

Trends in Self-Posting of Research Material Online by Academic Staff

1. Introduction

With the rapid uptake of digital media changing the way scholarly communication is perceived, we are in a privileged position to be part of a movement whose decisions now will help to ultimately decide future courses of action. A number of strategies have recently emerged to facilitate greatly enhanced access to traditional scholarly content, e.g. open access journals and institutional repositories.

In this spirit of promoting access to scholarly resources, JISC has funded a number of projects under the banner of its Focus on Access to Institutional Resources (FAIR) Programme¹. The University of Edinburgh is in a distinguished position to be involved in two of these projects, namely the Theses Alive! and SHERPA (Securing a Hybrid Environment for Research Access and Preservation) projects. The Theses Alive! project², based at the University of Edinburgh, is seeking to promote the adoption of a management system for electronic theses and dissertations (ETDs) in the UK, primarily by initiating an easily searchable and accessible online repository of ETDs at Edinburgh and at five pilot institutions across the UK. The main thrust of the SHERPA project, led by the University of Nottingham, is the creation, population and management of several e-print repositories based at several partner institutions³.

However, the uptake of digital media being used for research dissemination is not just limited to institutions and organisations. Individual scholars, seeing an opportunity to easily distribute their work to a potential wide audience and as a by-product raise their research profiles, have also seized the opportunity to use new technology by posting research material online. Currently the scale of this activity is not yet known. For the institutional repository concept to be successful then it would make great sense to look and learn from current trends of digital media usage.

Prior to the implementation of these projects at the University of Edinburgh, it was decided that a baseline survey of research material already held on departmental and personal web pages in the ed.ac.uk domain would be beneficial in a number of ways:

- it would provide a qualitative view of web usage across different subject areas, something that at the present time is poorly understood.
- it would aid the initial population of the repositories by identifying ready material and willing scholarly contributors.
- such a survey would provide an invaluable baseline upon which progress of the projects can be measured during evaluation.

¹ http://www.jisc.ac.uk/index.cfm?name=programme_fair

² <http://www.thesesalive.ac.uk>

³ <http://www.sherpa.ac.uk>

2.Methods

A number of survey methods to obtain information on the nature and volume of available online research material in the ed.ac.uk domain were initially considered. However, it was felt that due to the need for prompt and accurate results a questionnaire or sampling technique would not be appropriate. Instead a systematic approach was taken, whereby each departmental and staff web page was visited and the content of self-archived material was noted.

The University of Edinburgh's academic structure is based on three Colleges containing a total of 21 Schools, each made up from a varying number of disciplines. The University's online presence follows a similar hierarchy, although there is around half a million pages published within the ed.ac.uk domain, which represents in excess of one-tenth of all web pages published within the UK HE sector. The survey looked at each College in turn, searching for content at each level of the hierarchy, down through the School and to individual levels. During the course of the survey, carried out between 27/5/03 and 13/6/03, over 2500 staff web pages were visited.

Initially the survey began with documenting formal research material (e.g. postprints, preprints, theses and dissertations) within the College of Science and Engineering domain, but when other Colleges within the University were surveyed it became apparent that the type of material archived online varies considerably between subjects. To represent the different research cultures and web usage across such a diverse institution, other content, such as book chapters, conference and working papers, were also considered when compiling the data for the other Colleges.

3.Results

The results of the survey are displayed and discussed in the following tables and figures. It is worth noting that only academic and research staff are considered in the staff totals, and that therefore these figures do not represent the total staff of the university.

Sometimes research material was stored in School-based web pages and not on individual staff pages. It was necessary to describe this scenario as a 'School-based resource' to distinguish between self-archiving individuals and subject-based archiving. Where possible the School-based resource is described in a footnote to the table.

Table 1: Formal research material held on departmental and staff WebPages for the College of Science & Engineering

College of Science & Engineering	Staff	Number of Self Archiving staff	Papers	Pre-prints	PhD	MSc	PhD in Departmental site	% of S.A Staff
School of Biological Sciences	177	14	193	0	0	0	0	7.91
Institute of Cell & Molecular Biology	54	7	78	0	0	0	0	12.96
Institute of Cell, Animal & Population Biology	54	7	115	0	0	0	0	12.96
Institute for Stem Cell Research	69	0	0	0	0	0	0	0
School of Chemistry	43	5	114	0	1	0	0	11.63
School of Engineering & Electronics	143	10	27	0	1	0	22	6.99
Inst. of Materials & Processes	30	2	1	0	1	0	0	6.67
Inst. of Integrated Micro & Nano Systems	44	3	6	0	0	0	0	6.82
Inst. of Digital Communications	29	4	14	0	0	0	22	13.79
Inst. of Energy Systems	15	1	6	0	0	0	0	6.67
Inst. of Infrastructure & Environment	25	0	0	0	0	0	0	0
School of GeoScience	107	13	78	0	0	2	0	12.15
Geography	0	0	0	0	0	0	0	0
Geology & Geophysics	43	5	41	0	0	0	0	11.63
Meteorology	17	5	8	0	0	2	0	29.41
Inst. of Ecology & Resource Management	17	3	29	0	0	0	0	17.65
School of Informatics	150	49	184	0	17	3	31	32.67
School of Mathematics	59	17	187	52	3	0	0	28.81
School of Physics	104	8	71	5	2	0	0	7.69
TOTAL	783	116	854	57	24	5	53	14.81

Table 2: Research material held on departmental and staff WebPages for the College of Humanities & Social Science.

College of Humanities & Social Science	Staff	S.A staff	Papers	Pre-Prints	Conference Papers	Working Papers	Reports	Book Chapters	PhD	MSc	Other	Dept. Resource	% of S.A
School of Arts, Culture & Environment	62	2	0	0	0	0	0	0	0	0	3	1	3.23
Archaeology	13	0	0	0	0	0	0	0	0	0	0	0	0
Architecture	18	0	0	0	0	0	0	0	0	0	0	0	0

College of Humanities & Social Science	Staff	S.A staff	Papers	Pre-Prints	Conference Papers	Working Papers	Reports	Book Chapters	PhD	MSc	Other	Dept. Resource	% of S.A
Fine Art	15	0	0	0	0	0	0	0	0	0	0	0	0
Music	16	2	0	0	0	0	0	0	0	0	3 ¹	1 ²	12.5
School of Divinity	27	0	0	0	0	0	8	0	0	0	9³	2	0
School of Economics & Management Studies	158	5	18	0	12	42	6	0	0	0	17	3	3.16
Economics	26	5	8	0	0	32	0	0	0	0	0	0	19.23
Management School	132	0	5	0	12	10	6	0	0	0	17 ⁴	3	0
School of History & Classics	69	2	0	0	4	1	0	0	0	0	4⁵	3	2.90
Classics	17	0	0	0	0	0	0	0	0	0	0	2	0
History	28	0	0	0	0	0	0	0	0	0	0	0	0
Economic & Social History	14	2	0	0	4	1	0	0	0	0	4	0	14.29
Scottish History	10	0	0	0	0	0	0	0	0	0	0	1 ⁶	0
School of Law	54	0	0	0	0	0	0	0	0	0	0	1⁷	0
School of Literature, Languages & Cultures	118	0	12	0	0	0	1	0	0	0	49	6	0
English Literature	30	0	0	0	0	0	1	0	0	0	2 ⁸	1	0
Islamic & Middle Eastern Studies	12	0	0	0	0	0	0	0	0	0	2 ⁸	1	0
Celtic & Scottish Studies	16	0	0	0	0	0	0	0	0	0	0	1 ⁹	0
European Languages & Culture	60	0	12	0	0	0	0	0	0	0	45 ¹⁰	3 ¹¹	0
School of Nursing Studies	24	0	0	0	0	0	1	0	0	0	0	1	0
School of Philosophy, Psychology & Language Sciences	126	16	160	18	107	7	16	19	13	4	7	1	12.70
English Language	10	1	0	1	0	0	0	0	0	0	0	0	10.00
Theoretical & Applied Linguistics	36	11	130	11	97	7	12	17	12	4	5 ¹²	1 ¹³	30.56
Philosophy	21	3	30	6	2	0	4	2	0	0	2 ¹⁴	0	14.29
Psychology	59	1	0	0	8	0	0	0	1	0	0	0	1.69
School of Social & Political Studies	124	4	5	4	3	0	0	6	0	0	0	0	3.23
Moray House School of Education	151	0	0	0	2	4	9	3	0	0	0	1	0
Total	913	29	195	22	128	54	41	28	13	4	86	19	3.18

¹ Sheet Music, ² Russel Collection Catalogue, ³ Public lectures, ⁴ Data sets/ Newsletters, ⁵ Maps, ⁶ Scottish Witchcraft Database, ⁷ Publication Abstracts Database, ⁸ Newsletters, ⁹ Sound Clips Archive, ¹⁰ Essays, ¹¹ Gadda Journal, Arachnophiles Journal, Christine de Pizan Database, ¹² Undergraduate Dissertation/ Reviews, ¹³ TAAL ePrints server, ¹⁴ Reviews.

Table 3: Research material held on departmental and staff WebPages for the College of Medicine & Veterinary Medicine.

	Staff	S.A Staff	Publications List with link to publishers	Papers	Abstracts only	Unpublished Papers	Book Chapters	Other	% of S.A Staff
School of Biomedical & Clinical Lab Sciences	156	0	130	0	0	0	0	0	0
Biomedical Science	83	0	112	0	0	0	0	0	0
Medical Microbiology	57	0	0	0	0	0	0	0	0
Neuroscience	16	0	18	0	0	0	0	0	0
School of Clinical Sciences & Community Health	>82	1	0	8	114	3	1	0	-
Medical & Radiological Sciences									
<i>Dermatology</i>	32	1	0	8	0	3	1	0	3.13
<i>Medical Physics</i>	20	0	0	0	0	0	0	0	0
<i>Cardiovascular Research</i>	30	0	0	0	114	0	0	0	0
<i>Medical Radiology</i>	n/a ¹								
<i>Respiratory Medicine</i>	n/a								
Community Health Sciences									
<i>General Practice</i>	35	0	0	0	0	0	0	0	0
<i>Public Health Sciences (incl. Med. Statistics)</i>	86	0	0	0	0	0	0	0	0
<i>Health Behaviour & Change</i>	13	0	0	0	0	0	0	0	0
Clinical & Surgical Sciences									
<i>Internal Medicine</i>	n/a								
<i>Locomotor Science</i>	n/a								
<i>Surgical Sciences</i>									
-A&E	n/a								
- Anaesthetics	6	0	0	0	0	0	0	0	0
- Cardiac Surgery	n/a								
- Ophthalmology	n/a								
- Otolaryngology	n/a								
- Surgery	n/a								
- Vascular Surgery	n/a								
Reproduction & Developmental Sciences									
<i>Child, Life & Health</i>	15	0	0	0	0	0	0	0	0
<i>Clinical Biochemistry</i>	n/a								
<i>Genito-Urinary</i>	n/a								
<i>Obstetrics & Gynaecology</i>	7	0	0	0	0	0	0	0	0

¹ n/a denotes no website or information not available on website.

	Staff	S.A Staff	Publications List with link to publishers	Papers	Abstracts only	Unpublished Papers	Book Chapters	Other	% of S.A Staff
School of Molecular & Clinical Medicine	>144	1	0	0	0	0	0	0	-
Clinical Neurosciences	19	0	0	0	0	0	0	Reference database	0
Medical Sciences	<i>n/a</i> ¹								
Oncology	73	0	0	0	0	0	0	0	0
Pathology	<i>n/a</i>								
Psychiatry	52	1	0	0	0	0	0	Conference Poster	1.92
Royal (Dick) School of Veterinary Studies	>94	0	0	0	0	0	0	0	0
Pre-Clinical Vet Sciences	48	0	0	0	0	0	0	0	0
Tropical Vet Medicine	33	0	0	0	0	0	0	0	0
Veterinary Clinical Studies	13	0	0	0	0	0	0	Information sheets	0
Veterinary Pathology	<i>n/a</i>							Links to publisher	
Total	638	2	130	8	114	3	1	-	0.32

¹ *n/a* denotes no website or information not available on website.

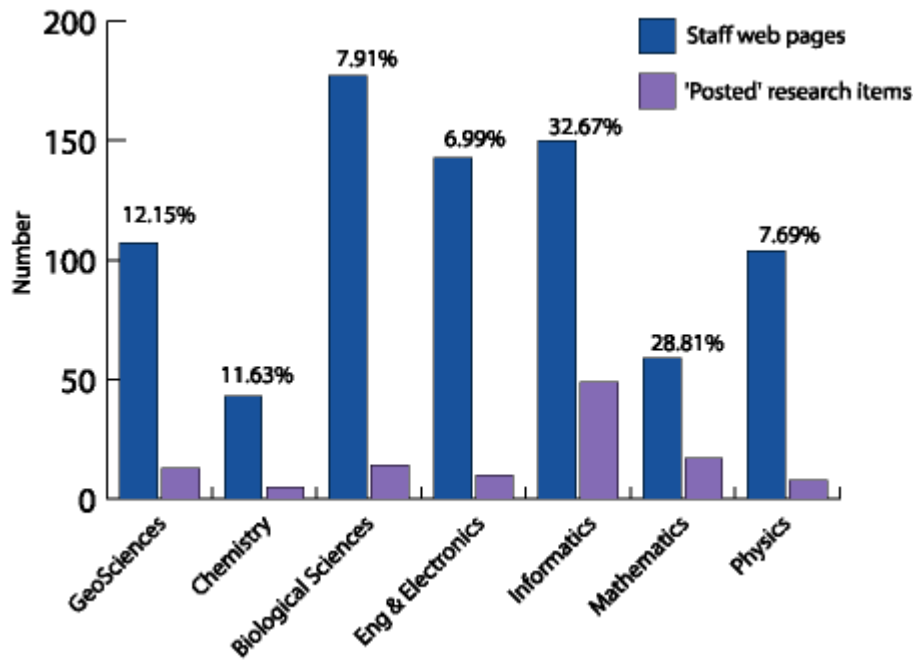


Figure 1 'Self-Archiving baseline' for the College of Science and Engineering.

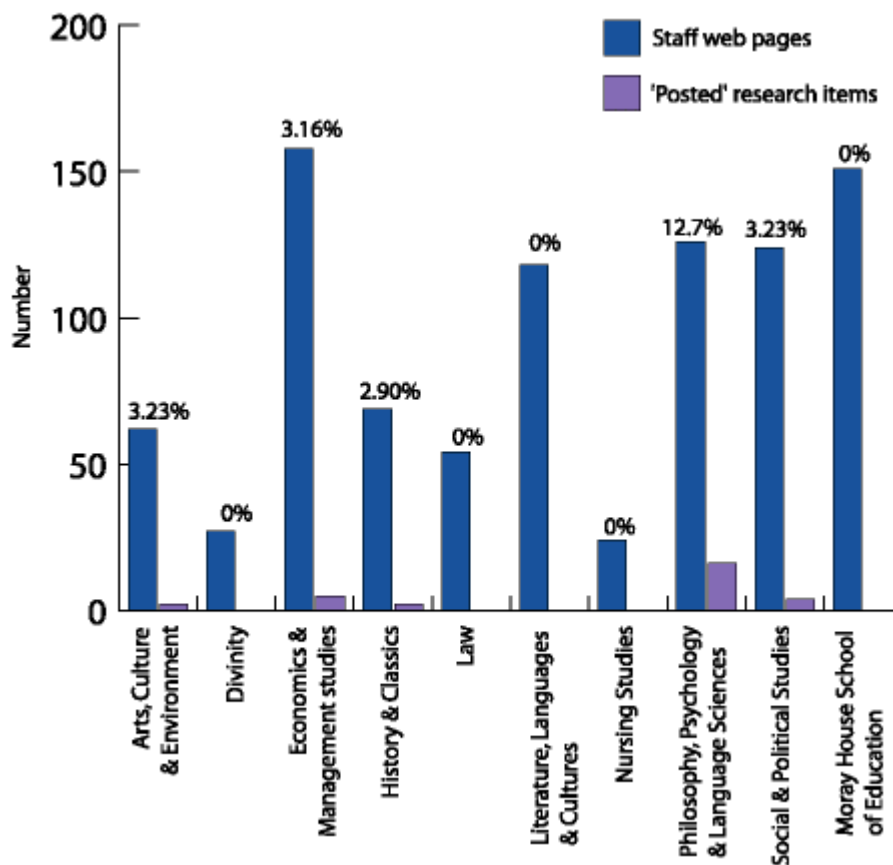


Figure 2 'Self-Archiving baseline' for the College of Humanities and Social Science.

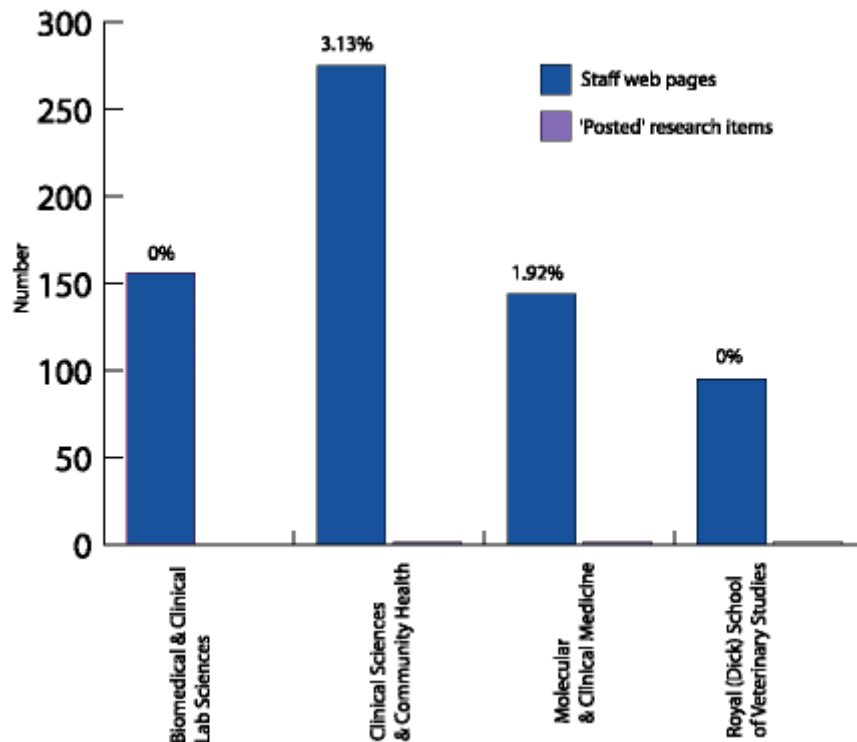


Figure 3 'Self-Archiving baseline' for the College of Medicine and Veterinary Medicine.

3.1 Existing trends between subject areas

Figures 1 through 3 show the volume and percentage of scholars currently self-archiving on personal and departmental websites in each College and School within the University of Edinburgh. As expected, there is a clear difference between academic areas. The average percentage of self-archiving scholars in each College supports this view. Within the College of Science and Engineering (S&E) this figure is 14.81%, which drops to 3.18% within Humanities and Social Science (HSS) and 0.32% within Medicine and Veterinary Medicine (MVM).

However, the situation is more complex than a simple trend of self-archiving being better established in S&E. Looking at the averages between Schools shows that even within Colleges there is a wide distribution of values. In S&E this ranges from 32.67% in Informatics to 6.99% in Engineering and Electronics (Figure 1) and in HSS from 12.70% in Philosophy, Psychology and Language Sciences to 0% in Divinity and Law (Figure 2). As self-archiving is rare in MVM this trend is less well defined; 3.13% in Clinical Sciences and Community Health to 0% in the School of Biomedical and Clinical Lab Sciences (Figure 3).

Even within individual Schools there is a noticeable change in self-archiving attitudes. For example, self-archiving percentages within the School of GeoScience range from 29.41% in Meteorology down to 0% in Geography (Table 1). This tendency is not just restricted to the College of S&E. In the School of Philosophy, Psychology and Language Sciences values span from 30.56% in Theoretical and Applied Linguistics to 1.69% in Psychology (Table 2).

3.2 Volume and type of self-archived research material

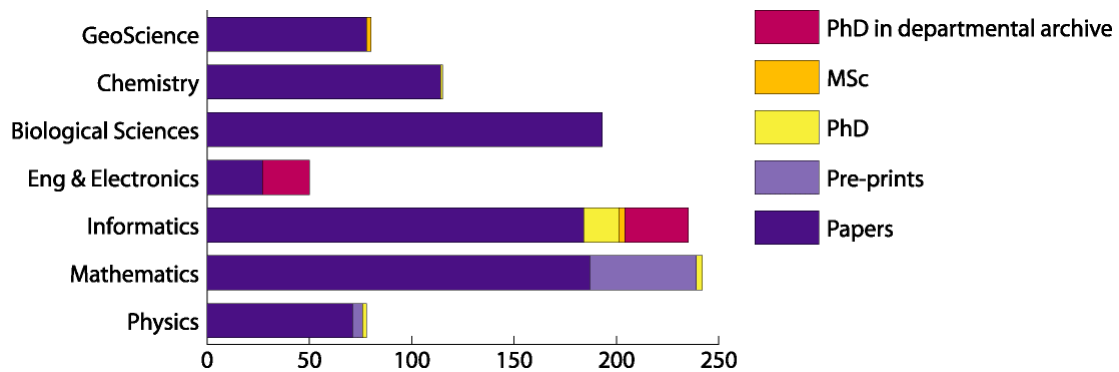


Figure 4 Volume and type of research material presently available in the S&E ed.ac.uk domain.

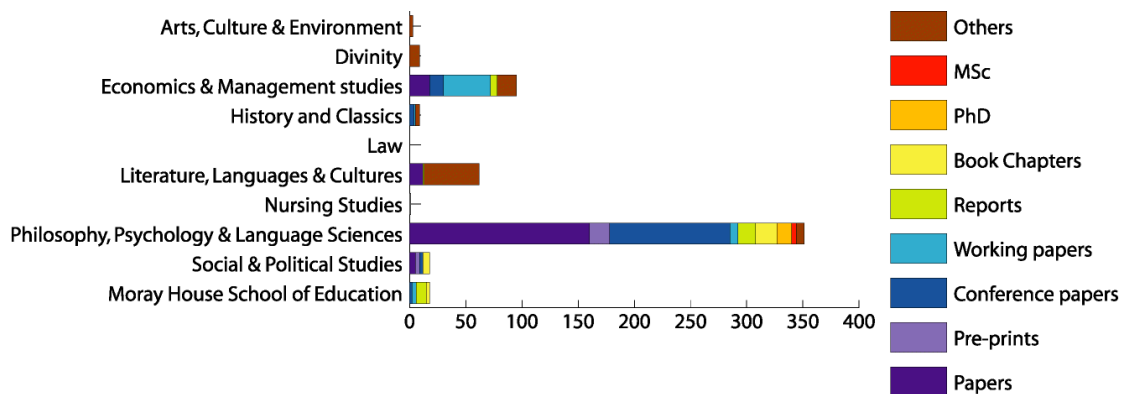


Figure 5 Volume and type of research material presently available in the HSS ed.ac.uk domain.

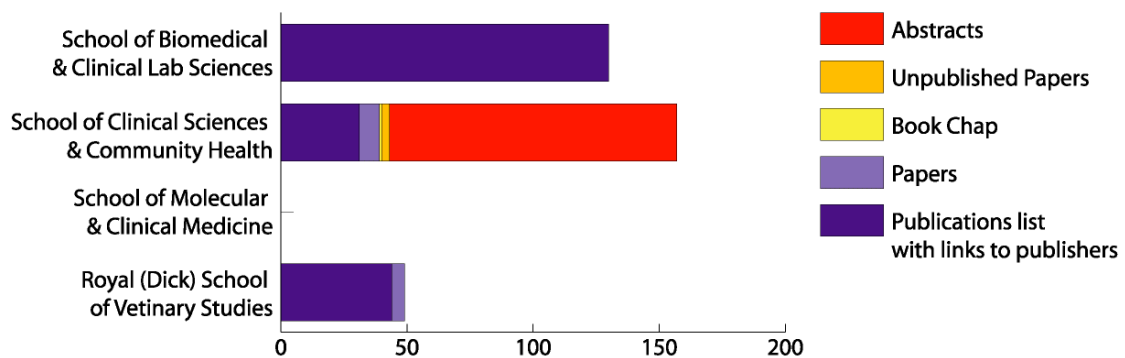


Figure 6 Volume and type of research material presently available in the MVM ed.ac.uk domain.

Figures 4 through 6 show the volume and breakdown of material presently available in the ed.ac.uk domain. In the S&E domain there is a total of 993 separate research items freely available (Table 1). This figure does not include conference papers and technical reports, as they were not included in the original survey for S&E. With these included the figure would be closer to 1500. The majority of surveyed research material, shown in Table 1, consisted of peer-reviewed journal papers (854), commonly in a PDF file format supplied by the publishers themselves. The next largest element of research material freely available was PhD theses (77), followed by preprints (57) and Masters dissertations (5).

As expected, the HSS domain contained a smaller, but still considerable, volume of research material, with 571 freely available items (Table 2). Generally, academic staff in HSS were less likely to self-archive journal papers (195) and PhD theses (13). However the volume of other material, e.g. essays, maps, sheet music, when combined, was significant (86).

Reflecting the trend of low scholarly self-archiving, the MVM domain contained a small amount of actual material. Academic staff in MVM were much less likely to place items online, rather favouring placing abstracts only (114) or citing references hyperlinked directly to publishers' web pages (130). This reluctance is displayed in the fact that only *one* scholar in the entire College of MVM placed freely available full-text journal papers and book chapters online.

4. Discussion and conclusions

Considering the wide-ranging self-archiving trends between academic Colleges and even within Schools, it seems there is a direct correlation between willingness to self-archive and the existence of subject-based repositories. Most of the academic units that have a high percentage of self-archiving scholars already have well-established subject repositories set up in that area. For example, the School of Informatics (32.67%) has CiteSeer¹, Theoretical and Applied Linguistics (30.56%) has Cogprints², Mathematics (28.81%) has the AMS Directory of Mathematics Preprint and e-Print Servers³, Economics (19.23%) has RePEc (Research Papers in Economics)⁴, Chemistry (11.63%) has the Chemistry Preprint Server⁵. The correlation does not hold, however, in the domain in which the world's most successful subject repository exists, Physics (7.69%) has a lower than expected result, despite the well-known success of the ArXiv⁶ subject repository. We would argue that this is because the ArXiv has become so successful in capturing and making persistently available a very high proportion of the output in the domains of high-energy physics and related fields, that academics trust it as their 'natural' repository for self-archived material. The same degree of trust may not yet obtain in the case of the subject repositories mentioned above, which leads to *additional* self-archiving in home institution repositories. So, it appears that, where there is a pre-existing culture of self-archiving eprints in subject repositories, scholars are more likely to post research material on

¹ <http://citeseer.nj.nec.com/cs>

² <http://cogprints.ecs.soton.ac.uk/>

³ <http://www.ams.org/global-preprints/index.html>

⁴ <http://repec.org/>

⁵ <http://www.chemweb.com/preprint/>

⁶ <http://xxx.arxiv.cornell.edu/>

their own web pages, until such time as those subject repositories become trusted for their comprehensiveness and persistence.

A surprising finding from the baseline survey is the relatively low volume of preprints found on personal web pages. This could be related to the success of eprint repositories, such as those described above. Another significant factor is that most papers or theses found online were part of a researcher's publication list in their online CV, which essentially showcases their research interests and credentials. Preprints do not have anywhere near the same impact factor as those papers from accredited journal titles, so it is possible that researchers would favour only putting their most impressive work in their online CV.

One aspect of the survey that is not shown in the results is the lack of consistency in dealing with copyright and IPR issues that scholars face when placing material online. Some academic units have responded by not self-archiving any material at all. A rather worrying example of this is the School of Law (- do they know something that we don't?). A small percentage of individual scholars have responded by using general disclaimers that may or may not be effective. Others, generally well-established professors, have posted material online that is arguably in breach of copyright agreements, e.g. whole book chapters. Most, however, take a middle line of only posting papers from sympathetic publishers who allow some form of self-archiving. It is apparent that if institutional repositories are going to work, then this general confusion over copyright and IPR issues needs to be addressed right at the source.

Summing up on a lighter note, it is extremely encouraging to see that such an unexpected high volume of research material (over 1000 peer-reviewed journal articles) exists online in the ed.ac.uk domain, suggesting that there is already a growing grassroots movement aiming towards freeing up scholarly communication through the use of digital media. The big problem is that this material is widely dispersed and therefore not easily found. This is not very useful for the wider dissemination of scholarly work. Also, personal websites tend to be ephemeral, so the long-term preservation of the research material held on them is extremely doubtful. This is where projects such as *SHERPA* and *Theses Alive!* can step in to help the process by providing a more stable platform for effective collation and dissemination of research. This study has shown that there is already a substantial corpus of research material available online. Contacting the pre-existing self-archiving authors and gathering initial content can overcome one of the main barriers in the creation of a successful institutional repository. The material is already out there; we just have to look for it.

Author details:

Dr. Theo Andrew

Project Officer for Theses Alive! and SHERPA

University of Edinburgh

<http://www.thesesalive.ac.uk/>