Working Paper

Faculty Self-Archiving Practices: A Case Study

Denise Troll Covey
Principal Librarian for Special Projects, Carnegie Mellon

July 2008

Abstract. Faculty web pages were examined to learn about self-archiving practice at Carnegie Mellon. More faculty are self-archiving their work and more work is being self-archived than expected. However, the distribution of self-archiving activity across the disciplines is not as expected. More faculty self-archive journal articles than other publications, but more conference papers are self-archived than journal articles. Many faculty who self-archive have self-archived fewer than ten publications. A small number of faculty has self-archived most of the work that is available open access from faculty web pages. Significant differences in faculty behavior within departments cannot be explained by disciplinary culture.

Introduction and Background

The success of initiatives to incorporate free-to-read versions of scholarly publications in the digital library hinges on the participation of faculty authors. Authors must retain the right to self-archive their work or publish in open access journals. Despite the availability of publisher self-archiving policies and research confirming the impact advantage of open access in many disciplines,[1] there is a broad gap between what could be available open access and what is available open access.[2] Though profound differences in disciplinary culture suggest that providing open access to scholarly work might not inevitably become standard practice in all fields,[3] a narrowing of the gap between opportunity and practice is likely, given educational initiatives to improve faculty understanding of access issues and copyrights, better tools and institutional support, and new public policies.

In 2006, I interviewed a stratified random sample of 87 Carnegie Mellon faculty to learn their approaches to publishing and disseminating their work.[4] Among the faculty interviewed:

- 77% did not consider copyright transfer terms when selecting a publisher.
- 52% knew what the term “open access” means, but many admitted they were guessing.
- 41% did not understand their copyright transfer agreements.
- 34% said copyright terms were not important.
- 22% had self-archived their work or published in an open access journal.
- 10% had tried to negotiate copyright transfer terms, primarily to retain the right to reuse their work; only 3% were interested in the right to self-archive their work.

Faculty track, age, gender and technological savvy appeared to be factors in faculty awareness of open access and their willingness and ability to self-archive their work. For example, the men I interviewed were more likely to be providing open access to their work than the women. Research-track faculty were more likely to be providing open access to their work than tenure- or teaching-track faculty. Younger faculty members were more likely to know about open access, to consider copyright terms when selecting a publisher, and to try to negotiate copyright transfer terms.

Prior to providing faculty with tools and instruction to facilitate self-archiving, the Libraries wanted to know what full-text material was already freely accessible from faculty web pages. In 2007-08 Carnegie Mellon University Libraries conducted two studies to improve our understanding of faculty self-archiving practices and the opportunity to self-archive in different disciplines. The purposes of these studies were to ascertain trends and to gather detailed baseline data that would inform strategic and tactical plans to facilitate self-archiving and enable assessing changes in self-archiving practice over time. After the studies started, the Provost provided funding for an institutional repository. Findings from the studies will aid the initial population of the repository by identifying content and willing contributors.

The first study examined faculty web pages to assess the types of publications produced and the types of access, if any, provided to the full text. The second study examined more closely the journal publications identified in the first study to determine whether the articles could have been self-archived in compliance with publisher policy and whether faculty practice is aligned with publisher policy. Neither study addressed why faculty do or do not choose to provide open

access to their work, what rights faculty had or thought they had when they self-archived their work, or where the open-access copies reside. This article reports key findings from the first study. Findings from the second study are in a separate article accepted for publication in *portal: Libraries and the Academy*.

Faculty behavior is influenced not only by significant differences in disciplinary culture,[5] but also by differences in the nature and mission of institutions, in departmental specialties and cultures, in available support and funding, and in age and other demographics. The differences make assembling comparative data difficult. For these reasons, studies of faculty self-archiving practice done elsewhere can provide a framework for interpreting the results of the study reported here, but cannot set a firm bar against which we can assess self-archiving activity at Carnegie Mellon. Activity at Carnegie Mellon is shaped in part by the characteristics of the campus community.

Founded by Andrew Carnegie in 1900 as the Carnegie Technical Schools and merged with the Mellon Institute in 1967, Carnegie Mellon is a relatively young institution of higher education. We are a mid-sized research university with an expanding global presence. Research and education emphasize collaboration and innovation across traditional disciplinary boundaries to solve social problems. The seven colleges in the university encompass the fine and performing arts, humanities and social sciences, physical sciences, computer sciences, engineering, business and management.

Given the passage of time and the Libraries' programmatic efforts to educate faculty about their rights, rising journal costs, declining readerships and the impact advantage of open access, we expected the study reported here to reveal that more faculty were providing open access to their work now than when I conducted the interviews in 2006. In "Open access self-archiving: An author study," Alma Swan and Sheridan Brown report an accelerating annual pace of adopting self-archiving with a 21% increase in 2004.[6] Extrapolating from their data, we expected our study to reveal that 35% to 40% of Carnegie Mellon faculty are self-archiving their work. In the absence of any data on which to base predictions, we expected only a small volume of the work produced by the faculty and cited on their web pages to be available open access. We expected

most of the self-archiving activity to occur in the School of Computer Science and the College of Engineering, with little self-archiving in the College of Humanities and Social Sciences and no self-archiving in the College of Fine Arts. We also expected faculty to pay little attention to copyright.

Following a discussion of the research method, sample and data collection process, this article presents the study findings in layers of increasing detail, beginning with the overall findings for the university, then the findings per college, and finally the findings per department and individual within the department. These are followed by some rough comparisons with the data gathered in other studies of self-archiving practice. The conclusions summarize the key findings.

Method and Sample

According to Kristin Antelman in "Do open-access articles have a greater research impact?" "Outside a few disciplines, the majority of freely available articles will not be found in a repository or in an open-access journal but, rather, on personal home pages."[7] Swan and Brown likewise report that more people self-archive using personal or departmental websites than disciplinary or institutional repositories.[8] By starting with faculty web pages, the Libraries could discover much of the material self-archived by our faculty, avoid the biases encountered in studies that start by selecting a sample of journals, and sidestep the inaccuracies associated with self-reporting.

To avoid the retrieval problems inherent in trying to find faculty web pages by searching the university web site, we started with department home pages, followed links to the departmental faculty directory and from there explored every link associated with each faculty member. Similar to the approach taken in the case study conducted by Theo Andrew at the University of Edinburgh,[9] from May 2007 through April 2008, we systematically visited each department home page and coded the publications found on linked faculty web pages. Though this approach does not capture all faculty publications or self-archived works, it does provide a detailed picture and suggest trends in faculty practice.

Many faculty members at Carnegie Mellon are affiliated with multiple departments in the university. To ensure that our study revealed as accurate a picture as possible of self-archiving practice in the university, each faculty member and his or her publications were counted only once. The data were assigned to the faculty member's home department identified by consulting the Human Resources faculty-staff directory. To the best of our ability, the study focused on full-time and emeritus faculty on the tenure, research and teaching tracks.

The term "publication" in this article means an authored work cited or referenced on faculty web pages, regardless of whether the work was officially published. The term "self-archiving" means that the faculty member or his or her designate provided a link to an open-access copy of the work. The location of the open-access copy – for example, on a local server, on a server at a co-author's institution, in a disciplinary archive or open-access journal – was not relevant.

Data Collection

The study entailed identifying the home departments of the faculty, locating and printing all of the faculty publication lists accessible from department home pages, and then for each faculty member, coding the publication and access types of the items listed. The term "list" is used loosely. Often there was a distinct list of (partial or complete) citations, but sometimes publications were simply mentioned and linked in paragraphs describing faculty interests or projects. Many faculty members had multiple publications lists, so redundant citations or references had to be eliminated. Care was taken not to eliminate the occurrence with a link to a full-text copy of the work. No data were eliminated based on date of publication.

With feedback from colleagues and administrators, we decided not to include presentations or invited lectures in the study, many of which were accompanied by self-archived PowerPoint presentations. We agreed to code the following publication types:

- Journal articles
- Conference papers – including conferences, symposia, workshops and annual meetings
- Technical reports and working papers – including reports to government agencies

- Book chapters
- Books – including theses and dissertations
- Other – including encyclopedia articles, book reviews, editorials, opinion pieces, testimonials, musical recordings and compositions, photographs, and works that could not be identified as fitting any other publication type

Faculty can provide different paths to the full text of a publication. We agreed to code the following access types:

- Open access – link to a freely available full-text version of the work, including pre-prints (author manuscripts prior to peer review) or post-prints (author manuscripts after peer review or publisher PDFs)
- On request – email link to author or statement to send email to the author to request a copy (with email address provided)
- Restricted access – link to a copy of the full-text work available by subscription or link that returned a prompt to login
- E-commerce – link to a commercial web site with a shopping cart
- Broken link – including file not found or corrupted file
- No link

One publication type and one access type were coded for each publication. If multiple access types were associated with a publication, for example, a link to an open-access copy and a link to a restricted-access copy, the open-access copy was counted.

Three librarians – Kristin Heath, Diane Covington and I, and graduate student Maureen Williams coded the publication and access types and entered the data into a spreadsheet. The data for each department were coded by only one person. I did the data analysis. Though I found and corrected some coding errors as I did the analysis, some errors likely remain. The Dean of Libraries and I agreed prior to the study that having multiple people code each publication would have increased the time and expense of the project with little if any gain for our purposes.

Overall Findings

Of the 1,018 faculty identified, 771 (76%) have one or more publication lists accessible from their department home page. Though departmental policies vary widely, the university policy on teaching-track faculty, unlike the policies for tenure- and research-track faculty, does not have a strong expectation of publication. Having done much of the coding in this study and conducted the interviews in 2006, I suspect that an analysis of the current study results by faculty track would reveal that faculty with no publication list linked to their department home page are primarily faculty on the teaching track.

Many faculty have web pages that provide lists labeled “recent” or “selected” publications. Often these lists are short and appear to be the result of efforts to provide a consistent look and feel to department web pages through the use of a template; the lack of more recent publication dates suggest that these lists are not maintained. Some faculty have personal websites that provide longer lists of publications. Some provide multiple lists, either organized around research topics or simply new presentations or updates of earlier lists that were apparently abandoned. Some faculty provide links to CVs that list hundreds of publications, perhaps covering their entire academic careers.

A total of 38,143 publications were coded from the lists accessible from department home pages. Of these, a surprising 40% (15,127) had links to an open-access copy of the work. Details per publication and access type are provided in table 1. Faculty attempted to provide access to the full-text of an additional 6% of the publications via links to copies restricted to subscribers (798 publications), email links or addresses (309 publications, copies available on request), links to e-commerce websites (194), and broken links (623). No attempt was made to provide access to the full text of the remaining 55% of the material (over 21,000 publications). Clearly faculty who choose to provide access to the full text of their work strongly prefer open access over any other method.

Table 1. Overall findings by publication and access type.

Total pubs	Publication type	Total	Open access	Restricted access	On request	E-commerce	Broken link	No link
39%	Journals	14,881	4,816	516	210	0	169	9,170
	Journals %	100%	32%	3%	1%	0%	1%	62%
38%	Conferences	14,417	7,267	231	52	1	278	6,588
	Conferences %	100%	50%	2%	0%	0%	2%	46%
8%	Reports	3,223	1,854	37	13	0	97	1,222
	Reports %	100%	58%	1%	0%	0%	3%	38%
7%	Chapters	2,700	600	5	30	20	32	2,011
	Chapters %	100%	22%	0%	1%	1%	1%	74%
3%	Books	989	124	3	2	128	24	708
	Books %	100%	13%	0%	0%	13%	2%	72%
5%	Other	1,933	466	6	2	45	23	1,413
	Other %	100%	24%	0%	0%	2%	1%	73%
100%	TOTAL	38,143	15,127	798	309	194	623	21,112
	TOTAL %	100%	40%	2%	1%	1%	2%	55%

Journal articles and conference papers account for most of the content listed on department and faculty web sites (39% and 38% respectively). Books and book chapters account for 10%; technical reports and working papers an additional 8%. In terms of sheer volume, more conference papers are available open access than any other publication type. In terms of self-archiving activity, the data suggest that conference papers and technical reports are more likely to be self-archived than journal articles. Journal articles and conference papers account for most of the restricted access and access on request. Books account for most of the e-commerce access.

The number of faculty who self-archive their work suggests the breadth of adoption of the practice. Table 2 shows the number of faculty who have self-archived anything and the number who have self-archived each publication type. Links to an open-access copy of at least one publication were provided by 42% of the faculty in the university and by 55% of the faculty with publication lists accessible from their department home pages. This exceeds our expectations. Despite the greater volume of conference papers self-archived by Carnegie Mellon faculty and the greater likelihood that conference papers and technical reports will be self-archived than journal articles, more faculty have self-archived journal articles than any other publication type. Nevertheless, hundreds of faculty have self-archived other types of publications in sufficient

numbers to warrant attention. Studies of self-archiving activity that focus strictly on journal articles provide an incomplete picture of the phenomenon.

Table 2. Faculty self-archiving activity by publication type.

Publication type self-archived	# Faculty self-archived	% Total faculty N = 1,018	% Faculty with pub list N = 771
Journal articles	344	34%	45%
Conference papers	258	25%	33%
Technical reports	260	26%	34%
Book chapters	149	15%	19%
Books	88	9%	11%
Other	86	8%	11%
Anything	424	42%	55%

Some self-archiving activity is well organized in online collections by a department or research institute. These repositories frequently acknowledge copyright transfer and restrictions on use and redistribution. When a discipline-based repository is available, on campus or elsewhere, some faculty publication lists link to the full-text work in the repository, for example in ArXiv or PubMed Central. In most cases, however, no disciplinary repository is available and self-archiving activity is more ad hoc. Many faculty who self-archive provide open access to only a small number of works and usually to only relatively recent work. Some, however, have scanned their older work and provide open access to dozens, even hundreds of works. (Details are provided later in this article.) As expected, few faculty web pages acknowledge copyright issues and possible restrictions on use or redistribution of the self-archived material.

Findings per College

Table 3 indicates the total number of publications coded per college and the frequency of the types of access provided to the full text. The data are sorted by the percentage of total publications that is available open access (OA). Most colleges do not commonly provide access to the full text. This is not surprising. What is surprising is the overall self-archiving activity in the colleges and how they rank based on the percentage of publications that are available open access. Ranked first, as expected, is the School of Computer Science with 63% of the publications available open access. Having a much smaller percentage of publications available

open access but in unexpected second place is the College of Humanities and Social Sciences with 35% of the publications available open access. The rankings in table 3 are not the rankings expected based on the faculty interviews conducted in 2006.

Table 3. Publications by college and access type.

Rank	College	Total pubs	OA	Restricted	On request	E-commerce	Broken link	No link
1	School of Computer Science	13,003	**63%**	2%	1%	0%	2%	31%
2	College of Humanities and Social Sciences	5,357	**35%**	0%	0%	1%	1%	63%
3	Heinz School of Public Policy and Management	971	**32%**	0%	0%	0%	1%	66%
4	Mellon College of Science	4,188	**31%**	6%	3%	0%	1%	59%
5	College of Engineering	10,190	**26%**	1%	0%	0%	1%	70%
6	Tepper School of Business	2,931	**22%**	3%	1%	0%	2%	71%
7	College of Fine Arts	1,503	**7%**	0%	0%	3%	1%	88%
	TOTAL	38,143	**40%**	2%	1%	1%	2%	55%

Providing restricted access to the full text (via copies available by subscription) and providing access on request (via email) are not very popular in any college, but most popular in the Mellon College of Science. Providing access to the full text via e-commerce links is even less popular, but is most popular in the College of Fine Arts.

Table 4 shows faculty self-archiving activity by college. The data are sorted by the percentage of faculty who self-archived at least one publication. Again, the School of Computer Science is in first place with 78% of the faculty self-archiving, but surprisingly Tepper School of Business is in second place with 64% of the faculty self-archiving. The College of Engineering and Mellon College of Science lag somewhat behind the College of Humanities and Social Sciences. This is not the ranking expected based on the faculty interviews conducted in 2006.

If the number of faculty who has self-archived their work suggests the breadth of adoption of the practice, then the number of publications self-archived by each faculty member suggests the depth of the practice. Table 5 shows the percentage of faculty in each college that has self-archived different volumes of publications. Throughout the university, 38% of the faculty who self-archive has self-archived no more than ten publications. In four of the seven colleges – Tepper School of Business, College of Humanities and Social Sciences, Mellon College of

Science and College of Fine Arts, over half of the faculty who self-archive has self-archived no more than ten publications. In contrast, in the College of Engineering and School of Computer Science at least 30% of the faculty has self-archived over fifty publications; at least 10% has self-archived over a hundred publications.

Table 4. Faculty self-archiving activity by college.

Rank	College	Total faculty	Faculty who self-archive	% Total faculty who self-archive
1	School of Computer Science	205	159	**78%**
2	Tepper School of Business	90	58	**64%**
3	College of Humanities and Social Sciences	148	61	**41%**
4	College of Engineering	179	64	**36%**
5	Mellon College of Science	158	56	**35%**
6	Heinz School of Public Policy & Management	54	17	**31%**
7	College of Fine Arts	184	9	**5%**
	TOTAL	1,018	424	**42%**

Table 5. Self-archiving faculty who self-archive various volumes of publications.

College / % faculty who self-archive		Number of publications self-archived						
		1 - 10	11 - 30	31 - 50	51 - 100	101 - 150	151 - 200	>200
School of Computer Science	**78%**	14%	26%	25%	21%	9%	3%	2%
Tepper School of Business	**64%**	72%	19%	5%	3%	0%	0%	0%
College of Humanities & Social Sciences	**41%**	51%	20%	10%	15%	0%	3%	2%
College of Engineering	**36%**	33%	23%	14%	19%	6%	3%	2%
Mellon College of Science	**35%**	59%	18%	7%	11%	4%	2%	0%
Heinz School of Public Policy & Mgmt	**31%**	41%	35%	12%	12%	0%	0%	0%
College of Fine Arts	**5%**	67%	11%	22%	0%	0%	0%	0%
TOTAL	**42%**	38%	23%	15%	15%	5%	2%	1%

Given different rates of publication across the disciplines and no information about when individual faculty members began their academic careers or began to self-archive, no claims can be made, based on the volume of material each has self-archived, about whether self-archiving has become a habitual activity. However, I am reluctant to conclude that self-archiving is standard, recurrent practice for all the faculty members who have self-archived fewer than ten publications. Some of this self-archiving behavior could actually deviate from the faculty member's standard practice, particularly with faculty whose only web page uses a departmental template and provides a short list of "recent" publications, one or more of which is linked to an open access copy but none of which were published within the past few years. Links to open

access copies of faculty work must be interpreted cautiously, without assuming that the faculty provided the links themselves or that they are even aware of the open access movement or impact advantage.

Examining self-archiving activity within the departments can help explain the college rankings in tables 3 and 4 and fine-tune the focus in table 5 by revealing striking departmental differences in self-archiving practice. Because the Tepper and Heinz Schools do not have departments, for the purposes of discussion they are treated as departments in the following analysis.

Findings per Department

Table 6 indicates the total number of publications coded per department and the frequency of the types of access provided to the full text. The data on open access in fields that have disciplinary repositories could be low because faculty might have self-archived their work in a disciplinary repository without linking it to the publication lists coded in this study.[10] Nevertheless, providing open access is the preferred means of providing access to the full text in all departments except Civil and Environmental Engineering and the Institute for Complex Engineered Systems, where providing restricted access occurs as frequently as providing open access.

In the College of Engineering, providing open access to the full text is more popular than providing no access in the Information Networking Institute (55% of publications self-archived), quite popular in Materials Science and Engineering (42%), and somewhat popular in Biomedical Engineering (36%) and Electrical and Computer Engineering (34%). In the College of Humanities and Social Sciences, providing open access is more popular than or almost as popular as providing no access in the Psychology (50%) and Philosophy (45%) departments and somewhat popular in Statistics (33%). In the Mellon College of Science, providing open access is equally or almost as popular as providing no access in Mathematics (48%) and Biology (40%). In the School of Computer Science, providing open access is more popular than providing no access in five of the seven departments.

Table 6. Findings by department and access type.

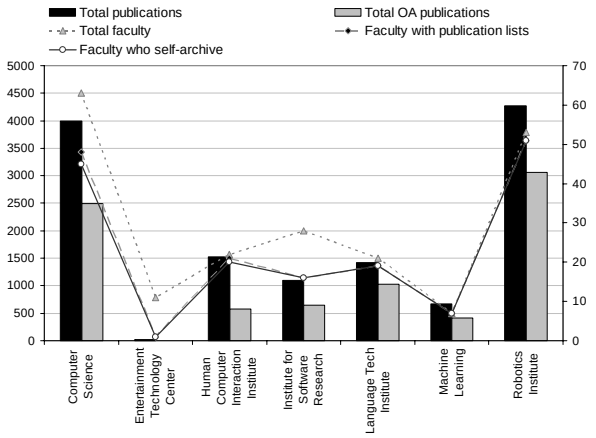
	Total pubs	OA	Restricted	On request	E-commerce	Broken link	No link
College of Engineering	**10,190**	**26%**	**1%**	**0%**	**0%**	**1%**	**70%**
Information Networking Institute	66	55%	8%	0%	0%	5%	33%
Materials Science & Engineering	1,504	42%	1%	0%	0%	1%	55%
Biomedical Engineering	188	36%	0%	0%	4%	3%	57%
Electrical & Computer Engineering	4,529	34%	1%	0%	0%	2%	63%
Mechanical Engineering	1,031	16%	1%	0%	0%	0%	82%
Engineering & Public Policy	644	13%	1%	3%	1%	2%	80%
Chemical Engineering	1,071	11%	5%	0%	0%	1%	84%
Instit for Complex Engineered Systems	31	6%	6%	0%	0%	0%	87%
Civil & Environmental Engineering	1,126	2%	2%	0%	0%	0%	95%
College of Fine Arts	**1,503**	**7%**	**0%**	**0%**	**3%**	**1%**	**88%**
Design	27	26%	0%	0%	0%	0%	74%
Architecture	875	11%	1%	0%	0%	1%	87%
Art	81	6%	0%	0%	4%	1%	89%
Drama	68	1%	0%	0%	0%	0%	99%
Music	452	0%	0%	0%	10%	0%	89%
College of Humanities and Social Sciences	**5,357**	**35%**	**0%**	**0%**	**1%**	**1%**	**63%**
Psychology	2,170	50%	0%	0%	1%	1%	48%
Philosophy	587	45%	0%	0%	1%	2%	52%
Statistics	1,312	33%	1%	0%	0%	1%	64%
English	262	10%	1%	0%	6%	2%	82%
Modern Languages	158	9%	0%	0%	3%	1%	88%
Social and Decision Sciences	633	4%	0%	0%	0%	0%	95%
History	235	0%	3%	0%	1%	1%	95%
Heinz School of Public Policy and Management	**971**	**32%**	**0%**	**0%**	**0%**	**1%**	**66%**
Mellon College of Science	4,188	31%	6%	3%	0%	1%	59%
Mathematics	917	48%	2%	0%	1%	2%	48%
Biology	388	40%	14%	0%	0%	1%	45%
Physics	1,274	34%	9%	11%	0%	0%	45%
Chemistry	1,609	16%	5%	0%	0%	1%	78%
School of Computer Science	**13,003**	**63%**	**2%**	**1%**	**0%**	**2%**	**31%**
Language Technologies Institute	1,420	72%	0%	0%	1%	3%	24%
Robotics Institute	4,272	72%	0%	0%	0%	1%	27%
Computer Science	3,999	62%	4%	0%	0%	2%	32%
Machine Learning	673	62%	2%	0%	0%	1%	35%
Instit for Software Research Intern	1,100	59%	3%	9%	1%	4%	24%
Human Computer Interaction Institute	1,519	38%	6%	0%	0%	6%	50%
Entertainment Technology Center	20	10%	0%	0%	0%	5%	85%
Tepper School of Business	**2,931**	**22%**	**3%**	**1%**	**0%**	**2%**	**71%**
CARNEGIE MELLON	**38,143**	**40%**	**2%**	**1%**	**1%**	**2%**	**55%**

In ascending order, the departments with the lowest percentage of self-archived publications (0% to 11%) are History, Music, Drama, Civil and Environmental Engineering, Social and Decision Sciences, Art, Modern Languages, English, Chemical Engineering and Architecture. Slightly above, at 16% of publications available open access, is the Chemistry department.

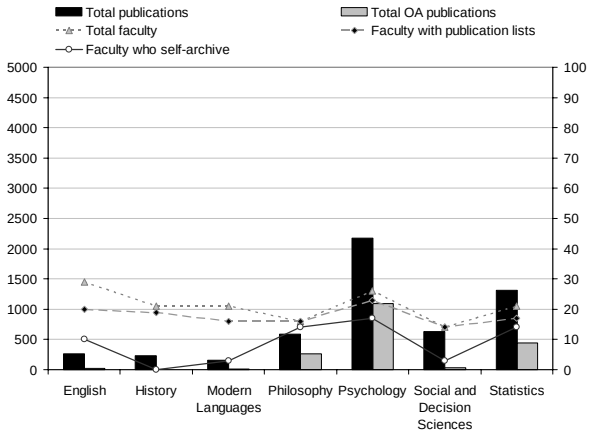
Providing access to restricted copies (available by subscription) is not very popular in any department, but most frequently occurred in Biology. Providing access on request (via email) is not common, but most frequently occurred in Physics and the Institute for Software Research International. E-commerce links are uncommon, but most frequent in Music and English. Broken links are most common in the Human Computer Interaction Institute.

Figure 1 shows the total volume of publications and open access publications per department in the context of the number of faculty in each department, the number of faculty who had publication lists linked to department home pages and the number who self-archived at least one publication. The numeric rankings are based on the percentage of total publications per college that are available open access. As expected, there is great variation in department size, volume of publications, provision of publication lists and self-archiving activity.

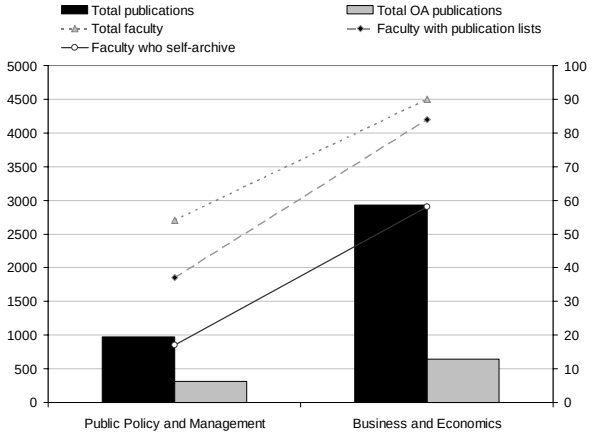
Figure 2 shows the distribution of publishing and self-archiving activity by department and publication type. With some exceptions, more conference papers, technical reports and miscellaneous other publications are self-archived than journal articles. Nevertheless, those who choose to self-archive journal articles appear to do it enthusiastically. In many cases the percentage of self-archived work that is journal articles is larger than the percentage of published work that is journal articles. In disciplines where peer-reviewed journal publications have higher status than other types of publications in terms of faculty reward systems, this disproportionate self-archiving of journal articles could be an indication that faculty choose to self-archive their best work. Stevan Harnad explains that the tendency of authors to self-archive their best work and the tendency of the best authors to self-archive their work introduce a "quality bias" (QB) in the work currently available open access.[11] This higher quality bias is one of several factors he identifies as contributing to the impact advantage of open access, and one of the factors that will disappear if 100% of authors self-archive 100% of their work.



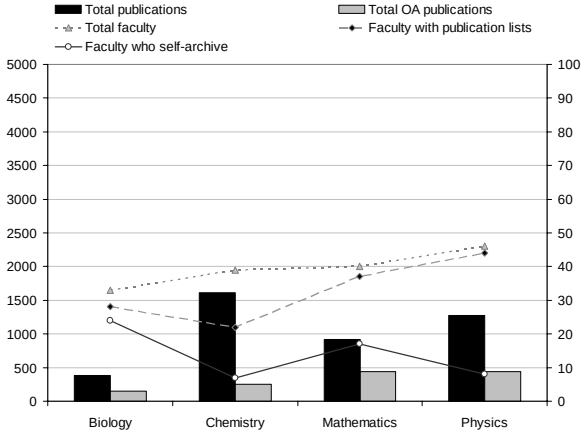
#1 School of Computer Science



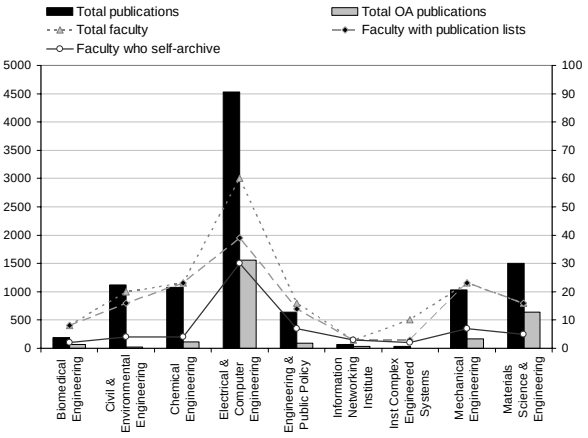
#2 College of Humanities and Social Sciences



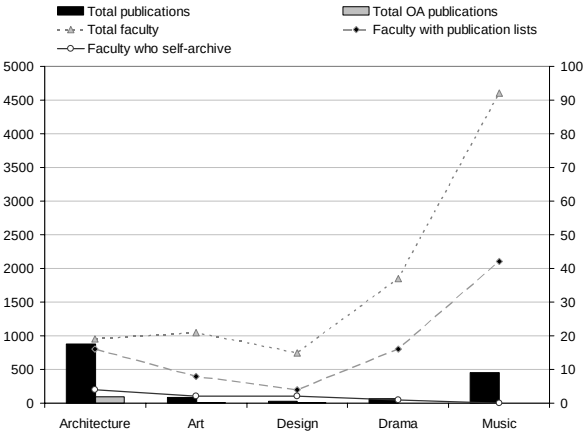
#3 Heinz School of Public Policy and Management
#6 Tepper School of Business



#4 Mellon College of Science

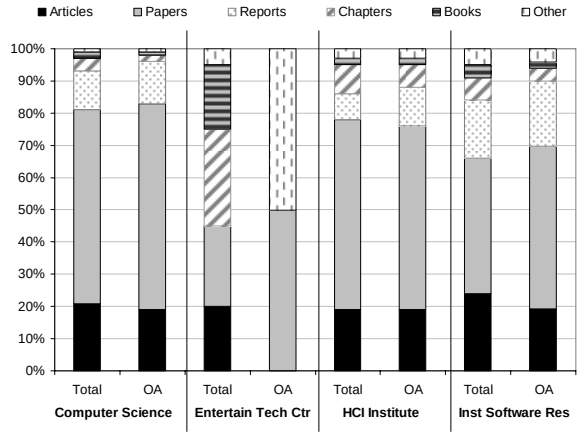


#5 College of Engineering

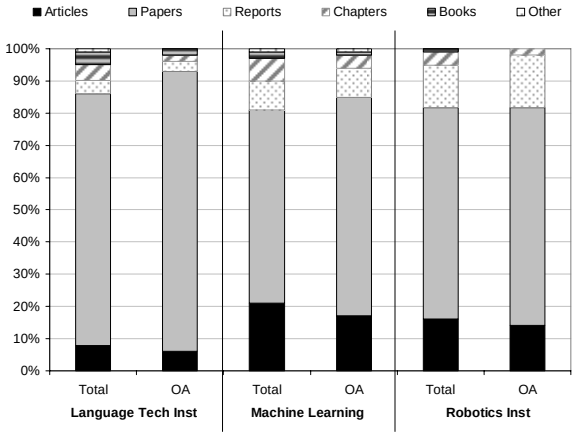


#7 College of Fine Arts

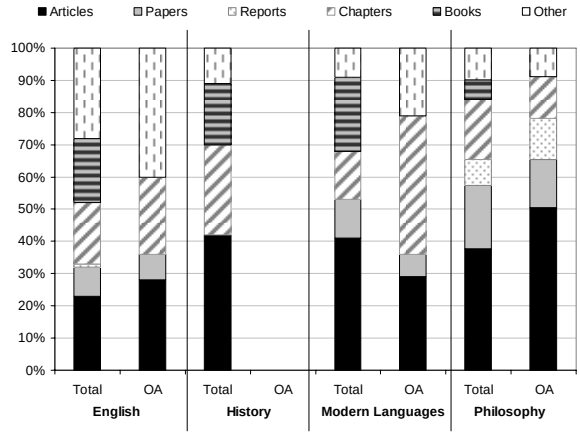
Figure 1. Total publications and open access publications per department.



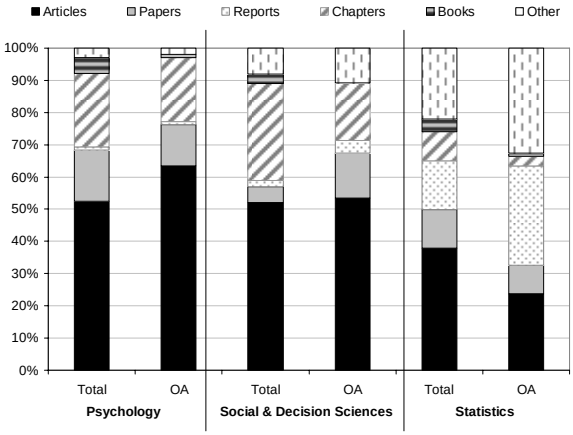
#1 School of Computer Science



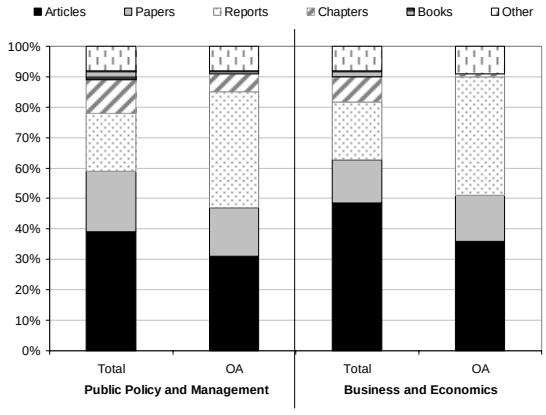
#1 School of Computer Science



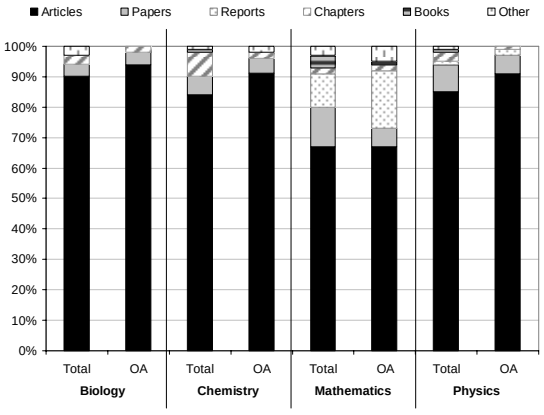
#2 College of Humanities and Social Sciences



#2 College of Humanities and Social Sciences

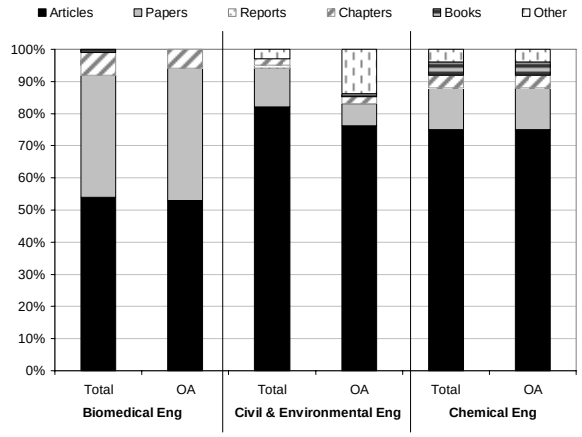


#3 Heinz School of Public Policy and Management
#6 Tepper School of Business

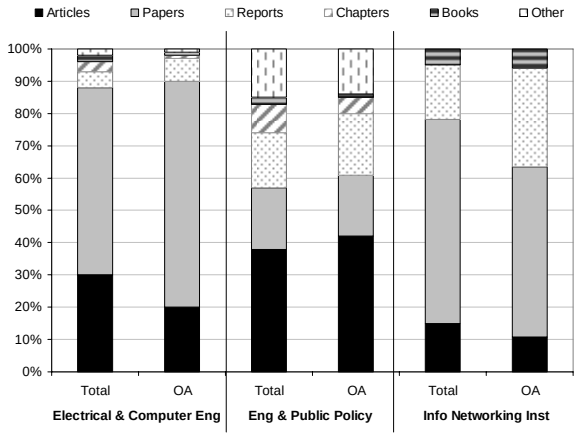


#4 Mellon College of Science

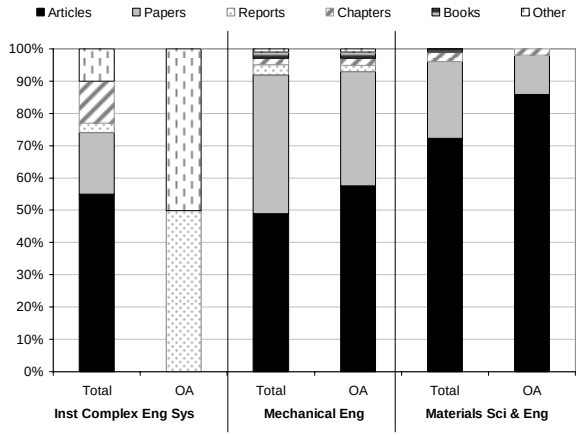
Figure 2. The distribution of publishing and self-archiving activity by department and publication type.



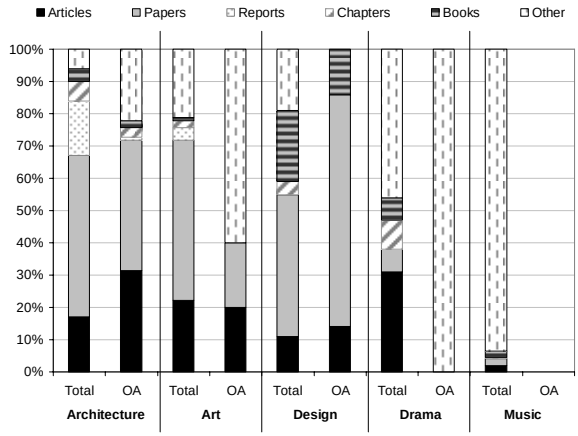
#5 College of Engineering



#5 College of Engineering



#5 College of Engineering



#7 College of Fine Arts

Figure 2 continued. The distribution of publishing and self-archiving activity by department and publication type.

Table 7 shows faculty self-archiving activity by department and publication type. In most departments, more faculty self-archive journal articles than any other type of publication. Notable exceptions are in the School of Computer Science where more faculty in five of the seven departments self-archive conference papers than journal articles. Across the university, more faculty in the Institute for Software Research International and the Engineering and Public Policy department self-archive technical reports than conference papers or journal articles. Roughly the same number of faculty self-archives journal articles and conference papers in the Information Networking Institute, Electrical and Computer Engineering, Biomedical

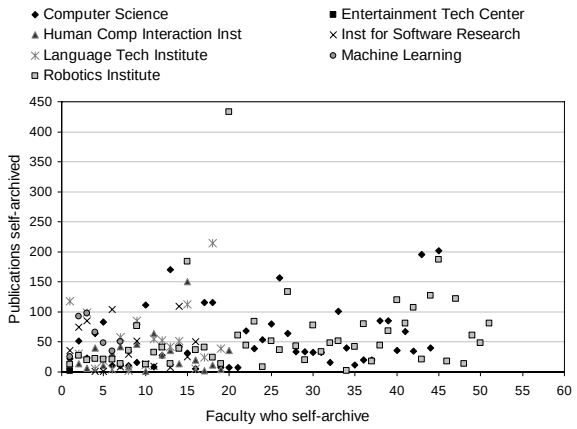
Table 7. Faculty self-archiving activity by department and publication type.

	Total faculty	Anything	Journal articles	Conference papers	Tech reports	Book chapters	Books	Other
College of Engineering	**179**	**36%**	**30%**	**26%**	**16%**	**11%**	**8%**	**9%**
Information Networking Institute	3	100%	100%	100%	100%	0%	67%	0%
Electrical & Computer Engineering	60	50%	43%	45%	25%	13%	10%	13%
Engineering & Public Policy	16	44%	31%	19%	44%	19%	6%	13%
Materials Science & Engineering	16	31%	31%	25%	0%	19%	13%	6%
Mechanical Engineering	23	30%	26%	22%	9%	9%	9%	4%
Biomedical Engineering	8	25%	24%	25%	0%	13%	0%	0%
Civil & Environmental Engineering	20	20%	10%	5%	0%	5%	5%	5%
Inst for Complex Eng Systems	10	20%	0%	0%	10%	0%	0%	10%
Chemical Engineering	23	17%	17%	9%	0%	9%	4%	9%
College of Fine Arts	**184**	**5%**	**3%**	**3%**	**1%**	**1%**	**1%**	**3%**
Architecture	19	21%	16%	16%	5%	11%	5%	16%
Design	15	13%	7%	7%	0%	0%	7%	0%
Art	21	10%	5%	5%	0%	0%	0%	5%
Drama	37	3%	0%	0%	0%	0%	0%	3%
Music	92	0%	0%	0%	0%	0%	0%	0%
College of Humanities and Social Sciences	**148**	**41%**	**34%**	**19%**	**17%**	**26%**	**4%**	**18%**
Philosophy	16	88%	88%	50%	50%	63%	0%	31%
Statistics	21	67%	62%	29%	62%	29%	10%	24%
Psychology	26	65%	65%	42%	12%	58%	15%	27%
English	29	34%	7%	3%	0%	7%	0%	24%
Social and Decision Sciences	14	21%	21%	7%	7%	21%	0%	7%
Modern Languages	21	14%	5%	5%	0%	10%	0%	5%
History	21	0%	0%	0%	0%	0%	0%	0%
Heinz School of Public Policy and Management	**54**	**31%**	**17%**	**7%**	**28%**	**11%**	**2%**	**2%**
Mellon College of Science	**158**	**36%**	**31%**	**8%**	**8%**	**7%**	**3%**	**4%**
Biology	33	73%	73%	6%	0%	6%	0%	0%
Mathematics	40	43%	28%	13%	25%	8%	8%	8%
Chemistry	39	18%	15%	3%	0%	5%	0%	5%
Physics	46	17%	17%	11%	7%	9%	2%	2%
School of Computer Science	**205**	**78%**	**67%**	**74%**	**61%**	**33%**	**28%**	**13%**
Machine Learning	7	100%	100%	100%	86%	57%	71%	29%
Robotics Institute	53	96%	89%	96%	83%	38%	25%	9%
Human Computer Interaction Inst	22	91%	68%	82%	55%	41%	32%	18%
Language Technologies Institute	21	90%	76%	90%	52%	38%	43%	10%
Computer Science	63	71%	67%	68%	59%	30%	25%	14%
Institute for Software Research Intern	28	57%	39%	43%	57%	25%	25%	14%
Entertainment Technology Center	11	9%	0%	9%	0%	0%	0%	9%
Tepper School of Business	**90**	**64%**	**46%**	**11%**	**58%**	**6%**	**3%**	**6%**
CARNEGIE MELLON	**1,018**	**41%**	**34%**	**25%**	**25%**	**15%**	**9%**	**8%**

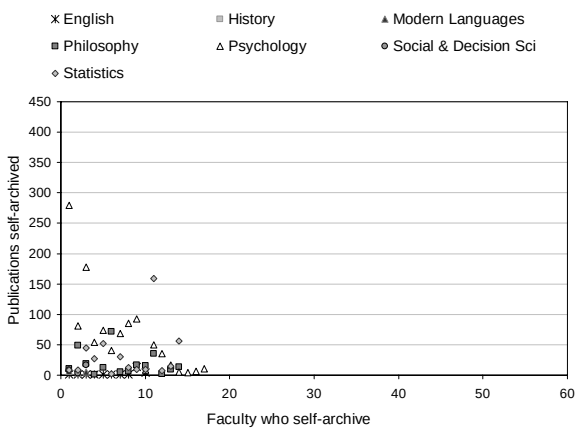
Engineering, Architecture, Design, Art, Machine Learning and the department of Computer Science. In Engineering and Public Policy, the same number of faculty self-archives book chapters as self-archives conference papers. With the exception of the Statistics department, more faculty self-archive book chapters than conference papers in all departments in the College of Humanities and Social Sciences. The significant point is the widespread self-archiving of work produced and disseminated outside of the journal literature.

The breadth and depth of self-archiving practice per faculty member is shown in figure 3. Table 8 provides the percentages of faculty who have self-archived different volumes of publications. With the exception of faculty in departments in the School of Computer Science, a relatively small number of faculty has self-archived more than 50 publications. In almost 40% of the departments with faculty who self-archive, at least half of the faculty who self-archive has self-archived no more than ten publications. In roughly 70% of the departments, at least half of the faculty who self-archive has self-archived no more than 30 publications. Yet in four departments where less than a third of the faculty self-archives – Biomedical Engineering, Materials Science and Engineering, Chemistry and Physics, there are some aggressive self-archivers who have provided open access to at least fifty if not a hundred or more of their publications. More important perhaps, in terms of interpreting the findings from this study, is the fact that a relatively small number of Carnegie Mellon faculty has self-archived most of the publications available open access from faculty web pages.

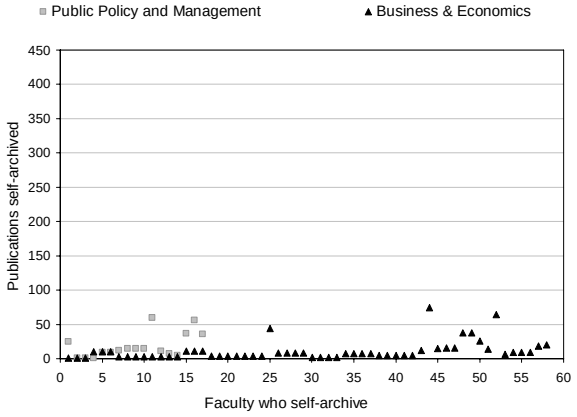
The 35 faculty in the university who have self-archived more than 100 publications – 8% of the total faculty who self-archive – have self-archived two-thirds of the total publications self-archived (15,127) by Carnegie Mellon faculty on their web pages. In the School of Computer Science 22 faculty have self-archived almost 23% of the total publications self-archived on faculty web pages; the 433 publications self-archived by a single faculty member in the Robotics Institute constitute almost 3% of the total publications self-archived on faculty web pages. Seven faculty in the College of Engineering – two in Materials Science and Engineering and five in Electrical and Computer Engineering, have self-archived almost 8% of the total publications self-archived on faculty web pages. In the College of Humanities and Social Sciences, two faculty members in Psychology and one in Statistics account for 4% of the total publications



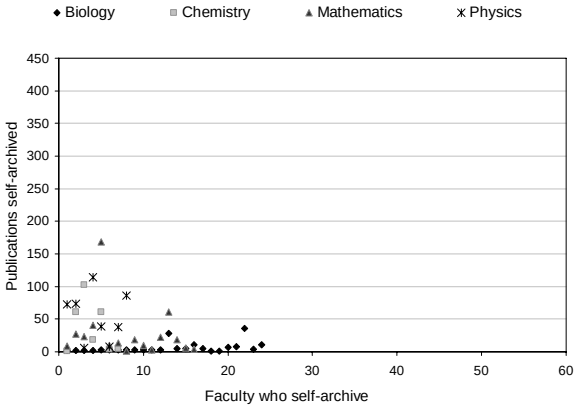
#1 School of Computer Science



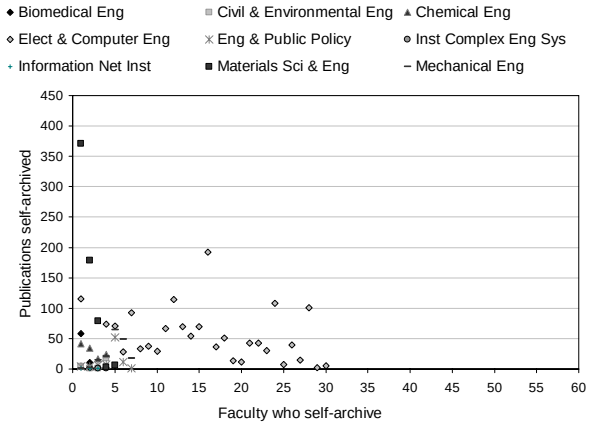
#2 College of Humanities and Social Sciences



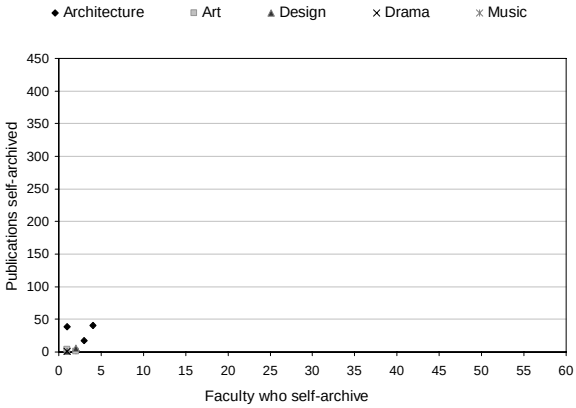
#3 Heinz School of Public Policy and Management
#6 Tepper School of Business



#4 Mellon College of Science



#5 College of Engineering



#7 College of Fine Arts

Figure 3. Depth and breadth of faculty self-archiving activity.

Table 8. Faculty by department who self-archive various volumes of publications.

	Faculty self-archive	Number of publications self-archived						
		1-10	11-30	31-50	51-100	101-150	151-200	>200
College of Engineering	**36%**	**33%**	**23%**	**14%**	**19%**	**6%**	**3%**	**2%**
Biomedical Engineering	25%	50%	0%	0%	50%	0%	0%	0%
Chemical Engineering	17%	0%	50%	50%	0%	0%	0%	0%
Civil & Environmental Engineering	20%	75%	25%	0%	0%	0%	0%	0%
Electrical & Computer Engineering	50%	17%	20%	20%	27%	13%	3%	0%
Engineering & Public Policy	44%	71%	14%	0%	14%	0%	0%	0%
Information Networking Institute	100%	33%	67%	0%	0%	0%	0%	0%
Instit for Complex Engineered Systems	20%	100%	0%	0%	0%	0%	0%	0%
Materials Science & Engineering	31%	40%	0%	0%	20%	0%	20%	20%
Mechanical Engineering	30%	29%	43%	14%	14%	0%	0%	0%
College of Fine Arts	**5%**	**67%**	**11%**	**22%**	**0%**	**0%**	**0%**	**0%**
Architecture	21%	25%	25%	50%	0%	0%	0%	0%
Art	10%	100%	0%	0%	0%	0%	0%	0%
Design	13%	100%	0%	0%	0%	0%	0%	0%
Drama	3%	100%	0%	0%	0%	0%	0%	0%
Music	0%							
College of Humanities and Social Sciences	**41%**	**51%**	**20%**	**10%**	**15%**	**0%**	**3%**	**2%**
English	34%	90%	10%	0%	0%	0%	0%	0%
History	0%							
Modern Languages	14%	100%	0%	0%	0%	0%	0%	0%
Philosophy	88%	43%	36%	14%	7%	0%	0%	0%
Psychology	65%	29%	6%	18%	35%	0%	6%	6%
Social and Decision Sciences	21%	67%	33%	0%	0%	0%	0%	0%
Statistics	67%	43%	29%	7%	14%	0%	7%	0%
Heinz School of Public Policy and Management	**31%**	**41%**	**35%**	**12%**	**12%**	**0%**	**0%**	**0%**
Mellon College of Science	**36%**	**59%**	**18%**	**7%**	**11%**	**4%**	**2%**	**0%**
Biology	73%	83%	13%	4%	0%	0%	0%	0%
Chemistry	18%	43%	14%	0%	29%	14%	0%	0%
Mathematics	43%	47%	35%	6%	6%	0%	6%	0%
Physics	17%	25%	0%	25%	38%	13%	0%	0%
School of Computer Science	**78%**	**14%**	**26%**	**25%**	**21%**	**9%**	**3%**	**2%**
Computer Science	71%	11%	24%	24%	22%	9%	7%	2%
Entertainment Technology Center	9%	100%	0%	0%	0%	0%	0%	0%
Human Computer Interaction Institute	91%	25%	40%	25%	5%	5%	0%	0%
Inst for Software Research International	57%	31%	19%	19%	19%	13%	0%	0%
Language Technologies Institute	90%	21%	16%	16%	32%	11%	0%	5%
Machine Learning	100%	0%	14%	43%	43%	0%	0%	0%
Robotics Institute	96%	4%	31%	27%	22%	10%	4%	2%
Tepper School of Business	**64%**	**72%**	**19%**	**5%**	**3%**	**0%**	**0%**	**0%**
CARNEGIE MELLON	**41%**	**38%**	**23%**	**15%**	**15%**	**5%**	**2%**	**1%**

self-archived on faculty web pages. Three faculty in the Mellon College of Science, one each in Chemistry, Mathematics and Physics, account for almost 3% of the total publications self-archived on faculty web pages. Self-archiving is no doubt routine practice for these faculty.

Department Rankings and Speculations

Table 9 ranks the top ten self-archiving departments out of the total 34 departments in the university based on four different views of the data. As expected, departments in the School of Computer Science dominate the various rankings and no department in the College of Fine Arts makes the top ten. The Heinz School of Public Policy and Management also does not appear in the top ten on any scale. Departments in the College of Humanities and Social Sciences appear as frequently in the rankings as departments in the College of Engineering and more often than some departments in Mellon College of Science. Tepper School of Business also appears more frequently than some departments in the Mellon College of Science.

The rankings based on the percentage of faculty who self-archive at least one publication and the percentage who self-archive more than 50 publications are an attempt to distinguish occasional or beginning practice from habitual practice. The top ten departments on these two scales are different, not just in the sequence of departments in the rankings but in the departments that appear in the rankings. The six departments that appear on both scales could be departments where self-archiving has become an accepted cultural practice at Carnegie Mellon. All but two of these departments, Psychology and Statistics, are in the School of Computer Science. Six departments that rank in the top ten based on the percentage of faculty who has self-archived more than 50 publications do not appear in the top ten based on the percentage of faculty in the department who self-archive. This finding suggests that at least some of the faculty in Materials Science and Engineering, Biomedical Engineering, Engineering and Public Policy, Physics, Chemistry and the Institute for Software Research International are habitual self-archivers. The most enthusiastic self-archivers appear to be in Materials Science and Engineering, where over half of the faculty who choose to self-archive have self-archived more than 50 publications.

Table 9. Top ten self-archiving departments.

	% Faculty who self-archive			% Self-archiving faculty with > 50 self-archived publications		
Rank	**Department**	**%**	**N**	**Department**	**%**	**N**
1	Information Networking Instit Machine Learning	100% 100%	3 7	Materials Science & Engineering	60%	5
2	Robotics Institute	96%	53	Physics	51%	8
3	Human Computer Interaction Institute	91%	22	Biomedical Engineering	50%	2
4	Language Technologies Institute	90%	21	Language Technologies Institute	48%	19
5	Philosophy	88%	16	Psychology	47%	17
6	Biology	73%	33	Chemistry Engineering and Public Policy Machine Learning	43% 43% 43%	7 7 7
7	Computer Science	71%	63	Computer Science	40%	45
8	Statistics	67%	21	Robotics Institute	38%	51
9	Psychology	65%	26	Inst for Software Research Intern	32%	16
10	Tepper School of Business	64%	90	Statistics	21%	14

	% Total publications self-archived			Total # self-archived publications	
Rank	**Department**	**%**	**N**	**Department**	**N**
1	Language Technologies Institute Robotics Institute	72% 72%	1,420 4,272	Robotics Institute	3,058
2	Computer Science Machine Learning	62% 62%	3,999 673	Computer Science	2,491
3	Inst for Software Research Intern	59%	1,100	Electrical and Computer Engineering	1,554
4	Information Networking Institute	55%	66	Psychology	1,090
5	Psychology	50%	2,170	Language Technologies Institute	1,029
6	Mathematics	48%	917	Inst for Software Research Intern	647
7	Philosophy	45%	587	Tepper School of Business	639
8	Materials Science and Engineering	42%	1,054	Materials Science and Engineering	638
9	Biology	40%	388	Human Computer Interaction Institute	578
10	Human Computer Interaction Institute	38%	1,519	Mathematics Statistics	439 439

The rankings based on the percentage of total publications that are self-archived and the total number of self-archived publications are another attempt to discern where self-archiving could be a habitual practice with at least some faculty in the departments. Two departments appear in these rankings that do not appear in the rankings based on the percentage of faculty who self-archive: Mathematics and Electrical and Computer Engineering.

Several additional observations are noteworthy. Psychology ranks in the top ten on all four scales, Statistics on three of the scales and Philosophy on two scales. Philosophy and Statistics rank in the top ten in terms of the percentage of faculty who self-archive. A larger percentage of faculty in the Philosophy department self-archives their work than faculty in the Biology,

Computer Science, Statistics and Psychology departments. Philosophy is also in the top ten in terms of the percentage of total publications self-archived. Statistics is also in the top ten in terms of the percentage of the faculty who has self-archived more than 50 publications and the total number of publications self-archived. The data suggest that self-archiving practice has strongly penetrated Psychology, Philosophy and Statistics at Carnegie Mellon and explain how self-archiving in the College of Humanities and Social Sciences equals or exceeds self-archiving in the College of Engineering and Mellon College of Science.

Rough Comparisons

Assembling comparative data is difficult, but it can provide a context for interpreting the data from the Carnegie Mellon study. The comparisons are apt only in broad strokes, not in the details. To use an analogy, we can talk about fruit as a general category of edible, seed-bearing plant crop, but comparing apples and oranges is problematic. In the same sense, we can discuss attempts to discern self-archiving rates, but comparing the rates derived from different studies is problematic. Many factors can account for discovered discrepancies, including the timeframes of the studies, the research methods and the characteristics of the samples. Significant differences thwart the ability to draw firm conclusions. Nevertheless comparisons can be fruitful as an exploratory exercise.

Table 10 aligns the data from the Carnegie Mellon study with the data from two studies conducted by Kristin Antelman,[12] published in 2004 and 2006, and a study conducted by Chawki Hajjem, Stevan Harnad and Yves Gingras published in 2005. [13] Because the Antelman and Hajjem et al studies looked strictly at the self-archiving of journal articles, the Carnegie Mellon data in the table are the data on self-archiving journal articles, not the other publication types examined in the study. The Carnegie Mellon sample includes all the journal titles and articles cited or referenced on faculty web pages, including the older works listed on CVs. In contrast, the Antelman and Hajjem et al studies examined selected journal titles, in Antelman's case no more than ten titles per discipline, and in the Hajjem et al study the journals covered by the ISI Science and Social Science Citation Indices. The significantly different timeframes and sample

sizes for the comparative studies are shown in the table. The table also indicates the disciplines where Antelman examined only high impact journals.

Table 10. Comparative rates of self-archiving journal articles in the Antelman, Hajjem et al and Troll Covey studies.
* Antelman's study included ten high-impact journals per discipline

Discipline	Study	Sample content (publication dates)	Total articles	Self-archiving rate
Biology	Hajjem et al	1992-2003	c. 600,000	15%
	Troll Covey	1971-2007	352	41%
Business / Economics	Antelman	2003-2005	300+	60%
	Hajjem et al	1992-2003	< 100,000	9% / 14%
	Troll Covey	1962-2007	1,415	16%
Electrical and electronics engineering *	Antelman	2001-2002	500 - 600	37%
	Troll Covey	1968-2007	1,350	23%
Management	Hajjem et al	1992-2003	< 40,000	7%
	Troll Covey	1971-2007	380	26%
Mathematics *	Antelman	2001-2002	500 - 600	69%
	Troll Covey	1952-2007	617	47%
Philosophy *	Antelman	1999-2000	500 - 600	17%
	Troll Covey	1970-2007	221	61%
Psychology	Antelman	2003-2005	300+	28%
	Hajjem et al	1992-2003	c. 220,000	7%
	Troll Covey	1963-2008	1,141	62%

In comparison with Antelman's findings, faculty at Carnegie Mellon appear to be less enthusiastic about self-archiving journal articles in Economics, Mathematics and Electrical and Electronic Engineering, but more enthusiastic about self-archiving journal articles in Philosophy and Psychology. In comparison with the findings of Hajjem et al, the self-archiving of journal articles in Business and Economics at Carnegie Mellon is slightly higher, self-archiving of articles in Biology and Management at Carnegie Mellon is roughly three times as high, and self-archiving in Psychology is almost nine times as high.

In addition to the different timeframes, research methods and samples in the studies, the fact that Carnegie Mellon faculty are more likely to self-archive more recent publications than older works, many of which are listed on their CVs or otherwise cited or referenced on their websites, could also skew the findings, as could self-archiving in disciplinary repositories without linking the open access copies to their web pages. The tendency of authors to self-archive their best work, the quality bias noted earlier in this article, could explain the higher self-archiving rate in

Mathematics and Electrical and Electronic Engineering found in Antelman's study of selected high impact journals. It could not explain the much greater self-archiving of Philosophy articles by Carnegie Mellon faculty than Antelman discovered in her study of articles published in high impact Philosophy journals. The much larger sample size in the Hajjem study provides a more comprehensive picture of overall academic practice, suggesting that self-archiving activity at Carnegie Mellon might be significantly different, perhaps because of the interdisciplinary nature of the institution or its focus on applied research. Another possibility is that faculty in selected disciplines at Carnegie Mellon and elsewhere could have adopted or accelerated the practice of self-archiving in the years since Antelman and Hajjem completed their studies.

Similar to the Carnegie Mellon case study, in 2003 Theo Andrew systematically examined personal web pages at the University of Edinburgh to assess faculty self-archiving of a variety of publication types.[14] In 2005 Alma Swan and Sheridan Brown gathered self-reports from 1,296 people (74% of whom were academics) on self-archiving article pre-prints and post-prints using personal web pages, institutional repositories and disciplinary repositories.[15] Table 11 compares the findings from these studies with the data from the Carnegie Mellon study. For the purposes of comparison, the data from the Swan and Brown study are restricted to self-archiving using personal web pages, which their survey revealed was where most of the self-archiving occurred.

With the exception of Music and Mathematics, a larger percentage of faculty at Carnegie Mellon self-archives in the comparable disciplines than at the University of Edinburgh. The disparity is striking in Economics, Philosophy and Psychology. In comparison with the findings in the Swan and Brown study, self-archiving in Chemistry, Mathematics and Physics appears to be somewhat lower at Carnegie Mellon, but somewhat higher in Computer Sciences and much higher in Psychology. Many of the factors described above could account for the differences.

Table 11. Comparative rates of faculty self-archiving practice in the Andrew, Swan and Brown, and Troll Covey studies.

Discipline	Study	Sample content	Total faculty	Self-archiving rate
Architecture	Andrew	Articles	18	0%
	Troll Covey	Articles	19	21%
Art	Andrew	Articles	15	0%
	Troll Covey	Articles	21	10%
Biology	Andrew	Articles, theses & dissertations	177	8%
	Troll Covey	Articles	33	27%
Chemistry	Andrew	Articles, theses & dissertations	43	12%
	Swan and Brown	Self-reports article pre-prints / post-prints	78	25% / 28%
	Troll Covey	Articles	39	15%
Computer sciences	Swan and Brown	Self-reports article pre-prints / post-prints	156	45% / 60%
	Troll Covey	Articles	205	67%
Economics	Andrew	All publication types	26	19%
	Troll Covey	All publication types	90	64%
Management	Andrew	All publication types	132	0%
	Troll Covey	All publication types	54	31%
Mathematics	Andrew	Articles, theses & dissertations	59	29%
	Swan and Brown	Self-reports article pre-prints / post-prints	78	33% / 44%
	Troll Covey	Articles	40	28%
Music	Andrew	All publication types	16	13%
	Troll Covey	All publication types	92	0%
Philosophy	Andrew	All publication types	21	14%
	Troll Covey	All publication types	16	88%
Physics	Andrew	Articles, theses & dissertations	104	8%
	Swan and Brown	Self-reports article pre-prints / post-prints	91	21% / 36%
	Troll Covey	Articles	46	17%
Psychology	Andrew	All publication types	59	2%
	Troll Covey	All publication types	26	65%
	Swan and Brown	Self-reports article pre-prints / post-prints	117	20% / 33%
	Troll Covey	Articles	26	65%
School of Engineering	Andrew	All publication types	143	7%
	Troll Covey	All publication types	179	36%

Conclusions

For Carnegie Mellon faculty who choose to provide access to the full text of their work, providing open access is the preferred method. Self-archiving activity in the university is much

more popular and widespread than expected. At least 42% of the faculty has self-archived one or more publications and the practice has penetrated all colleges and all but the History and Music departments. A surprising 40% of the content cited on faculty web pages is available open access, including half of the conference papers and over half of the technical reports. However, the distribution of self-archiving activity within the university is not as expected. We expected more activity in the College of Engineering and Mellon College of Science and much less activity in the College of Humanities and Social Sciences.

Though more faculty have self-archived journal articles than any other publication type, many faculty have self-archived other types of publications in sufficient number to warrant attention. Studies of self-archiving practice that focus strictly on journal articles provide an incomplete picture of the phenomenon. Journal-based studies not only introduce a sampling bias into the research, but also a publication-type bias likely driven by traditional reward systems. Reward systems are slowly changing. In the School of Computer Science at Carnegie Mellon, for example, peer-reviewed conference publications now carry the weight of journal publications in the review process for promotion and tenure. More importantly, studies that privilege journal publication over other types of publication exhibit a very narrow view of scholarly communication, not the broad, process-oriented view currently driving discussions about the creation, transformation, dissemination and preservation of knowledge.

Significant differences in self-archiving activity, not only across but within departments, suggest that disciplinary differences alone do not account for faculty behavior. This finding supports the conclusions drawn from the faculty interviews I conducted in 2006. Within most departments, some faculty have self-archived nothing, some have self-archived a volume of material sufficient to suggest that self-archiving has become or is becoming a habit, and some have self-archived so little that it could have been a one-time occurrence, perhaps performed by an assistant or co-author. Interdisciplinary collaborations, disciplinary specialties, departmental priorities, available support, age, gender, faculty track, rank on the track and technological savvy could be influential factors in determining whether, when or what faculty choose to self-archive.

Reminiscent of the 80-20 rule, prolific self-archiving by a small number of faculty accounts for a large percentage of the self-archived work. Specifically, 8% of the faculty who self-archive have self-archived two-thirds of the material available open access from faculty web pages. If we assume that faculty who have self-archived more than 30 publications are habitual self-archivers, then most Carnegie Mellon faculty who self-archive their work have not yet developed the habit. Self-archiving does appear to be a broadly accepted cultural practice not only in the School of Computer Science, but also in Psychology, Philosophy and perhaps Statistics. Enthusiastic or habitual self-archivers are sprinkled throughout many other departments, including some surprising niches where less than a third of the faculty in the department self-archive, for example, Architecture, Chemistry, Biomedical Engineering, Materials Science and Engineering and the Heinz School of Public Policy.

Comparative data across institutions are helpful in developing strategic and tactical plans because faculty are influenced by what their peers are doing and what their peers value. Where Carnegie Mellon faculty appear to be in step with their peers or even more aggressive in terms of self-archiving, we can identify champions to help us educate colleagues in their department and, given the interdisciplinary nature of research at Carnegie Mellon, colleagues in collaborating departments. Where Carnegie Mellon faculty appear to be lagging behind their peers in self-archiving their work, we can target efforts to educate them about the importance of open access and what their peers are doing. More studies, studies in a broader array of disciplines and studies that encompass a variety of publication types are needed to facilitate comparisons and to understand and assess penetration of self-archiving activity across the disciplines.

Lack of consistent faculty attention to copyright issues related to self-archiving is an area of concern, as is the ephemeral nature of personal web pages. Moving Carnegie Mellon faculty from their ad hoc approach to self-archiving via personal web pages to depositing their work in our new institutional repository will be a slow transition. We respect disciplinary differences in whether and how faculty share their work and in the degree to which the rising cost of journal subscriptions impacts readership. Our goals are to help faculty understand the issues so that they can make informed choices and, if the choice is to self-archive, to provide tools and support that will help them better showcase, disseminate and preserve their work.

Notes

1. Alma Swan, "Open Access and the Progress of Science." *American Scientist* 95 (May-June 2007): 198-200, http://eprints.ecs.soton.ac.uk/13860/1/American_Scientist_article.pdf (accessed May 22, 2008); Chawki Hajjem, Stevan Harnad and Yves Gingras, "Ten-Year Cross-Disciplinary Comparison of the Growth of Open Access and How it Increases Research Citation Impact," *Bulletin of the IEEE Computer Society Technical Committee of Data Engineering* 28, 4 (2005): 39-47, http://eprints.ecs.soton.ac.uk/11688/ (accessed May 22, 2008); The Open Citation Project, *The Effect of Open Access and Downloads ("Hits") on Citation Impact: A Bibliography of Studies*, http://opcit.eprints.org/oacitation-biblio.html (accessed May 28, 2008).
2. Hajjem, Harnad and Gingras, "Ten-Year Cross-Disciplinary Comparison of the Growth of Open Access and How it Increases Research Citation Impact"; Stevan Harnad, Tim Brody, François Vallières, Les Carr, Steve Hitchcock, Yves Gingras, Charles Oppenheim, Heinrich Stamerjohanns, and Eberhard R. Hilf, "The access / impact problem and the green and gold roads to open access," *Serial Review* 30, 4 (2004): 310-314.
3. Rob Kling and Geoffrey McKim, "Not Just a Matter of Time: Field Differences and the Shaping of Electronic Media in Supporting Scientific Communication," *Journal of the American Society for Information Science* 51, 14 (December 2000): 1306-1320, http://arxiv.org/ftp/cs/papers/9909/9909008.pdf (accessed May 22, 2008).
4. Denise Troll Covey, "Faculty Rights and Other Scholarly Communication Practices," (presentation at the Digital Library Federation Forum, Boston, MA, November 2006), http://www.library.cmu.edu/People/troll/DLF_FacultyStudy_2006.ppt (accessed May 28, 2008); Denise Troll Covey, "Faculty Rights and Other Scholarly Communication Practices," (presentation in the Digital Library Colloquium, Pittsburgh, PA , January 17, 2007, http://www.library.cmu.edu/People/troll/FacultyStudy_DL_Colloquium.ppt (accessed June 3, 2008).
5. Jenny Fry and Sanna Talja, "The Cultural Shaping of Scholarly Communication: Explaining E-Journal Use Within and Across Academic Fields," American Society for Information Science and Technology Annual Meeting (Providence, RI: November 2004),

http://www.uta.fi/~sanna.talja/FryTalja_asistfinal_konv.pdf (accessed May 23, 2008); Jouni Kekale, "Conceptions of Quality in Four Different Disciplines," *Tertiary Education and Management* 8, 1 (2002): 65-80.

6. Alma Swan and Sheridan Brown, *Open access self-archiving: An author study* (Cornwall, UK: Key Perspectives Limited, May 2005): 39, http://eprints.ecs.soton.ac.uk/10999/1/jisc2.pdf (accessed May 23, 2008).
7. Kristin Antelman, "Do open-access articles have a greater research impact?" *College & Research Libraries* 65, 5 (September 2004): 380, http://www.ala.org/ala/acrl/acrlpubs/crljournal/crl2004/crlseptember/antelman.pdf (accessed May 22, 2008).
8. Alma Swan and Sheridan Brown, *Open access self-archiving: An author study*, 27.
9. Theo Andrew, "Trends in Self-Posting of Research Material Online by Academic Staff," *Ariadne* 37 (October 2003), http://www.ariadne.ac.uk/issue37/andrew/intro.html (accessed May 22, 2008).
10. For example, faculty conducting research in psychology, neuroscience, linguistics and selected areas in computer science, philosophy and biology could have self-archived their work in CogPrints (http://cogprints.org). Faculty in mathematics could have self-archived their work in the Mathematics Preprint and e-Print Servers provided by the American Mathematical Society (AMS) (http://www.ams.org/global-preprints/index.html). Research in economics could be self-archived in Research Papers in Economics (RePEc, http://www.repec.org/). Research in physics, mathematics, computer science, quantitative biology and statistics could be self-archived in ArXiv (http://arxiv.org/). Research in computer science could also be self-archived in CiteSeer (http://citeseer.ist.psu.edu/).
11. Stevan Harnad, "OA Impact Advantage = EA + (AA) + (QB) + QA + (CA) + UA" (September 2005), http://eprints.ecs.soton.ac.uk/12085/01/OAA.html (accessed May 23, 2008); Chawki Hajjem and Stevan Harnad, *The Self-Archiving Impact Advantage: Quality Advantage or Quality Bias?* (November 2006), http://eprints.ecs.soton.ac.uk/13193/ (accessed May 22, 2008).
12. Kristin Antelman, "Self-Archiving Practice and the Influence of Publisher Policies in the social sciences," *Learned Publishing* 19 (2006): 85-95,

http://www.lib.ncsu.edu/staff/kaantelm/antelman_self-archiving.pdf (accessed May 22, 2008); Antelman, "Do open-access articles have a greater research impact?"

13. Hajjem, Harnad and Gingras, "Ten-Year Cross-Disciplinary Comparison of the Growth of Open Access and How it Increases Research Citation Impact."
14. Andrew, "Trends in Self-Posting of Research Material Online by Academic Staff."
15. Swan and Brown, *Open access self-archiving: An author study*.