to a significant degree. This factor might indicate a perception of greater credit for authorship for 1<sup>st</sup> and 2<sup>nd</sup> authors within sociology, leading subsequent position authors to leave self-archiving decisions to authors with higher positions.

Single authors have been associated with lower percentages of self-archiving, which makes sense from the perspective of simply increasing the odds that someone will self-archive. Norris, Oppenheim, & Rowland (2008b) determined that high numbers of non-archiving single authors within sociology had contributed to an overall low percentage of open access articles within their study. Given the lower percentages of self-archiving in sociology and anthropology, groups of co-authors might discuss among themselves which of the authors will self-archive the article, if they are not all able to do so. The self-archiving responsibility could be given to whichever author was most

comfortable with the technology, perhaps giving self-archiving responsibility to more junior authors.

### **Publisher Self-Archiving Policies**

Publisher self-archiving policies varied widely, even within this small group of high-impact journals. Publishers are still working out the mechanics of how they will comply with the new funder self-archiving mandates, and this factor is contributing to more changes in self-archiving policies.

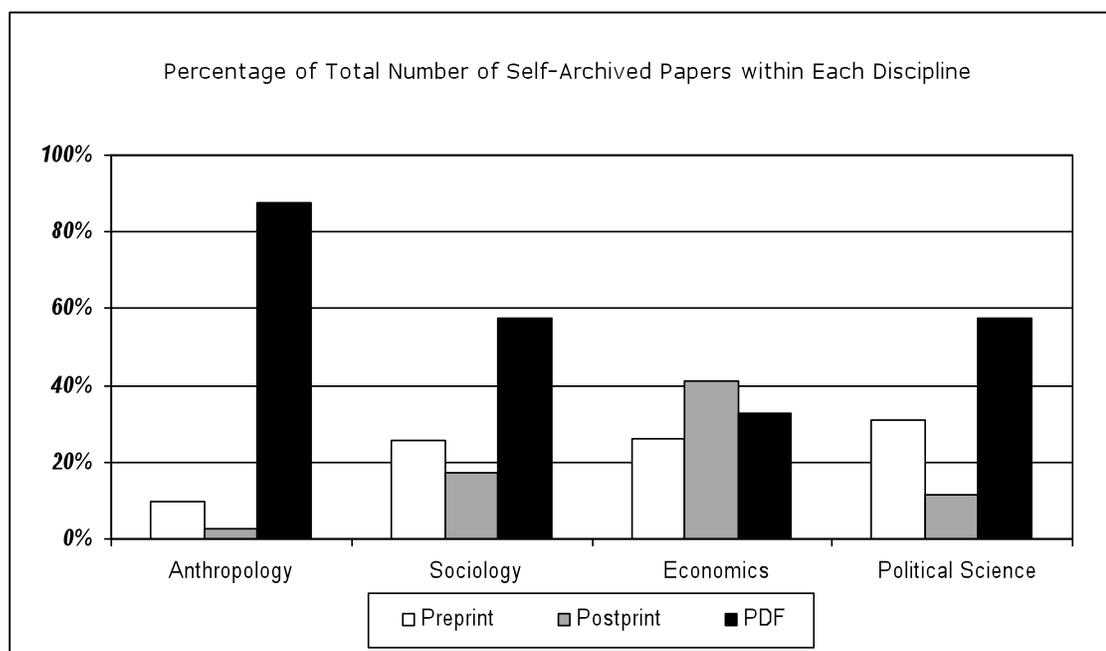
Thirteen (34%) of the 38 journals in the study either expressly allowed posting of the publisher's pdf version of the article or were silent, and 25 journals (66%) explicitly prohibited the self-archiving of the publisher's pdf version of the article. It was not possible to determine if any authors in the study negotiated for special permissions outside of the standard policy, but this is assumed to be a rather low number.

Is author self-archiving of a publisher's pdf more likely if the publisher's policy permits it? The data showed that there was a statistically significant difference in the self-archiving rates of pdf's by authors depending upon the publisher's policy (Author self-archived pdf (yes/no) by Publisher policy permits pdf (yes/no)  $p < .0001$ ). There was a tendency toward self-archiving the publisher's pdf by authors who published in journals that allow posting of the publisher's pdf, and a tendency toward not self-archiving the publisher's pdf among authors who published in journals that disallow use. However, the difference in self-archiving of publisher's pdf versions could be explained by other factors, including disciplinary differences or regional distinctions between authors. For example authors who self-archive frequently may select journals that allow them to self-

archive the final published versions of their articles. Economists, who had higher self-archiving rates overall, also published in a greater number of journals that permitted self-archiving the publisher's pdf. Of concern, however, is that publisher policies do not seem to match well with many of the established self-archiving habits of authors, or that the policies are too varied to allow authors to create a consistent self-archiving system.

Figure 8 shows self-archiving rates of different versions of articles (preprint, postprint, publisher's pdf) show rather marked differences by discipline. Publisher's pdf's were the preferred version to self-archive within anthropology, political science and sociology, while economics authors used a greater mixture of versions, reflecting their patterns of self-archiving early preprints or working papers and then updating versions during the publishing process.

Figure 8: Comparison of Self-Archived Versions Among Disciplines



### **Self-Archiving Rates by Year of Publication**

In the four social science disciplines, there was not a significant difference between the number of articles self-archived in 2007 and 2008 ( $p < 0.2591$ ). There are a number of factors which might lead to a delay between publication and self-archiving. First, many scholars may not be in the habit of updating web pages and uploading papers on a regular basis. A second factor is that many articles included in the study were still within 12 or 24-month embargo periods when the searches were conducted. Even with these possible delaying factors, 2008 articles were not self-archived at significantly lower rates than 2007 articles.

### **Scholars' Web Pages**

Over 3,400 social science authors and their individual self-archiving habits were the primary focus of this study. As discussed above, self-archiving habits show disciplinary differences, as well as country and regional differences. The authors in the study were affiliated with a very large number of institutions from across the globe. While the groups of authors from different regions had numbers large enough to ascertain statistical differences in self-archiving behavior, comparing groups of authors from different schools and departments did not produce numbers large enough to be statistically reliable.

However, there were a number of individual and institutional self-archiving success stories that were evident from the data. There were groups of scholars in departments whose self-archiving behavior was far more consistent than other

departments in the same field. The self-archiving behavior of these groups is not generalizable to the whole group of authors, but may nevertheless be instructive.

There were a few standout departments that are worth mentioning. The first is the group of authors affiliated with the Max Planck Institutes. The Max Planck Society has been a leader in the open access movement, so it is not surprising that the Institutes would promote self-archiving among their scientists. Within the study there was a group of 38 authors from the Max Planck Institutes who published articles in the set of anthropology journals during the months measured. This group represented a little over 3% of all anthropology authors in the study. The anthropologists affiliated with the Max Planck Institutes self-archived at a rate of 55.3%. This is significant for the following reasons: 1) the overall self-archiving rate for the total group of anthropologists was 14.64%, and 2) the authors from Max Planck comprised 3% of the total anthropology authors in the study, but collectively represented 12% of the self-archiving authors.

Visiting the websites of the Max Planck Institutes reveals information that helps to explain the high rate of self-archiving. It is evident that self-archiving is the norm among the scholars. There are several distinct places on the website where self-archived publications can be located. Each scholar has a personal Web page, and many have a separate page for their publications. If self-archived articles are not found on the scholar's website, a number of the individual laboratories at Max Planck have created repositories housing the publications of their scholars. Two examples are the Molecular Genetics Laboratory ([http://www.eva.mpg.de/primat/files/genetic\\_pubs.htm](http://www.eva.mpg.de/primat/files/genetic_pubs.htm)), and the Department of Human Evolution (<http://www.eva.mpg.de/evolution/files/pubs.htm>). The

Max Planck Institutes deserve praise not only for self-archiving, but also for outstanding web design and interesting content that elaborates on their research.

Two departments within Economics were standouts. They were the MIT Department of Economics (<http://econ-www.mit.edu>) and the Harvard University Department of Economics (<http://www.economics.harvard.edu>). Both departments are located within influential research institutions that have expressed support for open access, and both schools have instituted mandatory self-archiving policies.

The authors from MIT Economics department had a self-archiving rate of 96.66%, and comprised approximately 4% of the total authors in economics. The authors from Harvard comprised 6.4% of the total economics authors, and self-archived at a 76.3% rate. Overall, economics authors self-archived at a rate of 58%. Both Harvard and MIT economists show certain similarities in self-archiving habits. Each faculty member has a well-developed personal Web page, often with additional pages dedicated to publications or working papers. Part of the successful self-archiving rates of both departments may be partly explained by the large number of faculty working papers that are later accepted for publication in high-impact journals. Because the economists at both departments consistently self-archive their working papers on their web pages, when papers are subsequently published, updating web pages to reflect publication information does not take much additional time or effort.

The sociology data did not produce significant groups of authors from one institution, as contributions were widely distributed among many institutions and many countries. The largest number of sociology authors from a single institution was 23. The self-archiving rate for that institution was lower than the average. Political science

authors were also widely distributed across many different institutions, and did not produce significantly large groups of authors from one institution.

General patterns emerged from the process of searching through scholar's Web pages. Scholars who self-archive, tend to do so fairly consistently. That is, they either self-archive all or most of their papers and articles, or they self-archive none of their articles. Authors also show consistency in self-archiving the same version of their articles. Authors who self-archived publisher's pdfs did so with a high degree of consistency regardless of the journal, indicating that they were not adapting their behavior based on publisher policies. A high percentage of authors linked from their homepage to their articles in subject repositories such as SSRN or RePEc. This behavior was most common among economists.

Within the study, there were a large percentage of scholars who created their own personal Web pages. The personally created web pages were individualized to suit what the scholar wanted to express about his or her teaching, research, writings and interests, and many deviated considerably from the standard university created Web page. For example, many of the anthropologists' web pages contained field notes, or photographs of lemurs and chimpanzees, or pictures of large groups of anthropologists at work on sites across the globe. Economists' pages had datasets, updates, and corrections. Some of more public intellectuals had websites linking to interviews in multimedia formats.

Younger scholars, including Ph.D. candidates and graduate students were single or co-authors of many of the articles in the study. Few of these scholars had a university-created location to self-archive their articles, which greatly lowered the likelihood that they had self-archived. However, some universities had created excellent job market web

pages for their candidates, and these sites often contained self-archived versions of papers. The most successful independent strategy for these younger scholars appeared to be the creation of a Google Pages web page. These web pages were easy to locate in Google with name searching or title searching, and many of the independently created Google Pages contained self-archived versions of articles.

### **Conclusion**

Although the reasons that published authors in certain social science disciplines, as well as authors from certain countries and regions have higher self-archiving rates than others has not yet been fully explained, it is evident that certain distinct groups of scholars have a tendency to self-archive more frequently than others. Self-archiving habits are to some extent influenced by discipline, region, country, and norms particular to institutions or academic departments. Primarily, however, self-archiving habits are individual, depending on a range of factors, incentives, and disincentives individual to the scholar. Authors follow the lead of their disciplines, their peers and their institutions when developing self-archiving habits (Davis & Connolly, 2007).

It is a laudable goal of research funders, universities and open access proponents to increase the amount of self-archived peer-reviewed scholarship through policies mandating deposit in subject or institutional repositories. However, the personal web pages, blogs and other informal scholarly communication technologies created and maintained by social scientists also provide an opportunity for important communication with peers and with the public. Allowing the scholar to mediate and explain his or her

own research, the way that many do so well on their personal web pages, increases the promise of more meaningful public access to social science.

Open access self-archiving of peer-reviewed articles is critical to increasing public access to social science research. Mandating deposit of published articles in institutional or subject repositories appears to be the most efficient way to eliminate the evident randomness in self-archiving habits. However, in creating large repositories of articles, it is hoped that the scholar as mediator is not completely decoupled from the research, and methods will be considered to link articles in repositories back to the scholar or scholars who created the work.

## References

- Antelman, K. (2006). Self-archiving practice and the influence of publisher policies in the social sciences. *Learned Publishing* 19(2), 85-95.
- Association of Research Libraries (2005). *ARL Statistics, 2004-2005*. Kyrillidou, M. and Young, M. (eds.) <http://www.arl.org/bm~doc/arlstat05.pdf> (accessed April 9, 2009).
- Association of American Universities, Association of Research Libraries, Coalition for Networked Information & National Association of State Universities and Land Grant Colleges (2009). *The University's Role in the Dissemination of Research and Scholarship – A Call to Action*. <http://www.arl.org/bm~doc/disseminating-research-feb09.pdf> (accessed April 11, 2009).
- Barjak, F. Li, X. & Thelwall, M. (2007). Which factors explain the web impact of scientists' personal homepages? *Journal of the American Society for Information Science and Technology*, 58(2), 200-211.
- Bergstrom, T. & Lavaty, R. (2007). *How often do economists self-archive?* <http://repositories.cdlib.org/ucsbecon/bergstrom/2007a> (accessed April 4, 2009).
- Brussels Declaration on STM Publishing* (2007). International Association of Scientific, Technical and Medical Publishers. <http://www.stm-assoc.org/brussels-declaration> (accessed April 10, 2009).
- Budapest Open Access Initiative (2002). <http://www.soros.org/openaccess/read.shtml> (accessed April 10, 2008).
- Calhoun, C. (2008). *Social Science for Public Knowledge*. [http://www.nyu.edu/ipk/files/social\\_science\\_for\\_public\\_knowledge\\_calhoun.pdf](http://www.nyu.edu/ipk/files/social_science_for_public_knowledge_calhoun.pdf) (accessed March 19, 2009).

- Campbell, B. (2007). Scientific Publishing in the European Research Area: Access, Dissemination and Preservation in the Digital Age. *Wiley-Blackwell Journal News*. <http://www.wiley.com/bw/journalnews/newsitem.asp?release=1125> (accessed April 9, 2009).
- Carr, L., & Harnad, S. (2005). Keystroke Economy: A Study of the Time and Effort Involved in Self-Archiving. <http://www.ecs.soton.ac.uk/~harnad/Temp/Keystroke3edited.doc> (accessed March 15, 2009).
- Coleman, A. (2006). Self-archiving and the Copyright Transfer Agreements of ISI-ranked library and information science journals. *Journal of the American Society of Information Science and Technology*, 58(2), 286-296.
- Davis, P. M. & Connolly, M.L.J. (2007). Institutional Repositories: Evaluating the Reasons for Non-Use of Cornell University's Installation of DSpace. *D-Lib Magazine* 13(3/4). <http://www.dlib.org/dlib/march07/davis/03davis.html> (accessed 16 March 2007).
- Directory of Open Access Journals. <http://www.doaj.org> (accessed 4 April 2009).
- ESRC & The Academy of Social Sciences (2008). *Learned Societies in the Social Sciences: Developing Knowledge Transfer and Public Engagement*. [http://www.esrc.ac.uk/ESRCInfoCentre/Images/Developing%20Dialogue\\_tcm6-27125.pdf](http://www.esrc.ac.uk/ESRCInfoCentre/Images/Developing%20Dialogue_tcm6-27125.pdf) (accessed April 10, 2009).
- Foster, N. & Gibbons, S. (2005). Understanding faculty to improve content recruitment for institutional repositories. *D-Lib Magazine*, 11(1). <http://www.dlib.org/dlib/january05/foster/01foster.html> (accessed March 15, 2009).
- Gandel, P., Katz, R. & Metros, S. (2004). The Weariness of the Flesh. Reflections on the life of the Mind in an Era of Abundance. *Educause Review* 39(2), 40-51.

- Guédon, J.C. (2001). *In Oldenburg's long shadow: librarians, research scientists, publishers and the control of scientific publishing*.  
<http://www.arl.org/resources/pubs/mmproceedings/138guedon.shtml> (accessed April 5, 2009).
- Gusterson, H. & Besteman, C. (2005). Introduction. *In: Besteman, C. & Gusterson, H. (eds.) Why America's top pundits are wrong: anthropologists talk back*. Berkeley, California: University of California Press.
- Haddow, G. (2008). Self-archiving to institutional repositories is improved by assisted and mandated deposit; disciplinary culture is not a factor. *Evidence-Based Library and Information Practice*, 3(2), 55-57.
- Hajjem, C., Harnad, S. & Gingras, Y. (2005). *Ten-year cross-disciplinary comparison of the growth of OA and how it increases citation impact*.  
<http://eprints.ecs.soton.ac.uk/12906/> (accessed April 11, 2009).
- Harnad, S. (1995). A subversive proposal. *In: Okerson, A. & O'Donnell, J. (eds.) Scholarly journals at the crossroads: A subversive proposal for E-publishing*.  
<http://www.arl.org/sc/subversive/index.shtml> (accessed April 7, 2009).
- Harnad, S. (2006). Opening Access by Overcoming Zeno's Paralysis.  
<http://eprints.ecs.soton.ac.uk/12094/2/harnad-jacobsbook.pdf> (accessed March 9, 2009).
- Harnad, S., Brody, T., Vallieres, F., Carr, L., Hitchcock, S., Gingras, Y., Oppenheim, C., Hajjem, C., & Hilf, E.R. (2008). The Access/impact problem and the green and gold roads to open access: an update. *Serials Review*, 34(1), 36-40.
- Kelty, C.M., Fischer, M.J.M., Golub, A.R., Jackson, J.B., Christen, K., Brown, M.E. & Boellstorff, T. (2008). Anthropology of/in circulation: The future of open access and scholarly societies. *Cultural Anthropology*, 23(3), 559-588.
- NIH Public Access Policy. <http://publicaccess.nih.gov> (accessed April 12, 2009).

- NISO (2008). *Journal Article Versions (JAV): Recommendations of the NISO/ALPSP JAV Technical Working Group*.  
[http://www.niso.org/committees/Journal\\_versioning/JournalVer\\_comm.html](http://www.niso.org/committees/Journal_versioning/JournalVer_comm.html)  
(accessed April 10, 2009).
- Norris, M., Oppenheim, C. & Rowland, F. (2008a). Finding open access articles using Google, Google Scholar, OAIster and OpenDOAR. *Online Information Review*, 32(6), 709-715.
- Norris, M., Oppenheim, C. & Rowland, F. (2008b). The citation advantage of open access articles. *Journal of the American Society for Information Science and Technology*, 59(12), 1963-1972.
- OAIster. <http://OAIster.umdl.umich.edu/o/OAIster/about.html>. (accessed March 24, 2009).
- OpenDOAR. (2008). About OpenDOAR. <http://www.opendoar.org/about.html> (accessed April 3, 2009).
- Pinfield, S. (2007). Can open access repositories and peer-reviewed journals coexist? *Serials*, 20(3), 163-171.
- PubMed Central. *Author Manuscripts in PMC*.  
<http://www.pubmedcentral.nih.gov/about/authorms.html> (accessed 6 April 2009).
- SHERPA/RoMEO. *Publisher copyright policies and self-archiving*.  
<http://www.sherpa.ac.uk/romeo.php> (accessed April 8, 2009).
- Suber, P. (2007). *Open Access Overview*.  
<http://www.earlham.edu/~peters/fos/overview.htm> (accessed April 11, 2009).
- Swan, A. & Brown, S. (2005). Open access self-archiving: An author study. Technical Report, Joint Information Systems Committee. *dLIST*.  
<http://dlist.sir.arizona.edu/814/01/jisc2.pdf> (accessed March 22, 2009).

- Thoms, L. & Thelwall, M. (2005). Academic home pages: Reconstruction of the self. *First Monday* 10(12). [http://firstmonday.org/issues/issue10\\_12/thoms/index.html](http://firstmonday.org/issues/issue10_12/thoms/index.html) (accessed February 27, 2009).
- Wellcome Trust (2008). *Position statement in support of open and unrestricted access to published research*. <http://www.wellcome.ac.uk/About-us/Policy/Spotlight-issues/Open-access/Policy/index.htm> (accessed April 10, 2009).
- Willinsky, J. (2004). As Open Access Is Public Access, Can Journals Help Policymakers Read Research? *Canadian Journal of Communication*, 29(3). <http://www.cjconline.ca/index.php/journal/article/viewArticle/1422/1526> (accessed April 5, 2009).
- Willinsky, J. (2006). *The Access Principle*. Cambridge, Massachusetts: MIT Press.
- Wilson, E.O. (1998). *Consilience: The Unity of Knowledge*. New York: Random House.
- Wren, J.D. (2005). Open access and openly accessible: a study of scientific publications shared via the internet. *BMJ*, 330(1128) <http://bmj.bmjournals.com/cgi/reprint/330/7500/1128> (accessed April 2, 2009).
- Wren, J.D., Kozak, K.Z., Johnson, K.R., Deakyne, S.J., Schilling, L.M. & Dellavalle, R.P. (2007). The write position. A survey of perceived contributions to papers based on byline position and number of authors. *EMBO Reports*, 8(11), 988-991.
- Xia, J. (2007). Assessment of Self-Archiving in Institutional Repositories: Across Disciplines. *The Journal of Academic Librarianship*, 33(6), 647-654.
- Zuccala, A. (2009). The Layperson and Open Access. In: Cronin, B. (ed.) *Annual Review of Information Science and Technology*. Medford, New Jersey: Information Today.

## APPENDIX

### Examples of Self-Archiving Policy Variations Among Publishers (Preprint Versions)

---

The most common policy allows self-archiving of Preprint Versions of papers accepted for publication.

Variations to this general principle include:

- Right to archive Preprint Version must be individually negotiated on an article-by-article and author-by-author basis
  - Archiving Preprint Version limited to not-for-profit or non-commercial websites
  - Archiving Preprint Version prohibited unless paper is accompanied with citation to published version and link to published version on Publisher's Website
  - Archiving Preprint Version prohibited unless there is an express acknowledgment that paper has been accepted for publication
  - Archiving Preprint Version prohibited unless there is an express acknowledgment that paper has been accepted for publication and the paper is expressly labeled as a "preprint"
  - Right to archive Preprint Version ends upon publication of the final paper and Preprint Version must be removed from all websites at that time
  - Archiving Preprint Version prohibited unless paper was archived prior to acceptance of paper for publication
  - Preprint Version may be archived, but the archived version cannot be updated to reflect changes made after date paper was accepted for publication
  - Finally, some publishers impose specific restrictions on the use of Preprint Versions in individual transfer agreements
-

### **Examples of Self-Archiving Policy Variations Among Publishers (Postprint Versions)**

---

Self-archiving policies applicable to Postprint Versions of papers accepted for publication are generally more restrictive than self-archiving policies applicable to Preprint Versions.

Below is a sampling of Postprint Version self-archiving policies:

- Right to archive Postprint Version must be individually negotiated on an article-by-article and author-by-author basis
  - Archiving Postprint Version prohibited unless paper is accompanied with citation to published version and link to published version on Publisher's Website
  - Archiving Postprint Version prohibited unless there is an express acknowledgment that paper has been accepted for publication
  - Archiving Postprint Version prohibited for an embargo period ranging from 12 to 24 months, and after the end of the embargo period self-archiving is only allowed on a not-for-profit or non-commercial website
  - Archiving Postprint Version prohibited for an embargo period ranging from 12 to 24 months, and after the end of the embargo period self-archiving is only allowed if the paper is accompanied with a link to published version on Publisher's Website
  - Archiving Postprint Version may be discouraged by the Publisher in favor of using the Publisher's PDF Version or link to published version on Publisher's Website
  - Some policies prohibiting self-archiving of Postprint Versions expressly defer to self-archiving requirements of the applicable funding organizations, others do not
  - Archiving Postprint Version limited to personal or institutional websites and archiving in subject repositories such as SSRN.com is prohibited
-

### **Examples of Self-Archiving Policy Variations Among Publishers (Publisher's PDF Versions)**

---

Self-archiving policy applicable to Publisher's PDF Versions of papers accepted for publication are consistently more restrictive than self-archiving policies applicable to either Postprint Version or Preprint Versions. The most common policy prohibits any self-archiving of Publisher's PDF Versions.

Below is a sampling of policies which permit self-archiving of Publisher's PDF Version:

- Archiving Publisher's PDF Version prohibited for an embargo period ranging from 12 to 24 months, and after the end of the embargo period self-archiving is only allowed on a not-for-profit or non-commercial website
  - Archiving Publisher's PDF Version prohibited for an embargo period ranging from 12 to 24 months, and after the end of the embargo period self-archiving is only allowed if the paper is accompanied with a link to published version on Publisher's Website
  - Archiving Publisher's PDF Version allowed individual author's website or departmental website once the paper has been published, but archiving in an institutional repository, in a subject repository or in a funding-body archive (such as PubMed Central) is prohibited for an embargo period of up to 12 months
  - Some policies prohibiting self-archiving of Publisher's PDF Version expressly defer to self-archiving requirements of the applicable funding organizations, others do not
  - Archiving of Publisher's PDF Version only allowed if Preprint Version was previously archived, and then Preprint Version must be replaced with Publisher's PDF Version once available
  - Archiving Publisher's PDF Version only allowed if author has paid acceptance fee to publisher
-