

Working Papers, Open Access and Cyber-Infrastructure in Classical Studies

David Pritchard

This is a pre-copy-editing, author-produced PDF of the revised version of an article which has been accepted for publication by *Literary and Linguistic Computing* following peer review. For information on the publisher authenticated version go to:

<http://llc.oxfordjournals.org/>

15 February 2008

Dr David Pritchard
Department of Classics and Ancient History
School of Philosophical and Historical Inquiry
The University of Sydney
NSW 2006
AUSTRALIA

Abstract

Princeton–Stanford Working Papers in Classics is a web-based series of work-in-progress scripts by members of two leading departments of classics. It introduces the humanities to a new form of scholarly communication and represents a major advance in the free availability of classical-studies scholarship in cyberspace. This article both reviews the initial performance of this open-access experiment and the benefits and challenges of working papers more generally for classical studies. After two years of operation *Princeton–Stanford Working Papers in Classics* has proven to be a clear success. This series has built up a large international readership and a sizeable body of preprints and performs important scholarly and community-outreach functions. As this performance is largely due to its congruency with the working arrangements of ancient historians and classicists and the global demand for open-access scholarship, the series confirms the viability of this means of scholarly communication and the likelihood of its expansion in our discipline. But modifications are required to increase the benefits this series brings and the amount of scholarship it makes freely available online. Finally departments wishing to replicate its success will have to consider other important developments, such as the increasing availability of postprints, the linking of research funding to open access, and the emergence of new cyber-infrastructure.

1 What are Working Papers?

1.1 Definition

Working papers are unpublished versions of journal articles, book chapters, book reviews or other research-related manuscripts which are collected and disseminated in a systematic way. They vary in terms of completion from work in progress, which has not been submitted to a journal or publisher, to a manuscript that has been accepted for publication but is not yet revised for its final submission. Authors normally contribute to such series in the hope of getting valuable feedback on current research, which they intend to publish formally. However, those coordinating working-paper series can never guarantee contributions will be published and a minority of them are never peer-reviewed nor make it into print (Kling, 2004, 608-15; Ober *et al.*, 2007, 233). Across the research sector papers circulating in this manner have been variously called grey literature, preprints, research manuscripts or memoranda, technical reports and working papers (Kling, 2004, 596-8). The last is the term of choice among social scientists.

1.2 Invention, Development and Spread

Working papers were invented by scientists several decades before the advent of the internet (Kling, 2004, 596-7). Although they were first circulated by electronics experts at the Massachusetts Institute of Technology (MIT) in 1946, this form of communication was developed most fully by high-energy physicists. In 1962 the first director of the Stanford Linear Accelerator Centre (SLAC), W.K.H. Panofsky, asked his library to begin collecting work-in-progress papers by high-energy physicists from around the world (Kreitz *et al.*, 1997, 24-32). With the assistance of the library of the European Organisation for Nuclear Research (CERN), the SLAC library quickly established manual systems for obtaining, cataloguing and archiving unpublished manuscripts.

As high-energy physicists needed a high level of computer literacy to operate in their computer-dependent field, in subsequent decades they had a leading part in the development of the computer-based management of their preprints, along the way inventing the computerised database, e-mail and the World Wide Web (Kreitz *et al.*, 1997, 26). In 1974 the completion of the Stanford

Public Information Retrieval System–High-Energy Physics (SPIRES–HEP) allowed working papers to be catalogued by computer (Addis, 2002; Kreitz *et al.*, 1997, 24-6). Run collaboratively with the Deutsche Elektronen Synchrotron, this database recorded the bibliographical details and abstracts of working papers and the general subjects they addressed. In 1982 the SLAC library became one of the first worldwide to get rid of its card catalogue, while, by 1985, high-energy physicists in 44 countries had access to its computerised database through 662 locally based servers that were connected to SPIRES–HEP via a precursor of the internet.

Rapid change in this mode of scholarly communication was set in train by two high-energy physicists in 1991 (Kreitz *et al.*, 27-31, 32 n.9). Mid year Paul Ginsparg of the Los Alamos National Laboratory (LANL) finalised a new system for collecting, publicising and archiving working papers electronically (Ginsparg, 1994). Physicists could now send versions of papers written in the newly invented formatting language of TeX, which reproduced mathematical formulae in basic keyboard characters, to an electronic archive at LANL, which was originally called xxx.lanl.gov. There they were grouped according to general subjects and assigned archive numbers. Each evening a list of the details and archive numbers of the working papers received during the day was sent via e-mail to subscribers of a list-server, who were free to transfer electronically from the archive any papers of interest. The librarians of SLAC very quickly automated the inclusion of the bibliographic details of these preprints into SPIRES–HEP and began posting electronically printer friendly versions of them on the database, which they continued to do until 1995, when this function was taken over by the LANL archive itself. In late 1991 Tim Berners-Lee and his team at CERN invented the World Wide Web by combining available hypertext and file-sharing protocols and building the first web-browser and server (Berners-Lee *et al.*, 1992). This was immediately recognised at SLAC as a superior way of providing remote access to SPIRES–HEP and, within a few months, one of its high-energy physicists had set up the first web-server in North America to do so (Addis, 2002). By the end of 1994 most searches of its database were conducted online, while around seventy percent of working papers were submitted electronically.

These two inventions of 1991 steadily increased the number of working papers by high-energy physicists and have brought about a complete change in the way they communicate with each other. The LANL archive has become *arXiv.org* (<http://www.arxiv.org/>), which is probably the best known discipline-based repository worldwide (Kling, 2004, 605, 608-15). Today owned and operated by the Cornell University Library, *arXiv.org* has gone beyond high-energy physics, now collecting working papers from mathematics, nonlinear science, computer science, statistics and other subdisciplines of physics as well. This archive currently holds more than four-hundred-thousand papers, gets tens of thousands of ‘hits’ or connections per hour, and processes more than twenty-million requests for preprints each year. These extraordinary numbers bear out a significant shift of scholarly communication among high-energy physicists: they communicate today largely via working papers, with peer-reviewed articles serving only as the credentials they require to get a post, tenure or a promotion (Ober *et al.*, 2007, 235-6).

At the same time as discipline-wide working papers were being developed in high-energy physics during the 1960s and 1970s, increasing numbers of individual departments and laboratories were setting up preprint series of their own (Kling, 2004, 597-603). The first to do so were in the fields of artificial intelligence, high-energy physics and mathematics. A decade later they began spreading into the social sciences where they appeared most commonly in departments of demography, economics and linguistics. In the mid to late 1990s many of these series were moved online, as the web became a cheaper and easier platform for distribution. In so doing departments made their preprints available to everyone with a browser and internet access, expanding profoundly the pool of potential readers. Strikingly however, the internet has not significantly expanded the number of disciplines using working papers, while a majority of university departments still do not have preprint series. In this area the humanities continues to perform particularly poorly, lagging well behind the sciences and social sciences in its use of cyberspace to circulate work-in-progress or published research (Houghton *et al.*, 2006, 54). As such the

Princeton–Stanford Working Papers in Classics (hereafter *PSWPC*) probably is the first department-based series of working papers in the humanities (Poynder, 2005).

2 The Princeton–Stanford Working Papers in Classics Series

2.1 Inception

The original idea for creating a working-paper series in classics and related disciplines came to Josiah Ober, as a result of his service on a university-wide committee for selections and promotions at Princeton University (Ober *et al.*, 2007, 231, 233). There he encountered listings of so-called working papers in the curricula vitae of social-science candidates. The operation of preprint series and their benefits were explained to him by another committee member, Gene Grossman, who had set up a department-based series of working papers in economic theory during the 1980s. As a longstanding advocate of the importance of social-science approaches for classical studies, Ober quickly grasped the potential of preprints for our discipline. Plans for setting up such a series were developed when Ober, as a visiting researcher at Stanford University, discussed this form of scholarly communication with Walter Scheidel, who knew of working papers because of a long research interest in demography. Scheidel pointed out how running such a series as a collaboration between their respective departments would enhance its benefits: the pooling of their working papers would attract a broader readership, while increasing the number of series-organisers would lower the administrative burden falling on any one individual. Scheidel has subsequently become the academic coordinator of the series at Stanford and its best represented author, with his working papers representing a fifth of its current offerings. Within the Department of Classics at Princeton University two colleagues have helped them bring their plans into effect. Brent Shaw has promoted the initiative publicly and serves as its other academic coordinator. As the department's IT manager Donna Sanclemente, using general web-design software, met her brief to set up an online series as quickly and cheaply as possible, with the *PSWPC* website going live in December 2005.

2.2 Expectations

The *PSWPC*-creators decided to run their initiative initially as an experiment in order to determine the viability of working papers in classical studies, get a better idea of costs and benefits, and receive feedback from users before deciding on its final form (Ober *et al.*, 2007, 230-1, 238-9). This desire of theirs to test the waters is easy to understand: web-based series of preprints normally thrive only when they are in tune with discipline-based working arrangements and perceived requirements and follow upon paper-based practices dating back decades (e.g. Kling and McKim, 2000). For example, although dispersed at three-thousand or so departments and laboratories around the world, high-energy physicists have long valued the speedy dissemination of new work as vital for building on each other's research, avoiding blind alleys in the development of theory and the costly duplication of experiments, and publicly registering new ideas as one's own (Kreitz *et al.*, 1997, 25-6). Consequently they have perceived the time lag of traditional publishing to be detrimental to their work and hence embraced working papers as a form of scholarly communication. By contrast, the practices and perceptions of the humanities are very different to those of high-energy physics and the other disciplines with working papers, a situation which renders the introduction of such a series into classical studies inherently risky: time and resources might be poured into the required cyber-infrastructure but there is no guarantee ancient historians and classicists will actually make use of it by posting their research-related scripts (Lane, 2006).

In these circumstances the *PSWPC*-creators understandably entertained only modest expectations about its initial performance. They only expected to get a few dozen or so working papers (Sanclemente, personal correspondence). They thought most of them would be downloaded by students and academics, while the series would only serve limited scholarly functions: it would cut down the time lag between the writing up of research and its availability to others and allow authors of work in progress to get helpful feedback (Poynder, 2005).

2.3 Design

The website of the *PSWPC* is visually appealing and easily navigated (<http://www.princeton.edu/~pswpc/index.html>). The home page of each department displays a hyperlink to the *PSWPC*-website, which is hosted by the computer server of the Department of Classics at Princeton University (Ober *et al.*, 2007, 230). As ancient historians and classicists are unfamiliar with this method of scholarly communication, the introductory web pages appropriately explain what working papers are, how they can be used, and the way they are collected. In these pages readers are granted permission to use its preprints in their own scholarship, given a suggested form for citing them, and encouraged to send feedback to the authors of individual papers via e-mail. The website allows readers to view its current offerings by individual author, the department to which contributors belong, the year papers were collected or the general subjects each working paper covers. Each of these viewing options generates a list of bibliographical details or so-called metadata paper by paper. Each dataset consists of a unique series number, the title of the working paper, the name of its author, his or her department and university, and an abstract of up to 150 words. By clicking the PDF icon beside each entry readers download a printer friendly version of the work-in-progress paper.

The introductory web pages also provide instructions for members of the Princeton and Stanford departments wishing to contribute manuscripts to the series. For a contributor the first step is the downloading of a standard cover page, on which they provide the metadata and version number of the manuscript, the date of the version's completion, and his or her e-mail address. The next steps are the attaching of this sheet to a soft copy of the paper and the converting of the file so created into a PDF, which is e-mailed as an attachment to the relevant departmental coordinator. Within this e-mail the author indicates which one or more of the 13 general subjects the series recognises he or she wishes the preprint to be listed under. The coordinator 'checks to see that the formatting is correct' before sending the PDF to the IT manager of the Princeton department, who

'mounts the file, ensuring that it is properly cross linked by author, date, institution, and subject area' (Ober *et al.*, 2007, 230 n.5). Since authors retain the copyright of their manuscripts, at any time they can request a preprint of theirs to be replaced with an updated version or removed from the website. Even in its experimental form the *PSWPC* series compares favourably to the established working papers of other departments, which usually do not provide explanations of what preprints are nor multiple viewing options.

3 The Performance of the *PSWPC* Series

By now this series has been running long enough to evaluate reliably its performance and the viability of working papers as a method of communication in classical studies. Evaluating the *PSWPC* is made easier by the range of usage data which its creators have made available. After two years of operation their experiment has proven to be a largely unqualified success: the series has exceeded initial expectations and performs important scholarly and community-outreach functions. In both departments very good numbers have taken advantage of this web-based initiative. As of 11 November 2007 the series has 99 working papers by 33 contributors. Eight of its preprints have been withdrawn because of their final publication in an edited collection or journal, while another 11 have been replaced with revised versions. As such the sponsoring departments are already making freely available a large number of work-in-progress papers, whose total extent is the equivalent of 8 or more edited collections. Around 70 per cent of these preprints may be by Stanford-based researchers, but the percentage of participating faculty there is some 25 points lower than at Princeton where 63 per cent of their colleagues have working papers.

The series has also built up a large and apparently diverse readership. In December 2006 when the downloads of its working papers were first measured, they were already running at 1750 per week (Ober *et al.*, 2007, 232). They were 1910 per week when counted again in September 2007 (Sanclimente, personal correspondence). A large number of these requests have come from universities and colleges across the United States. In turn a good number of academics around the

world are already e-mailing back comments; for although some *PSWPC* contributors are still waiting for feedback, others have received from those researching similar topics valuable tips, bibliography, ancient testimonia, and, at times, unpublished papers (Scheidel, personal correspondence). Clearly these working papers are fulfilling the two scholarly tasks which the creators of the series initially foresaw: they are making available to ancient historians and classicists a lot of research in advance of final publication and successfully soliciting feedback for their authors.

Yet the *PSWPC* is also performing other important functions, which were not fully anticipated at its inception. Tellingly the connections to its website not resulting in downloads has remained several times larger in number than the requests for preprints; for when the total number of hits were measured in November 2006, they came to a surprisingly large 1600 per day (Ober *et al.*, 2007, 232). As of September 2007 they were running at the lower rate of 1100 per day (Sanclemente, personal correspondence). Although many of these connections have been made by private networks and private sources within Australia, Britain, Canada and the United States, the sum of the hits from these larger Anglophone countries is balanced by those coming from developed and developing countries from across the globe (Ober *et al.*, 2007, 232-3). Admittedly some of the connections from non-university sources are presumably coming from academics and students who are using their own internet service providers. However, their apparent volume and the high ratio of hits to downloads suggest members of the general public are also browsing the website. As such the *PSWPC* contributors are probably picking up a wider range and a greater number of readers than they would have done if they had continued to have published only in scholarly books and journals. Consequently this series is providing two other benefits which its sponsoring departments did not anticipate: it is raising their international profiles higher and increasing the so-called research impact of their members.

4 The Reasons for Its Success and the Viability of Working Papers

4.1. Meeting of Disciplinary Requirements

There appear to be four significant reasons why the *PSWPC* has already secured large numbers of contributors and readers. Two of these concern its congruency with practices and perceptions within classical studies. Although ancient historians and classicists collaborate with others and co-author publications much less frequently than scientists and social scientists, they have always valued catching up on work in progress by colleagues and receiving informed feedback on their own. But because of the modest size of our departments and their dispersal worldwide, communication with other specialists has required long-distance correspondence, which, before e-mail, was done by individual scholars posting each other hardcopies of papers and memoranda (Kling, 2004, 596; Ober *et al.*, 2007, 234). Moreover, the sharing of work in progress has long been the formal justification for our attending or convening of conferences. Consequently this preprint series might not enjoy a specific paper-based precedent but it clearly builds on longstanding practices of our discipline and provides a good way for meeting one of its perceived needs: the sharing of in-progress research among specialists.

4.2 Standing of Computers in Classical Studies

A second reason for its success is the entrenched use of computers by ancient historians and classicists, which has positively disposed our discipline more than others in the humanities to computer-based innovations. Many outside of classics are surprised to learn of its leading role in computing and the widespread use of computer-based tools by its philologists for two decades. The relationship between classics and the computer began in 1949 when Robert Busa convinced the president of IBM to collaborate on the production of the first complete concordance of the Latin writings of Thomas Aquinas (Hockey 2004, 4-7). The subsequent history of computers within classical studies and the multiple ends to which they have been put are well studied by others (e.g.

Brunner, 1993; Crane, 2004). As such this article focuses on two of the most important developments for the shaping of perceptions of computers among ancient-world experts, namely the integration of computerised corpora of primary sources into classical philology and the heavy reliance on a so-called open-access archive of reviews of classical-studies books.

In 1972 the Thesaurus Linguae Graecae (TLG) Project was set up by Theodore Brunner of the University of California at Irvine, who would serve as its first director for twenty-five years (Berkowitz, 1993; Brunner, 1987; Hughes, 1987). The goal of this project was the producing of a computer-readable database of the surviving literature of the ancient Greeks from Homer to the seventh century of our era. The possibility of this had been opened up by the invention of an ASCII-based code by David W. Packard a few years earlier, which reproduced Greek letters, accents and breathings in basic keyboard characters (Packard, 1973). As an accomplished classicist and the son of a co-founder of the Hewlett Packard Corporation, Packard has remained a central figure in the integration of computers into classical studies for decades. While the TLG Project had access to his code, the data processing of more than a millennium of literature was a truly herculean task: 60 million words of Greek had to be entered manually on IBM punch cards, consolidated work by work onto magnetic tape, and checked for accuracy by a team of classicists (Brunner, 1987, 1, 6-7). All up this took over 15 years and cost 6 million US dollars, with the TLG first released in tape and CD-ROM formats (Hughes, 1987, 4-5).

During this long period of production the interest of classicists in this computer-based corpus grew strongly, as increasing numbers realised its potential for transforming the traditional practices and scope of philological research (Brunner, 1993, 17-27). However, as few outside of the TLG Project had the expertise, hardware and mainframe access to run such a searchable database, this potential would not have been realised as quickly as it was if not for the ongoing work of Packard in the 1970s. He modified the design of the new Hewlett-Packard Minicomputer and its operating system to create a portable computer, which could display and edit Greek, Hebrew and Latin scripts and search the TLG and other databases (Crane, 2004, 48-8; Hughes, 1986). The first

commercialised version of what Packard called the Ibycus was showcased at the annual conference of the American Philological Association (APA) in 1979. Despite costing over 10 thousand US dollars, a dozen or so departments quickly purchased one. By the time of the release of the third CD-ROM of the TLG in 1987, the cost of the Ibycus had fallen considerably to 4000 US dollars, allowing 200 or more other departments worldwide to make use of this database (Brunner, 1987, 9; Hughes, 1987, 6). The spread of computer-aided philology right across our discipline was ensured by two related developments of the late 1980s. Gregory Crane of Harvard University and others developed new programs for running computerised corpora on any general purpose computer, such as the Apple Macintosh or IBM PC, while the Packard Humanities Institute produced and marketed CD-ROMs of classical Latin literature (Brunner 1993, 27; Crane 2004, 48-9). By 1993 around 1400 institutions worldwide had paid a licensing fee for use of the TLG CD-ROM (Brunner, 1993, 33 n.76). The distribution of these disks ceased in 2001 when the TLG was successfully moved online (<http://www.tlg.uci.edu/>).

The dissemination of the TLG and other textual databases not only impacted positively on philology but quickly confirmed the utility of computing within classical studies. As a result ancient historians and classicists as whole have been well disposed to subsequent computer-based innovations, whether they have been in research, teaching or scholarly communication. Of these the most important has probably been the *Bryn Mawr Classical Review* (hereafter *BMCR*), which is the second oldest electronic journal in the humanities (<http://ccat.sas.upenn.edu/bmcr/>). Established by Richard Hamilton and James O'Donnell at Bryn Mawr College in 1990, this journal began distributing its contents exclusively via e-mail (Crane, 2004, 48; Hardwick, 2000, 285). Each review was submitted electronically, refereed and edited within only a few months, and sent within an electronic message to the subscribers of the journal's list-server. For the sake of speeding up scholarly communication the e-mail address of the reviewer was normally included, while the journal published the formal exchanges of what could be very colourful responses between reviewers and authors. Several months after the invention of the web by high-energy physicists in

late 1991, *BMCR* set up an internet-based archive where printer friendly versions of reviews could be posted initially and remain available for readers to search either via the date of their publication or the name of the reviewer or author. A few years later the hyperlink to each review's web page was included in the email versions sent to subscribers.

In its first full year of operation *BMCR* garnered some 300 subscribers and 124 reviews (Hamilton, personal correspondence). Numbers steadily rose, however, as more and more scholars came to realise the advantages of this computer-based initiative: its free and easy circulation, the utility of its archive, and its publishing of reviews in only a fraction of the time print-based journals needed. Five years later, in 1995, the journal had 1342 subscribers and 256 reviews. Since then the rate of growth has not abated: by 2006 *BMCR* had around 9000 subscribers and published 572 reviews, while its website processes extraordinary numbers of daily requests for archived reviews. In October 2007, for example, requests were running at 8892 per day, with the grand total of annual requests to the end of this month coming to 3,371,608.¹ Although a portion of the web pages which are requested are probably never read, these numbers point strongly to the integration of this open-access archive into the daily work-routines of ancient historians and classicists. As such *BMCR* has probably done more than any other computer-based initiative to habituate our discipline to the free availability of scholarship on online and hence to lay the ground for the success of the *PSWPC*.

4.3 Demand for Open-Access Research

Admittedly a good proportion of connections to the websites of the *BMCR* and *PSWPC* and to other open-access initiatives of classical studies are not made by academics and their students. Indeed we have reason to believe there already is solid interest in this series of working papers from outside of the research sector. In fact, this happens to be a third reason for its success. By making its offerings available to anyone with a browser and internet access the *PSWPC* helps satisfy worldwide demand for open-access research on antiquity. Many of these so-called interested readers are simply members of the public; as the rediscovery and appropriation of Graeco-Roman art and literature have long been mainstays of modern cultures, there is an abiding popular interest

in the subject matter of our discipline, with the curiosity of some piqued enough for them to seek out 'hard' information on the ancient Mediterranean. Serious interest also comes from our former students as well as secondary-school pupils and their teachers (e.g. Doherty, 2005).

This general interest in the ancient world may be strong but a sobering reality is that our lay readers have relatively limited access to our written research (Houghton *et al.*, 2006, 33-4). The borrowing of books and edited collections from university libraries are preserves usually of salaried academics and enrolled students, while most outside of this circle do not have the opportunity or wherewithal to access these collections by walking in. Consequently members of the public with a serious interest in our discipline largely rely on their own purchases of books, the usually slim pickings of community or school libraries and whatever internet-searches throw up. For the last several years the staff and students of universities with well-resourced research libraries have been able to download articles from leading journals via local networks, which might give the impression that a lot of our scholarship is freely available in cyberspace. In fact, such requests are only possible because of the annual subscriptions our libraries pay for so-called bundles of titles in electronic format (Cox, 2006; Houghton *et al.*, 2006, 3-4). Indeed classical studies, like other disciplines in the humanities, provides open access to only a tiny fraction of its published research (Suber, 2005, 234-6). Fifteen years after it became possible, only 8 of the close to a thousand journals in our discipline are freely available on the internet.² In these circumstances the *PSWPC* is clearly helping to meet unmet demand for open-access scholarship on the ancient world.

So far the success of the *PSWPC* has been explained in terms of the contexts of its operation: the first two reasons for this concern its congruency with the working arrangements and perceptions of classical studies, while the third focuses on its meeting of the global demand for open-access scholarship. Critically these reasons relate more to the means of scholarly communication of which these preprints are an example than they do to any unique features of the *PSWPC* itself. Consequently the success of this experiment strongly suggests this is a viable form of scholarly

communication in classical studies and that the number of departments setting up similar series will steadily increase.³

4.4 The High Quality of the Working Papers

The last significant reason for the success of the *PSWPC* is due to the quality, number and subject matter of its offerings. As two of North America's wealthiest and most prestigious universities Princeton and Stanford boast large departments of classics and some of the world's foremost ancient historians and classicists. As a result, their working papers are generally of a very high standard, contain important findings and insights, and are very likely to make it into print. Together the two departments have been able to make a sizeable number of preprints available, most of which canvass the more popular subjects and periods of our discipline. These characteristics go a long way towards explaining the growing readership of the *PSWPC*: it is providing a wide range of academics and postgraduates timely and free access to up-to-date and high-quality scholarship on which they can productively build their own research.

5 Improving the *PSWPC* Series

Clearly the *PSWPC* has proven to be a success and confirms the viability of preprints in classical studies. After two years of operation its creators can also be given feedback as they mull over its final form. In particular they should seriously consider improving the coordinating of its metadata and website and the scope of its interoperability. Such improvements would make the *PSWPC* easier to use and increase the positive publicity it brings its sponsoring departments and the quantity of scholarship to which it provides open access.

5.1 Metadata and Website

Glitches bedevil the metadata and website of the *PSWPC* in its current form. Many of its abstracts are considerably shorter or longer than the maximum set extent of 150 words and a proportion do not summarise their papers satisfactorily. Completeness and consistency of publication details, felicity of expression, and spelling remain problems across the metadata.

Technical problems are apparent on the web page listing the names of authors: some of its hyperlinks are malfunctioning and this list as a whole gives access to a smaller number of authors and working papers than the corresponding drop-down menu on the home page. The ameliorating of these issues is important: in contrast to the work-in-progress papers, this metadata exists only because of the series and is essential for the wide circulation of its contents. In addition we have seen how these working papers are probably attracting large numbers of non-academic readers, for whom they are presumably serving in many cases as their first contact with the Princeton and Stanford departments. Such readers would get a better impression of the relative standing of these departments if there were fewer glitches, which could be achieved through small changes to the work plan of the *PSWPC*: its academic coordinators could check the correctness not only of the formatting of papers but also their metadata, while its web master might spend more than the few hours currently being devoted to site maintenance each week (Ober *et al.*, 2007, 230 n.5; 231 n.6).

5.2 Interoperability

Serious redesign of the website is required to ensure its compliance with the Open Archives Initiative–Metadata Harvesting Protocol (OAI–MHP), which is vital for enhancing both the interoperability of this series and the value of this form of scholarly communication as it emerges in our discipline. The initial impetus for this data-sharing protocol arose again in high-energy physics: in 1999 members of LANL and other labs met at Sante Fe (New Mexico) to canvass ways of aggregating the metadata of all scripts held by disciplinary archives and the growing number of department-based series so that they could be browsed more easily (Lynch, 2001). However, as electronic working papers and other digital resources had spread into several other disciplines during 1990s, it was soon apparent theirs was a challenge which many beyond high-energy physics also faced. Consequently, in the following year, a number of research organisations provided funding for the so-called Open Archive Initiative (<http://www.openarchives.org/>). Headquartered initially at Cornell University, this group developed the OAI–MHP, a protocol which basically consists of two parts. Firstly, metadata needs to be formatted in the 15-element dataset of Dublin

Core, which is the internationally accepted standard for describing working papers and other digital resources (Hardwick 2000, 285). Secondly, the archive they are stored in needs a web interface which automatically provides metadata via pre-existing hypertext and file-sharing protocols, whenever requested by another machine.

The integration of the OAI-MHP into archives as a matter of course has spawned the creation of so-called metadata-harvesters, such as OAIster (<http://www.oaister.org/>), at the University of Michigan, which 'deal with all the separate, compliant archives, as if they formed one grand virtual archive' (Suber, 2005, 233). At the time of writing OAIster has 'harvested' the metadata of some 12 million digital objects from close to 900 repositories worldwide. This protocol also allows Google Scholar and other commercial search engines to collect and display systematically the descriptions of research-related scripts. As a result, web-users 'can now find articles in OAI-compliant archives even if they don't know which archives exist, where they are located, or what they contain' (Suber, 2005, 233).

The *PSWPC*-creators hoped this problem of fragmented holdings which are not easy to search would be avoided if one of our large professional organisations set up a discipline-wide repository for working papers, such as *arXiv.org*, or a website where different series of preprints could be searched simultaneously (Ober, 2007, 233, 238-41). However, this hope is not particularly well founded. Firstly, although there are already a handful of advocates for all-encompassing repositories in our discipline, working papers will surely need to become much better established before the American Philological Association, the Australasian Society for Classical Studies or some other professional body commits the resources and personnel for establishing and maintaining such a series. Secondly, the aggregating of the metadata of our preprints will only be possible if the *PSWPC* and all subsequent department-based series incorporate the OAI-MHP. The integration of this protocol into our preprint series has the added advantage of allowing simultaneous searches of our papers via pre-existing harvesters before any search engine is set up specifically for classical studies.

6 Concluding Remarks: Towards Open Access in Classical Studies

Clearly the *PSWPC*-creators drew the idea for their experiment from the communication practices of social-science departments, which dated back to the 1970s, rather than from more recent developments in so-called open access (Poynder, 2005). As departments planning to set up their own series of electronic scripts will need to take into account this shift across the research sector towards freer access to our scholarship, my concluding remarks focus on the implications of three important developments of the last several years in this area.

6.1 Increasing Availability of Postprints

Open access is formally defined as the making freely available to every web-user both digital objects and the copyright permission to download these at a minimum for personal use (Suber, 2005, 231). Such access was put on the agenda in the mid-1990s when Stevan Harnad made a well-publicised proposal for scholars to abandon print-based communication immediately by making as many of their scripts as possible available in cyberspace (Harnad, 1995). This cognitive scientist from the University of Southampton has since become the leading figure in a loosely organised coalition of academics and librarians campaigning for free forms of scholarly communication, which is called the open-access movement (<http://www.eprints.org/openaccess/>). In 2000 this movement successfully convinced 34 thousand scientists to boycott journals which refused to give open access to articles within 6 months of their publication (Kling, 2004, 591-2). Concerned about the impact of doing so on their subscriptions, publishers refused to comply, which ultimately led to the collapse of the boycott. Since then, however, growing calls for open access have resulted in many journals giving a lot of ground. Today around 90 percent of academic publishers provide web-based access to journals for subscribers, while more than 60 percent allow authors to make freely available on the net the final versions of their articles as soon as they are published in print (Cox, 2006, 273-5).

This new availability of postprints allows ancient historians and classicists to take the initiative in expanding open access to their published research as long as they have a suitable platform for self-archiving. The initial performance of the *PSWPC* strongly suggests that a department-based series for electronic scripts would indeed be a viable way to disseminate published articles on the web: in general, members of our discipline seem willing to use such a series for posting and downloading research-related papers, while it would appear to increase the impact of research in and outside of academia.

The setting up of department-based websites for working papers and postprints could also help to breakdown other discipline-specific obstacles to open access. In contrast to the situation of many other disciplines a large number of our journals are published by small professional bodies or individual departments, whose limited budgets permit few or no specialist editors (Ober *et al.*, 2007, 237). Consequently they have found it more difficult than university or commercial presses to keep apace of recent changes in scholarly communication. A case in point is the situation of classical-studies journals in Australia and New Zealand: of the 9 journals in this region 2 only have an electronic version available to subscribers, 3 allow the self-archiving of postprints, 5 others have no policy on this, and 1 rules it out entirely.⁴ Another obstacle is that although the journal article is the most important publication for career-building in the sciences and social sciences, in the humanities this role is shared by monographs and edited collections, out of which publishers seek to make profits (Suber, 2005, 237). As a result they do not volunteer to make books freely available in cyberspace, while we as authors are reluctant to broach this possibility in view of the benefits accruing to us from final publication. If a growing number of our departments set up their own series for working papers and postprints, our discipline as a whole would probably become more fully aware of the benefits of web and open access. Likely consequences are that smaller publishers of journals would be asked more and more frequently to ameliorate their policies on web access and self-archiving and that increasing numbers of authors would be emboldened to ask for open access

to at least a portion of their books. In time, these requests would no doubt expand the proportion of our published research which is free and online.

6.2 Linking of Funding to Open Access

Departments of classical studies contemplating open access need to take into account two more recent developments. The first concerns the move towards the linking of funding to open access. In the last few years governments and funding bodies have commissioned reports on the pros and cons of making research funding dependent on the posting of the results of funded projects in web-accessible repositories (Suber, 2007a). One such report by John Houghton, Peter Sheehan and Colin Steele was released by Australia's Department of Education, Science and Training in September 2006. The advantages of open access identified in this report are very similar to the benefits which have been canvassed for working papers (Houghton *et al.*, 2006, 31-4). Open access reduces the unnecessary duplication of research, speeds up the process of inquiry, and allows researchers to build more productively on the research of others. By making research available to those outside of the small circle with borrowing rights at well-resourced research libraries, it also improves the quality of policymaking, the practice of professionals, such as high-school teachers and doctors, and the knowledge and hence decision-making of the general public. Among its recommendations are that national funding bodies make the provision of open access to funded research compulsory and that each university introduce a similar policy for the research every staff member produces (58-60).

Many will be surprised to learn that some of these recommendations have already been adopted by the Australian government: the Australian Research Council (ARC) and the National Health and Medical Research Council (NHMRC) now ask the researchers they fund to place their data and results in a disciplinary archive or so-called institutional repository wherever they exist (ARC, 2007). In addition both councils require grant-recipients who do not provide open access to spell out the reasons for not doing so in their final funding reports. Although the ARC and NHMRC have not yet made this self-archiving compulsory, the intensity of competition for their funds and

the required inclusion of final reports in new grant applications will ensure a high level of compliance. Moreover, as Australian universities usually model their own research policies on the funding rules of these councils, this new open-access requirement is likely to lead to internal pressure on academics to deposit their working papers, postprints and digitised data into university-based archives. This linking of funding to self-archiving is taking place around the world: in 2006 alone comparable soft mandates were introduced by the national funding bodies of Austria and Germany and the Research Councils of the United Kingdom (Suber, 2007a). In the United States of America the National Endowment of the Humanities now favours grant applications which promise open access for research results, while Congress, in spite of serious lobbying from publishers, has recently passed an appropriations bill which makes funding from the National Institutes of Health dependent on compulsory self-archiving (Suber, 2007b).

6.3 New Cyber-Infrastructure

At first glance this move to mandated self-archiving seems another good reason for our departments to follow the lead of Princeton and Stanford. In addition to the open-access benefits the *PSWPC* confirms, department-based websites for preprints and postprints would surely help ancient historians and classicists to meet this new condition of funding. However, the situation is complicated by the emergence of a new and possibly superior platform for open access: the institutional repository. This is an open-access archive where the members of a university or research centre can deposit their working papers, published articles and, depending on its software, other research-related resources (Lynch, 2003). It is usually managed by a university library, which takes care of cataloguing and compliance with OAI-MHP and provides self-archiving academics advice on copyright and IT issues. The setting up a repository also entails a long-term commitment on the part of an institution to preserve its contents, which includes the upgrading of software and the transferring of data into different formats, once old file types become obsolete.

This third open-access development was made possible by a number of events in the early years of this millennium. In 2002 Raym Crow of the Washington-based Scholarly Publishing and

Academic Resources Coalition (<http://www.arl.org/sparc/>) released an influential paper on the ways such a platform could catalyse new forms of scholarly communication and provide to its sponsoring institution significant benefits in terms of publicity and research impact (Crow, 2002). Putting one up had also become feasible because of the new OAI-MHP, the exponential decline in storage costs, and the free availability of software for its management. In 2002, for example, MIT and the Hewlett Packard Corporation launched Dspace (<http://www.dspace.org/>), which became open-source software a year later. Around the same time the University of Southampton also made freely available Eprints (<http://www.eprints.org/>), which Harnad had developed for the managing of working papers and postprints. Since then a large number of institutional repositories have been set up: in Australia 19 of its 39 universities now have this cyber-infrastructure.⁵ The percentage of universities in North America with repositories may be lower but here and elsewhere they are going live at an increasingly rapid rate (Suber, 2007a).⁶

This proliferating of institutional repositories raises serious questions about the best path towards open access in classical studies. We may have sound reasons for making our research-related scripts freely available in cyberspace and be coming under pressure to do so. But is a departmental website now a second-best option for open access? Despite the initial success of the *PSWPC*, has this platform been superseded by the university-wide archive? On balance both of these questions should probably be answered in the negative. The department-based series and the institutional repository are not mutually exclusive and can complement each other. For some classical-studies departments the approach taken by the *PSWPC*-creators will remain the best means of open access in the short to medium term, as they wait for their institutions to set up repositories. For most departments, however, it is the institutional repository itself which will make feasible the circulation of their research-related scripts in cyberspace. The design of many pre-existing archives, such as the Sydney eScholarship Repository (<http://ses.library.usyd.edu.au/>), in fact involves departments in the coordinating of their series. As a Dspace-based archive the Sydney repository does not accept submissions directly from individual academics. Instead this function is given to

each department or centre of the university wishing to give open access to their scripts or other digital resources. The repository supports a department coordinating such a series by providing for it a plan for work flow which guarantees OAI–MHP compliance and the proper copyediting of metadata before it is placed on the web. In the repository each department gets its own introductory web pages where statements about the department and its series can be displayed. Here too web-users can search the contents of each series by author, date, title and customised general subjects.

The Sydney eScholarship Repository gives individual departments the infrastructure, work plan and technical support to set up their own series of working papers and postprints, which have the same functions as the *PSWPC*. Admittedly departments cannot customise the design of the web pages this archive gives them. But this open-access service is provided free of charge and includes an undertaking to preserve submitted files in perpetuity. Moreover, if it has the wherewithal, a department can lay a custom-built website over the repository, incorporating the metadata and search options of the latter and providing hyperlinks to its securely archived files. As such once Princeton University or Stanford University sets up an institutional repository, the *PSWPC*-creators could, if they wished, keep their website but migrate its contents to this cyber-infrastructure where they would have the central support they deserve for long-term preservation and OAI–MHP compliance. Finally the complementary relationship between these two platforms is not one way; for department-based series help solve an ongoing problem with institutional repositories: the reluctance of academics to use such university-wide initiatives (Houghton *et al.*, 2006, 8; Suber 2005, 8). The success of the *PSWPC* in getting a very high rate of participation among academics demonstrates the importance of involving departments in open access. Where the circulation of preprints and postprints is congruent with discipline-based perceptions and practices, researchers appear to be much more willing to be part of a series which is set up and managed by their own department.⁷

References

- Addis, L.** (2002). Brief and Biased History of Preprint and Database Activities at the SLAC Library, 1962-1994, SPIRES–HEP. <http://www.slac.stanford.edu/spires/papers/history.html> (accessed 24 March 2007).
- Australian Research Council.** (2007). Media Release: ARC and NHMRC Encourage Access to Research Findings, Canberra: Australian Government. http://www.arc.gov.au/media/releases/media_18Jan07.htm (accessed 25 June 2007).
- Berkowitz, L.** (1993). Ancilla to the Thesaurus Linguae Graecae: The TLG Canon. In Solomon, J. (ed.), *Accessing Antiquity: The Computerization of Classical Studies*. London and Tucson: University of Arizona Press, pp. 34-61.
- Berners-Lee, T., Cailliau, R., Groff, J.-F. and Pollermann, B.** (1992.) World-Wide Web: The Information University, *Electronic Networking*, 2(1): 52-8.
- Brunner, T.F.** (1987). Data Banks for the Humanities: Learning from Thesaurus Linguae Graecae, *Scholarly Communication*, 7: 1 and 6-7.
- Brunner, T.F.** (1993). Classics and the Computer: The History of a Relationship. In Solomon, J. (ed.), *Accessing Antiquity: The Computerization of Classical Studies*. London and Tucson: University of Arizona Press, pp.10-33.
- Cox, J.** (2006). Scholarly Publishing Practices: A Case of plus ça change, plus c'est la même chose?, *Learned Publishing*, 19: 273-76.
- Crane, G.** (2004). Classics and the Computer: An End of History. In Schreibman, S., Siemens, R. and Unsworth, J. (eds), *A Companion to Digital Humanities*. Malden, Melbourne and Oxford: Blackwell Publishing, pp. 46-55.
- Crow, R.** (2002). The Case for Institutional Repositories: A SPARC Position Paper, Washington (DC): Scholarly Publishing and Academic Resources Coalition. http://www.arl.org/sparc/bm~doc/ir_final_release_102.pdf (accessed 12 June 2007).

- Doherty, L.** (2005). Ancient Times Back in Vogue, *The Sydney Morning Herald*, 18 August: 13.
- Ginsparg, P.** (1994). First Steps towards Electronic Research Communication, *Computers in Physics*, 8 (4): 390-6.
- Hardwick, L.** (2000). Electrifying the Canon: The Impact of Computing on Classical Studies, *Computers and the Humanities*, 34: 279-95.
- Harnad, S.** (1995). The Post-Gutenberg Galaxy: How To Get There from Here, *The Information Society*, 11: 285-92.
- Hockey, S.** (2004). The History of Humanities Computing. In Schreibman, S., Siemens, R. and Unsworth, J. (eds), *A Companion to Digital Humanities*. Malden, Melbourne and Oxford: Blackwell Publishing, pp. 3-19.
- Houghton, J., Steele, C. and Sheehan, P.** (2006). *Research Communication Costs in Australia: Emerging Opportunities and Costs*. Department of Education, Science and Training, Canberra: Australian Government. http://www.dest.gov.au/NR/rdonlyres/0ACB271F-EA7D-4FAF-B3F7-0381F441B175/13935/DEST_Research_Communications_Cost_Report_Sept2006.pdf (accessed 4 May 2007).
- Hughes, J.J.** (1986). The Ibycus SC: A Multilingual Computer System for Scholars, *Bits and Bytes Review*, 1(1) 1-8.
- Hughes, J.J.** (1987). From Homer to Hesychius: The Thesaurus Linguae Graecae Project, *Bits and Bytes Review*, 1(7): 1-6.
- Kling, R.** (2004). The Internet and Unrefereed Scholarly Publishing, *Annual Review of Information Science and Technology*, 38: 591-631.
- Kling, R. and McKim, G.** (2000). 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: 1306-20.

- Kreitz, P.A., Addis, L., Galic, H. and Johnson, T.** (1996). The Virtual Library in Action: Collaborative International Control of High-Energy Physics Pre-Prints, *Publishing Research Quarterly*, 13(2): 24-32.
- Lane, B.** (2006). Benefits of Free Access, Higher Education Supplement, *The Australian*, 18 October: 21-2.
- Lynch, C.A.** (2001). Metadata Harvesting and the Open Access Initiative, *Association of Research Libraries: A Bimonthly Report*, 217 (August): 1-9.
- Lynch, C.A.** (2003). Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age, *Association of Research Libraries: A Bimonthly Report*, 226 (February): 1-7.
- Ober, J., Scheidel, W., Shaw, B. and Sanclemente, D.** (2007). Toward Open Access in Ancient Studies: The Princeton-Stanford Working Papers in Classics, *Hesperia*, 76: 229-42.
- Packard, D.W.** (1973). Computer-Assisted Morphological Analysis of Ancient Greek. In Zampolli, A. and Calzolari, N. (eds), *Proceedings of the Fifth Conference on Computational Linguistics*. Pisa: Association for Computational Linguistics, pp. 343-55.
- Poynder, R.** (2005). OA as Instrumental Good. *Open and Shut?*, 22 December.
<http://poynder.blogspot.com/2005/12/oa-as-instrumental-good.html> (accessed 16 April 2007).
- Suber, P.** (2005). Promoting Open Access in the Humanities', *Syllecta Classica*, 16: 231-46.
- Suber, P.** (2007a). Open Access in 2006. *SPARC Open Access Newsletter*, 105 (2 January).
www.earlham.edu/~peters/fos/newsletter/01-02-07.htm (accessed 20 June 2007).
- Suber, P.** (2007b). More on the Victory in the Senate. *Open Access News: News from the Open Access Movement*, 27 October.
http://www.earlham.edu/~peters/fos/2007_10_21_fosblogarchive.html (accessed 19 November 2007).

Notes

¹ <http://ccat.sas.upenn.edu/bmcr/stats/index.html> (accessed 11 November 2007).

² These are *Aigis*, *BMCR*, *Didaskalia*, *Electronic Antiquity*, *Frankfurter elektronische Rundschau zur Altertumskunde*, *Leeds International Classical Studies*, *Rosetta*, and *Classics@*. While the published volumes of *Arachnion* and *Histos* are still available online, these are no longer going concerns.

³ The first to do so in the United Kingdom was the Department of Classics at the University of Wales at Lampeter, whose series of working papers went live in November 2007 (<http://www.lamp.ac.uk/classics/workingpapers/>).

⁴ *Ancient History: Resources for Teachers* does not have an electronic version nor a policy on postprints. The same applies to *Iris*, *Mediterranean Archaeology*, *Prudentia*, and *Ramus. Classicum* and the *Bulletin of the Australian Centre for Egyptology* do allow the self-archiving of postprints but lack electronic versions. *Antichthon* rules out self-archiving entirely, while its volumes from 1993 to 2004 are available in electronic format via a local subscription service. By contrast, *Scholia* allows self-archiving, provides open access to its reviews, and an electronic version of every volume via Proquest and LOCKSS.

⁵ <http://roar.eprints.org> (accessed 15 November 2007).

⁶ In August 2006 the US-based Association of Research Libraries (<http://www.arl.org/>) reported that 43 percent of its members had institutional repositories, while some 35 percent intended to set one up in 2007. Its membership is restricted to research-intensive US universities and currently numbers 123.

⁷ For their assistance with this article special thanks go to Lea Beness, Susanne Binder, Gary Browne, Sten Christensen, William Dominik, Richard Hamilton, David Konstan, Donald Mastronarde, Elizabeth Minchin, Kit Morell, Camilla Norman, Josiah Ober, John Penwill, Ian Plant, Donna Sanclemente, Walter Scheidel, Brent Shaw, Colin Steele and Marcus Wilson as well as this journal's editor and her anonymous readers.

