A Call for Open Scholarship in Archaeology

Brian Hole

UCL Institute of Archaeology and Ubiquity Press

Abstract

Open Scholarship is a term that covers a range of information sharing mechanisms, including e-Publishing, Open Access and Open data. This chapter argues that these behaviours need to become more widely adopted in archaeology, especially in the context of an environment where funding is difficult to obtain and there is strong competition from other disciplines. It makes nine recommendations on how to achieve this, including the use of data journals to maximise research impact.

Introduction

Along with other disciplines in the humanities, archaeology is in a difficult position when the government funding it depends upon is provided under competitive conditions that ask it to demonstrate its 'impact' relative to fields in Science, Technology and Medicine (STM, sometimes also including Engineering to become STEM) that have closer ties to industry. As this chapter will discuss, Open Scholarship incorporates mechanisms such as e-Publishing, Open Access (OA) and open data, which lead to greater exposure, recognition and funding. But it is also much more than this – it is a culture of sharing, collaboration and efficiency among researchers that leads to more rapid advancement in knowledge.

The Need to Demonstrate Greater Impact for Funding

In an environment of financial austerity the humanities must begin following developments in this direction within STM in order to remain competitive. The competition for resources in the UK is very real, acute and serious:

> '…the enormous achievements of non-STEM disciplines are often
> overlooked – even when these involve, as so often, vital
> interdisciplinary research spanning the natural and social sciences.
> These misconceptions are potentially very damaging, especially at a
> time of diminishing resources. As modern research has become
> more and more interdisciplinary, and we move increasingly beyond
> the sterile and outdated notion of a society of 'two cultures', the
> mutual dependencies of 'hard' science and the humanities and social
> sciences have become ever clearer.' (British Academy 2010, 5)

This misconception of the relative value of the humanities continues to exist even at a purely economic level, despite estimates from economists that for every £1 spend on humanities research, £10 of immediate benefit is derived, with another £15-£20 in the long term (AHRC 2009, 3). Archaeology contributes a significant share to this, especially through its integral support of the heritage and tourism

industries, alone worth over £20.6 billion in GDP and supporting around 466,000 jobs (HLF 2010, 5). Despite this, when it comes to making budgetary decisions the humanities often come out on the losing side. A stark example of this is the announcement by London Metropolitan University that it plans to shut down 400 of its 577 courses, almost all of which are in the arts and humanities (Swain 2011). Archaeology has also been hit directly, with the Council for British Archaeology (CBA) set to loose all of its funding by 2015 (British Academy 2011).

In order to mitigate further cuts, and even reverse them if possible, it is therefore essential that all areas of the humanities including archaeology find ways to create and exhibit greater impact. This is especially important for university archaeology departments, who have to compete for funding by demonstrating impact in the Research Excellence Framework (REF) exercise. This can be done by embracing recent developments more prevalent in STM as well as the principles of Open Scholarship. In order to do this, it is important to note that there are significant differences in the way that research is both carried out and disseminated between the humanities and STM. Humanities researchers publish a far greater proportion of their work (around 50% compared to 20%) in monograph form for example (Huang and Chang 2008, 1820), and exhibit '… only limited uptake of even simple, freely available tools for data management and sharing' (RIN 2011, 7).

Open Scholarship

There is a general trend towards more open practices in science and technology, which has impacted both the scholarly process and publishing industry in several ways. Open source software has made it affordable for academic societies and smaller companies to enter the e-publishing space. At the same time demand from researchers, Higher Education Institutions (HEIs) and funding bodies for more affordable access to research outputs has resulted in a rapidly increasing number of HEI research paper repositories ('Green' Open Access), Open Access journals ('Gold' Open Access), and more recently, open data repositories. A growing body of evidence is now accumulating that demonstrates that open practices lead to greater collaboration (e.g. Zhang and Chen 2010), increased efficiency of research (e.g. Houghton *et al.* 2009), and higher and wider impact (e.g. Evans and Reimer 2009) A great benefit of open systems having been designed to enable sharing is that, albeit imperfectly, they provide mechanisms for tracking the usage of scholarly outputs and providing statistical measures of their impact. This not only motivates individual researchers to take part by providing mechanisms for recognition and reward, it also has the potential to provide policy making and funding bodies with more accurate decision-making metrics when comparing the value and impact of various disciplines. All of this together can be termed Open Scholarship (figure 1).
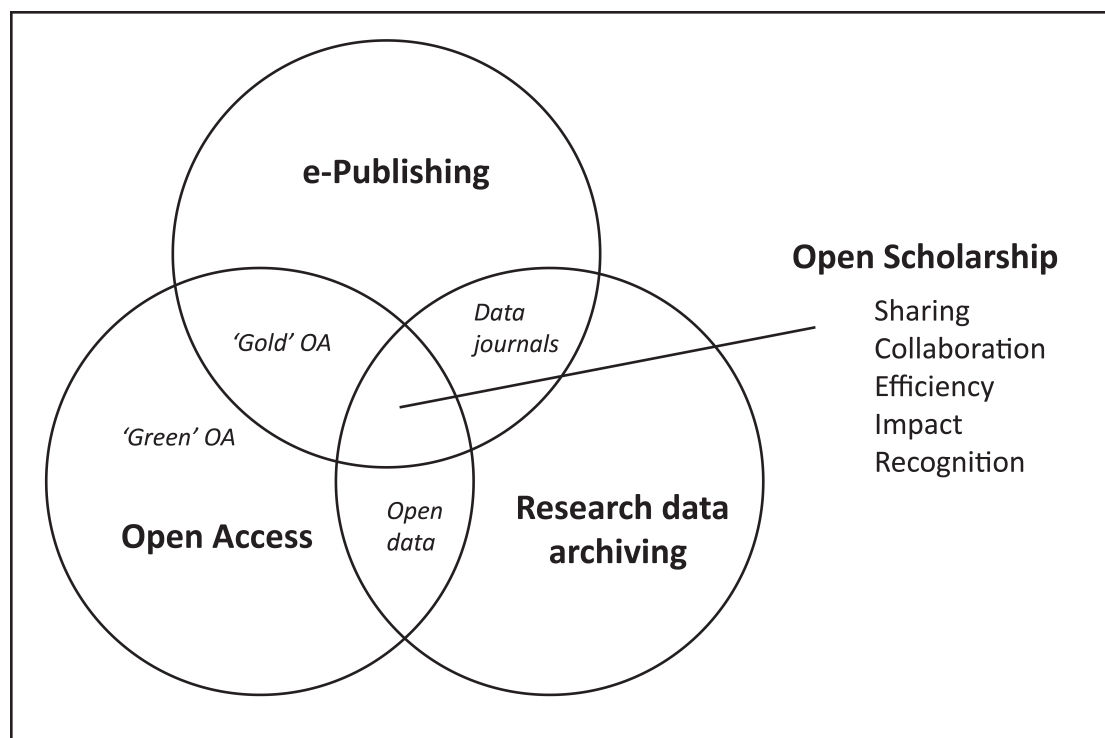
Figure 1 Venn diagram of Open Scholarship factors.

There are many reasons for archaeology to make a cultural shift and embrace Open Scholarship, including the benefits of greater efficiency, impact and collaboration. The most pressing however is the need to be able to fully articulate its value and impact to government in order to receive a fair share of funding and protect its future. To do so it is necessary to keep apace with developments in other academic areas, especially STM, for which a funding bias currently exists and needs to be addressed. The following sections look at what embracing Open Scholarship will require in each of its main component areas – e-publishing, Open Access and open data, from which a series of recommendations are then drawn.

e-Publishing

Along with conferences, publishing is the most prevalent means of disseminating research in the humanities. Publishing practices however are significantly behind those in the STM when it comes to wide distribution, tracking of impact, and rewarding of researchers. In order to reach as many readers as possible, it is essential to publish a paper electronically. Over 96.1% of STM journals have an electronic edition, but the humanities and social sciences are still trailing with 86.5% (RIN 2009, 14). The importance of this is highlighted by several studies that have shown that readership trends are strongly favouring e-journals. Tenopir *et al* have found that electronic versions of articles now account for the majority of readings in the sciences (2008: 5), and a recent study titled *Information Behaviour of the Researcher of the Future* by the UCL CIBER group found that the younger the scholar, the stronger the preference for e-materials:

> 'The age differences are startling and they suggest that the shift away from the physical to the virtual library will accelerate very rapidly…'
> (CIBER 2008, 13)

What this means is that work published in print-only journals will not only be less likely to be read, but that in the near future it will hardly have any audience at all. The outlook is the same for monograph publishing, also of great importance to archaeologists. The CIBER study also mentions the 'inexorable rise of the e-book', predicting that they will soon be the 'primary format for educational textbooks and scholarly books and monographs' (CIBER 2008, 26). This is backed up by the JISC *National e-Books Observatory Project* report that found that nearly 65% of UK teaching staff and students have used an e-book, and recommends a 'major expansion of library provision' in this area (JISC 2009, 9). Paper-only monographs will have readers in the short-term, but in line with journals they will lose their primary position and have many less readers in future unless digitised.

In addition to being the main preference of readers for journals and the emerging one for books, electronic publications have other major benefits: they can more easily be found and citations to them can be tracked, which can in turn be used to help assess impact. Electronic articles are assigned a Digital Object Identifier (DOI) that is usually displayed at the top of the article (e.g. http://dx.doi.org/10.5334/bha.2114). Unlike a regular web URL that is likely to stop working with time as web pages move, a DOI is a *persistent identifier* that is kept updated and should always point to the referenced article. If it is included in a bibliographic reference to the article, then readers will always be able to easily navigate to the article in future.

DOIs also enable citations of an article to be tracked, which helps to form an idea of the impact it has had. This is not a perfect measure, but it is the best quantitative measure currently available. Future improvements will include the ability to automatically assess the sentiment of a citation (i.e. positive or negative), but at the moment simply being able to rank articles from those with no citations to those with many is of great advantage. This is extremely useful for an assessor from a UK Research Excellence Framework (REF) panel or a US university's tenure review committee for example.

The use of DOIs is fundamental to unlocking the benefits of Open Scholarship, yet lacking in many archaeology journals. Readers perusing the references at the end of this chapter will notice that in stark contrast to the articles from STM journals, a reference (Richards 1997) from one of the leading archaeological journals, *Antiquity,* does not have a DOI. Authors who publish in journals that do not use such standards, no matter how prestigious, will find themselves disadvantaged in the longer term as article metrics become more established.

Citations using DOIs can also be aggregated to assess the impact of a journal overall, the most well known statistic being the Journal Impact Factor from Thomson Reuters. It is very common in the STM world to assess the quality of an article based on the Impact Factor of the journal it is published in, even though its accuracy is hotly contested (e.g. Rossner *et al* 2007). This is something of a problem for the humanities as the Impact Factor is used primarily for journals in the sciences and social sciences (Hubbard and McVeigh 2011, 133). Not all humanities journals therefore are able to get an Impact Factor, and when they do it is often very low, as citations from the other humanities journals are not included in the calculations, and only citations for the last two years are taken into account (the typical timeframe of high-citation intensity for a sciences paper), but for humanities disciplines the majority of citations may be spread out to a much greater degree in the future. This is greatly compounded by the fact

that at the present time citations in books, of great importance in the humanities, are also not counted. The presence, absence or value of a Journal Impact Factor is therefore not a good way of determining the impact of an article, with the number of citations of each article in isolation providing a much more accurate picture. This is something that researchers often misunderstand to their detriment in the UK's REF exercise, often submitting work from Impact Factor journals even though the Higher Education Funding Council for England (HEFCE) and the REF assessment panels explicitly tell them that this is not taken into consideration (SCST 2004, par. 209).

In order to counter the impact factor bias, publishers in the humanities and archaeology need to provide additional article level metrics such as those used in Public Library of Science (PLoS) journals (PLoS 2011) and recommended by the Alt-Metrics manifesto (Priem *et al* 2010) and others (e.g. Neylon and Wu 2009). Citation data can be complimented with metrics on links to the article from social networking sites like Facebook and Twitter, social bookmarking services such as CiteULike and Connotea, reference tools such as Mendeley and Zotero, and in blogs and other media. This information provides a much wider context with which to judge the impact of research on both the academic and the wider community.

Fortunately for archaeology, with e-publishing exactly the same solution as that described above can be applied to monographs. Social Sciences and Humanities researchers publish significantly less journal articles and more books than do their colleagues in STM (Huang and Chang 2008, 1820). By switching to e-book publishing (with the possibility of retaining print editions or using print on demand), publishers can treat books and book chapters exactly the same way as articles, with all of the same advantages. It is becoming standard practice to assign DOIs to book chapters, so that they can be easily found, cited, and tracked for impact as well.

The significance of this trend in publishing should not be underestimated. Many archaeologists publish as much as 50% of their research in book form, and if they do not publish these electronically they will find themselves at a significant disadvantage when their work is judged for impact against that of colleagues who do. Equally important in the UK, with print-only publications the overall impact of archaeology will not be assessed as highly as it might be, with a corresponding loss of government funding through the REF. Readers will see the irony in the fact that the book in which this chapter appears may be available on paper only. I have therefore provided a reference[1] that points to an electronic version in the UCL institutional repository, an example of green OA.

Open Access

The benefits to authors of publishing electronically are multiplied when this is Open Access (OA). OA refers to electronic papers that are published in proper academic journals and subject to full peer review as usual, but are also made freely available to anyone in the world with an internet connection. Based on figures from the Directory of Open Access Journals (DOAJ), the number of OA titles is increasing at a rate more than three times faster than for all journals on average (figure 2, updated from Morrison 2011a). This indicates that OA is a major trend that will soon account for a significant amount of published journals.
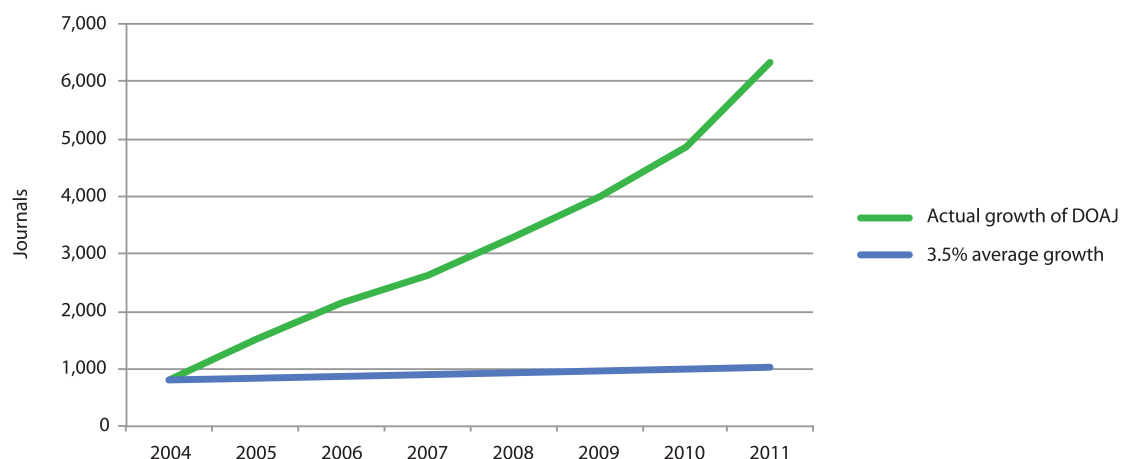
Figure 2 Chart of OA growth compared to the average for all journals, based on data collected by Morrison (2011b) detailing the growth of the Directory of Open Access Journals (DOAJ), and the estimate of 3.5% average growth from Ware and Mabe (2006, 5).

The humanities and archaeology in particular still lag significantly behind STM in OA publishing, and once more need to make up ground in order to remain competitive. OA articles have been shown to be downloaded 89% more frequently in the first year of publication than those from subscription journals (Davis *et al* 2008, 1), thus reaching a wider audience more quickly and having a more immediate impact. This extra distribution subsequently results in significantly higher citation rates, e.g. 157% higher in computer science (Lawrence 2001), and up to 290% higher in the sciences in general (Eysenback 2006, 692).

This is not only important for increasing institutional funding by demonstrating greater impact, but also for the funding of individual researchers and projects. The Research Councils UK (RCUK) *Position Statement on Access to Research Outputs,* to which the Arts and Humanities Research Council (AHRC) is a signatory, states that:

> 'Ideas and knowledge derived from publicly-funded research must be made available, and accessible for public use, interrogation and scrutiny, as widely, rapidly and effectively as practicable.' (RCUK 2006)

Funders cannot force an author to use an open access publisher instead of another one, but their preference is reflected in their policies, which now overwhelmingly mandate OA archiving ('Green' OA) as a minimum (see figure 3 for the UK). This is now also a requirement of the majority of UK HEIs (e.g. UCL 2010).

Funding Agency	Article must be available free of charge without any access restrictions	It must be the final full version
Action on Hearing Loss	X	X

Arthritis Research UK	X	X
Arts and Humanities Research Council (AHRC)	X	
Biotechnology and Biological Sciences Research Council (BBSRC)	X	
British Heart Foundation (BHF)	X	X
Cancer Research UK	X	X
Chief Scientist Office, Scottish Executive (CSO)	X	X
Department of Health (DoH)	X	X
Dunhill Medical Trust (DMT)	X	X
Economic and Social Research Council (ESRC)	X	X
Engineering and Physical Sciences Research Council (EPSRC)	X	X
JISC (Joint Information Systems Committee)	X	X
Medical Research Council (MRC)	X	X
National Institute for Health Research (NIHR)	X	X
Natural Environment Research Council (NERC)	X	X
Parkinson's UK	X	X
Science and Technology Facilities Council (STFC)	X	X
Stroke Association		
Wellcome Trust	X	X

Figure 3 Funding policies of UK research funders for peer-reviewed research outputs (derived from SHERPA-JULIET 2011)

OA is above all an integral part of Open Scholarship because it prevents academic and economic exclusion. The OA funding model (for fully OA journals) involves charging an Article Processing Charge (APC) upon article acceptance, which is typically paid by the author's institution or funder. The article is then made freely available to anyone anywhere in the world, as opposed to the traditional model in which a reader or their institution must first either purchase a subscription or pay a relatively high fee to download it. The OA model is perfectly sustainable, as has been shown by highly successful publishers such as BioMed Central, and it has been estimated that its universal adoption in the UK could result in savings of £200 million per year (Houghton *et al.* 2009, xx).

The need for open access to scholarly resources in the developing world in particular is something that is widely recognised and supported in the academic community, and often this support is underestimated by non-OA publishers. This was demonstrated by the massive outcry against a leading group of publishers who chose to withdraw their journals from the World Health Organisation's (WHO) Health Internetwork for Access to Research Initiative (HINARI) in January of 2011 (Pérez Koehlmoos and Smith 2011). The argument that these countries could now afford to begin paying subscriptions to the most expensive medical journals in the world, despite having the highest levels of avoidable mortality caused outrage among the scientific community. An editorial in *The Lancet* (one of the journals that had been withdrawn by its publisher, Elsevier) summarised this, and presented something of an ultimatum:

> ' If publishers are genuine about their mission to improve health through partnerships with medical and research communities, they need to send a stronger signal of commitment to countries that most need the knowledge they control.' (The Lancet 2011, 272)

While archaeology journals cannot promise to help reduce mortality in poor countries, the arguments for providing free access to information are nonetheless strong. By withholding access to the research results of Western institutions, publishers are effectively ensuring the exclusion of researchers in developing countries from the global academic community. This is essentially a case of a privileged club protecting its own membership, despite outward claims to the contrary. As a telling example, as of August 17th 2011, the World Archaeological Congress (WAC) charges over £30 for a single 15 page article from its *Archaeologies* journal through Springer, and £81 for books in it's *One World Archaeology* series. WAC claims that it aims '… to foster international academic interaction' and that it recognises 'the need to make archaeological studies relevant to the wider community', yet in the current economic climate these prices are already too much for researchers and libraries in the UK, and well beyond the reach of those in developing countries at any time.

Open Research Data

Research Data is any information collected and processed for the purposes of academic study. It can be in many sizes and formats, from large relational databases to small Microsoft Excel files. Some examples of archaeological research data are GIS data, geophysical survey data, image files and qualitative survey notes. The archiving of research data is another important aspect of Open Scholarship. As with electronic publishing and Open Access for research articles, there are major benefits, both to the originators of the data and to the wider community of this approach. Several open data repositories have appeared in STM such as GenBank, which specialises in gene sequencing data (Benson *et al* 2007) and Dryad, which archiveslong-tail data in the biosciences (Vision 2010). While far fewer open data repositories exist in the humanities, archaeology is better served than most disciplines, with the Archaeological Data Service (ADS) in York for example (Richards 1997).

Despite the availability of these repositories, 80% of research data is not openly available (PARSE insight 2010, 19). Many reasons for this are cited by researchers, often including the lack of time to prepare data for archiving or to make it understandable to others, a problem that increases with time. The post doctoral researcher who collected the data may have left for example. Even after as little as a year those who produced the data files may no longer easily understand them, and within 2-5 years the file formats themselves may no longer be readable, while the disks or other media they are stored on may have decayed or become obsolete. Other scholars, particularly those who have assembled large datasets over a long period of time, often cite competitive advantage as a reason for not archiving. A recent study has shown that even in disciplines with a relatively strong culture of data archiving such as genetics, authors of only 25% of papers had placed the associated research data in appropriate repositories (Piwowar 2011). In another study, only 10% of researchers were willing to supply the data underlying their publications in medicine when requested (Savage and Vickery 2009). This is especially important as the data upon which a paper is based needs to be made available in order to validate its conclusions, including even the values used to make a simple chart (many graduate students know what it is like to have to extract such data from a graph with the help of a ruler, with all of the associated room for error). The data behind the charts and tables in this chapter are openly available in the UCL institutional repository (Hole 2011).

Besides validation, there are many compelling reasons to openly archive data. From a purely selfish point of view, researchers who do so can expect to be well rewarded in the near future. As with research articles, research datasets receive a DOI when archived. It is now becoming established as a best practice in STM to include the DOI to the archived data on which an article is based in the article itself, so that readers can navigate to it easily. Similarly datasets from other researchers that have been used should be cited in the reference section. This means that the data also becomes findable, citable and trackable, like any other e-publication. This is extremely important for archaeology in the UK, as from 2014 research data will be a recognised research output for the REF. The REF assessment panels will find it easiest to assess the impact of data if it has similar metrics to those for articles. It is highly probable that citation statistics for data are even more valuable. Where a research dataset and a research article each have ten citations for example, it can be argued that the dataset has had a greater impact due to the fact that it has been reused. This will only happen if both a culture of data archiving and data citation are established in archaeology, and if evidence from a recent study in the social sciences, where only 29% of data used was found to be referenced (Mooney 2011) is any indication, then we still have some way to go. Overcoming the reluctance of some researchers to share data despite the above benefits could be achieved by offering an embargo period, through which they would retain sole access to the data for 2-5 years following deposition.

The incentives to openly archive and cite data are extremely compelling however. Not only do researchers get credit for having made the data available, but it has also been shown that this can lead to a significant increase in citations of associated research papers, found to be 69% higher in genetics (Piwowar *et al* 2007). It is the actual reuse of archived data that has the greatest value. By providing their data as a platform for others to build upon, researchers can have a truly lasting impact. To borrow an extended analogy that Heather Piwowar from the University of British Colombia uses (see Cheng 2011 for a nice presentation of this), if you want others to see further by standing on your shoulders, then you need to have broad shoulders, and sharing your data is the most effective way to achieve this. This makes research more efficient overall, as researchers are able to reuse data instead of recreating it, effectively creating more output for less budget.

Data sharing not only means more citations and a greater legacy, it can also result in an increased number of opportunities for collaboration with others in the same field, research that in turn has been found to rank more highly than average for impact (Whitfield 2008, 721). Data repositories can also make datasets machine readable as 'linked data', resulting in new kinds of multidisciplinary research being done, which the creators of the data have not even envisaged (Van de Sompel and Lagoze 2009; Bizer *et al* 2009). Examples could involving combining archaeological datasets with those from linguistics, genetics or environmental science.

In order to archive data effectively so that it will be available to other researchers in the long term, it is important to choose an appropriate repository. Making data available as supplementary files along with a research article is not archiving, and there is no guarantee that it will still be available on the publisher's website in future. One study on biomedical publications for example found that only 83% of links to supplemental material were still functional just one year after publication

(Anderson *et al* 2006), and another that found that publishers only had preservation plans for 12% of supplemental files (PARSE insight 2010, 13). Repositories specialise in the care of data. They curate it, add metadata, and index it with other services. They also keep back-ups, perform checks on its state over time to make sure that it hasn't changed, and migrate it to new file formats if the original ones become obsolete. Finally they provide the data with a permanent identifier such as a DOI so that it can be easily found and cited, and its usage tracked. It is also important that a repository is open, as it is of little benefit to the research community if simply used as a back up, and will not attract any significant number of citations. The ADS is one such suitable repository for archaeology. Unfortunately specialised research data repositories hosted by HEIs are still several years away in most cases. Some institutional research article repositories may be extended to include data in the short term, important in the UK if datasets are to be archived in time to be assessed for the next REF. It is however essential that such a repository provide a DOI (not all currently do), and this should be checked before depositing data.

The kind of open license under which data is archived is also extremely important. Many open licenses such as "share-alike" and "copyleft" create legal obligations involving copyright and attribution that make sharing of data impractical, and their use is to be strongly discouraged (e.g. Science Commons 2011, par. 4.1). The potential of open and linked data is huge, but in some scenarios this may involve the creation of new datasets by federating data from anywhere from 10 to 100,000 separate data sources. If each of those original datasets has a different license that a researcher is required to follow for their new dataset, then they are already lost. At the same time it may be very difficult to provide attribution for all of the source datasets in a myriad of different ways as requested under law. The solution is to choose a license that waives all rights to the data, and requests citation through scholarly norms, such as the Creative Commons Zero licence (CC0) (see Creative Commons 2011).

Finally, it is important to ensure that your data is in a sufficiently usable form that anyone downloading it can access and understand it. If they cannot then they are unlikely to reuse and cite it. This means that the data needs to be well formatted and described – for example a spreadsheet with ambiguous column headers will only cause confusion. The data should also be provided in an open format, which may mean exporting a Microsoft Excel file to CSV for example. This is important because you cannot guarantee that the recipient will have the same software, and a proprietary format is likely to be difficult to read 5-10 years down the line. This will also make it easier for the repository to automatically migrate your file to a new format later on if necessary. The deposited data must also be actionable, which means that if any unique software has been created to process and make sense of it, then this should also be deposited.

In summary, in order to broaden their shoulders and make their data open (and receive credit for it), researchers need to deposit it in an appropriate open repository that provides a DOI, make sure that it has a CC0 or equivalent license, and ensure that it is in usable form.

At the same time it is necessary to follow best practices in data citation. Once a dataset has been given a DOI, this should be cited in all publications associated with the data in the same way as a reference to a paper, in the references section. If placed elsewhere in the paper it is unlikely that the citation will be tracked and added to impact metrics. Ideally the DOI should be given with the

'http://dx.doi.org/' prefix, so that readers understand that it is usable as a web link. For example:

> Smith, J. 2011. ABC Project Osteological Data. *ADS Data Repository.* DOI: http://dx.doi.org/10.1234/98765.1

Although the benefits of these practices are substantial, it still takes time for them to be fully embraced, as seen in STM where adoption levels are still relatively low and increasing only slowly (Piwowar 2011). In the UK this is partly due to the absence of mandates from funding bodies (see figure 4).

Funding Agency	Data archiving required	Data must be deposited within five years
Action on Hearing Loss		
Arthritis Research UK		
Arts and Humanities Research Council (AHRC)		
Biotechnology and Biological Sciences Research Council (BBSRC)	X	X
British Heart Foundation (BHF)		
Cancer Research UK	X	
Chief Scientist Office, Scottish Executive (CSO)		
Department of Health (DoH)		
Dunhill Medical Trust (DMT)		
Economic and Social Research Council (ESRC)	X	X
Engineering and Physical Sciences Research Council (EPSRC)	X	
JISC (Joint Information Systems Committee)		
Medical Research Council (MRC)	X	
National Institute for Health Research (NIHR)		
Natural Environment Research Council (NERC)	X	
Parkinson's UK		
Science and Technology Facilities Council (STFC)		
Stroke Association		
Wellcome Trust	X	

Figure 4 Funding policies of UK research funders for data archiving (derived from SHERPA-JULIET 2011)

A good way to further encourage data archiving, sharing and citation is through the use of data journals, a relatively new concept in publishing. Ubiquity Press plans to begin publishing data journals in the humanities and social sciences from October 2011, beginning with the Journal of Open Archaeological Data (JOAD).

A data journal such as JOAD helps researchers to understand the value of archiving data by providing them with an additional citable publication in the form of an article, for which they are already used to gaining credit. This article, or data paper, is a concise description of the methodology by which the data was produced, and provides details of its reuse potential. The data paper and the dataset are submitted together to the journal and undergo peer review. In this case peer review is to ensure that the paper describes the dataset accurately, and that the data is in open form with an open license, well formatted and labelled, and is actionable. If the paper is accepted then user is given a provisional DOI for it, and asked to choose from a list of suitable repositories. The data is then

deposited in the chosen archive, along with the DOI of the paper. In return the repository will assign a DOI to the dataset, which is then incorporated into the paper in the data journal, and this is published (along with any other DOIs for research papers based on the data). The metadata about the dataset is also exposed in the published paper as linked data, enhancing its chances of discovery and probability of reuse. This process is outlined in figure 5.
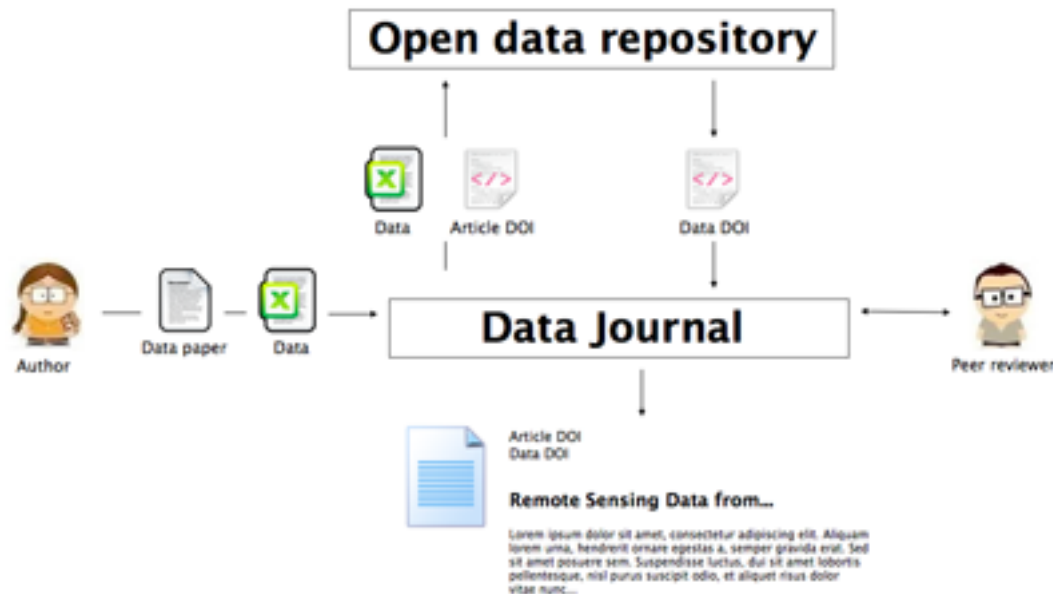

Figure 5 Proposed data journal workflow

As researchers are currently more likely to directly cite an article than a dataset, the data paper provides the author with a way to gain credit for open archiving. By engaging more readers with the data, the paper can be expected to also result in more citations of the dataset and the associated research articles, resulting in a more comprehensive demonstration of overall impact.

Summary: Recommendations for Open Scholarship in Archaeology

1. Publish research articles electronically. The impact of work in print-only journals will not be apparent to anyone assessing it.
2. Publish books electronically too. If you are contributing a chapter encourage the editors to go with a publisher who can produce an electronic edition.
3. Ensure that your publisher gives you a DOI for your article, book or chapter. If they cannot then it is best to go elsewhere as your citations will not be tracked.
4. At minimum archive a copy of your article, book or chapter in your institutional repository ('Green' OA).
5. If you have a choice, publish your article in an Open Access journal or your chapter or book with a publisher who will release it as OA ('Gold' OA).
6. Archive your research data in an appropriate open repository, ensuring that it has a CC0 or similar licence and is in an open, usable form.

7. Ensure that the repository can give you a DOI or similar persistent identifier, otherwise consider another repository that can do so, so that use of the data can be tracked.
8. Cite the data in the reference section of any research articles based on it.
9. Publish a data paper if the option is available to you in order to ensure that the academic community are aware of its availability and reuse potential. This will also help you to achieve the widest dissemination of your work and the greatest overall impact.

Conclusion

This chapter has sought to demonstrate that practising Open Scholarship has important benefits for the advancement of knowledge, but also that doing so can greatly increase the visibility of research and benefit a researcher's career. Most importantly, these practises should improve the future prospects for the humanities and archaeology, enabling them to better assert their value and impact in a competitive world. It is hoped that the recommendations set out above can be of immediate and practical effect in moving towards this goal.

Acknowledgements

Many thanks to Heather Morrison for sharing her data on the growth of DOAJ (Morrison 2011b).

Notes

1 The electronic version of this paper can be found in the UCL Discovery repository: http://discovery.ucl.ac.uk/.

References

AHRC 2009. *Leading the World: The Economic Impact of UK Arts and Humanities Research.* Bristol: Arts and Humanities Research Council, http://www.ahrc.ac.uk/About/Policy/Documents/leadingtheworld.pdf (accessed 16 August 2011).

Anderson, N. R., Tarczy-Hornoch, P. and Bumgarner, R. E. 2006. On the Persistence of Supplementary Resources in Biomedical Publications. *BMC Bioinformatics* 7:260, DOI: http://dx.doi.org/10.1186/1471-2105-7-260.

Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J. and Wheeler, D. L. 2007. GenBank. *Nucleic Acids Research,* 2008, 36 (Database issue D25–D30), DOI: http://dx.doi.org/10.1093/nar/gkm929.

Bizer, C., Heath, T. and Berners-Lee, T. 2009. Linked Data – The Story So Far. *International Journal on Semantic Web and Information Systems* 5(3), DOI: http://dx.doi.org/10.4018/jswis.2009081901.

British Academy 2010. Past Present and Future: The Public Value of the Humanities and Social Sciences. London: The British Academy, http://www.britac.ac.uk/templates/asset-relay.cfm?frmAssetFileID=9608 (accessed 15 October 2011).

British Academy 2011. British Academy Funding for CBA. London: The British Academy, http://www.britac.ac.uk/news/news.cfm/newsid/464 (accessed 16 August 2011).

Cheng, C. 2011. Data Sharing Makes Our Shoulders Broader: A Prezi about Dryad for web visitors. http://prezi.com/x0rcxhxqgtwf/dryad-data-sharing-makes-our-shoulders-broader/ (accessed 16 August 2011).

CIBER 2008. *Information Behaviour of the Researcher of the Future: A CIBER Briefing Paper.* University College London.

Creative Commons 2011. About CC0 — "No Rights Reserved". http://creativecommons.org/about/cc0 (accessed 16 August 2011).

Davis, P. M., Lewenstein, B. V., Simon, D. H., Booth, J. G. and Connolly, M. J. L. 2008. Open Access Publishing, Article Downloads, and Citations: Randomised Controlled Trial. *BMJ* 337:a568, DOI: http://dx.doi.org/10.1136/bmj.a568.

Evans, J. A. and Reimer, J. 2009, Open Access and Global Participation in Science. *Science* 323(5917):1025, DOI: http://dx.doi.org/10.1126/science.1154562.

Eysenbach, G. 2006. Citation Advantage of Open Access Articles. *PLoS Biology* 4(5):e157. DOI: http://dx.doi.org/10.1371/journal.pbio.0040157.

Swain, H. 2011. London Met VC Explains Why He is Cutting 400 Courses. *The Guardian* Tuesday 3 May 2011, http://www.guardian.co.uk/education/2011/may/03/london-metropolitan-gillies-course-cuts (accessed 16 August 2011).

HLF 2010. *Investing in Success: Heritage and the UK Tourism Economy.* London: Heritage Lottery Fund, http://www.hlf.org.uk/aboutus/howwework/Documents/HLF_Tourism_Impact_single.pdf (accessed 16 August 2011).

Hole, B, 2011. Data from "A Call for Open Scholarship in Archaeology" fig. 2. UCL Discovery repository, http://discovery.ucl.ac.uk/1325645/.

Houghton, J., Rasmussen, B., Sheehan, P., Oppenheim, C., Morris, A., Creaser, C., Greenwood, H., Summers, M. and Gourlay, A. 2009. *Economic Implications of Alternative Scholarly Publishing Models: Exploring the Costs and Benefits.* JISC, http://www.jisc.ac.uk/media/documents/publications/rpteconomicoapublishing.pdf (accessed 15 October 2011).

Huang, M. and Chang, Y. 2008. Characteristics of Research Output in Social Sciences and Humanities: From a Research Evaluation Perspective. *Journal of the American Society for Information Science and Technology,* 59(11):1819-1828, DOI: http://dx.doi.org/10.1002/asi.20885.

Hubbard, S. C. and McVeigh, M. E. 2011. Casting a Wide Net: the Journal Impact Factor Numerator, *Learned Publishing* 24(2), DOI: http://dx.doi.org/10.1087/20110208.

JISC 2009. *National e-Books Observatory Project: Key Findings and Recommendations.* London: Joint Information Systems Council, http://observatory.jiscebooks.org/files/2011/08/JISC-national-ebooks-observatory-project-final-report.pdf (accessed 15 October 2011).

The Lancet 2011. Editorial: Bad decisions for global health. *The Lancet* 37(9762):272, DOI: http://dx.doi.org/10.1016/S0140-6736(11)60066-4.

Lawrence, S. 2001. Free online availability substantially increases a paper's impact. *Nature – Web Debates* http://www.nature.com/nature/debates/e-access/Articles/lawrence.html (accessed 16 August 2011).

Mooney, H. 2011. Citing Data Sources in the Social Sciences: Do Authors Do It? *Learned Publishing,* 24:99–108, DOI: http://dx.doi.org/10.1087/20110204.

Morrison, H. 2011a. Dramatic Growth of Open Access March 31 2011. The Imaginary Journal of Poetic Economics (blog), http://poeticeconomics.blogspot.com/2011/03/dramatic-growth-of-open-access-march-31.html (accessed 16 August 2011).

Morrison, H. 2011b, Dataset: "Dramatic Growth of Open Access",
    http://hdl.handle.net/1902.1/14446 V12.
Neylon, C. and Wu, S. 2009. Article-Level Metrics and the Evolution of Scientific
    Impact. *PLoS Biology* 7(11): e1000242, DOI:
    http://dx.doi.org/1010.1371/journal.pbio.1000242.
PARSE insight, 2010, *Insight into Digital Preservation of Research Output in
    Europe: Insight Report.* http://www.parse-insight.eu/downloads/PARSE-
    Insight_D3-6_InsightReport.pdf (accessed 16 August 2011).
Pérez Koehlmoos, T. and Smith, R. 2011. Big publishers cut access to journals in
    poor countries. The Lancet, January 18th 2011, DOI:
    http://dx.doi.org/10.1016/S0140-6736(11)60067-6.
Piwowar H. A., Day R. S. and Fridsma D. B. 2007. Sharing Detailed Research
    Data Is Associated with Increased Citation Rate. *PLoS ONE* 2(3):e308, DOI:
    http://dx.doi.org/10.1371/journal.pone.0000308.
Piwowar H. A. 2011. Who Shares? Who Doesn't? Factors Associated with Openly
    Archiving Raw Research Data. *PLoS ONE* 6(7):e18657, DOI:
    http://dx.doi.org/10.1371/.journal.pone.0018657,
PLoS 2011. Article Level Metrics. http://article-level-metrics.plos.org/ (accessed
    16 August 2011).
Priem, J.*,* Taraborelli, D., Groth, P. and Neylon, C. 2010. Alt-Metrics: A Manifesto.
    Version 1.0 – October 26, 2010, http://altmetrics.org/manifesto/ (accessed 16
    August 2011).
Richards, J. D. 1997. Preservation and Re-Use of Digital Data: The Role of the
    Archaeology Data Service. *Antiquity* 71(274):1057.
RIN 2009. *E-Journals: Their Use, Value and Impact - Phase One Report.* London:
    The Research Information Network,
    http://www.rin.ac.uk/system/files/attachments/E-journals-report.pdf
    (accessed 15 October 2011).
RIN 2011. *Reinventing Research? Information Practices in the Humanities.*
    London: The Research Information Network,
    http://www.rin.ac.uk/system/files/attachments/Humanities_Case_Studies_for_
    screen_2_0.pdf (accessed 15 October 2011).
Rossner, M., Van Epps, H. and Hill, E. 2007. Show Me the Data. *Journal of
    Experimental Medicine*, 204(13), DOI:
    http://dx.doi.org/10.1084/jem.20072544.
Savage, C. J. and Vickers, A. J. 2009. Empirical Study of Data Sharing by Authors
    Publishing in PLoS Journals. *PLoS ONE* 4(9):e7078, DOI:
    http://dx.doi.org/10.1371/journal.pone.0007078.
SCST 2004. *Tenth Report.* House of Commons Select Committee on Science and
    Technology,
    http://www.publications.parliament.uk/pa/cm200304/cmselect/cmsctech/399/
    39902.htm (accessed 16 August 2011).
Science Commons 2011. Protocol for Implementing Open Access Data.
http://sciencecommons.org/projects/publishing/open-access-data-protocol/
    (accessed 17 August 2011).
SHERPA-JULIET 2011. Research Funders' Open Access Policies.
    http://www.sherpa.ac.uk/juliet/ (accessed 16 August 2011).
Tenopir, C., King, D. W., Edwards, S. and Wu, L. 2009. Electronic Journals and
    Changes in Scholarly Article Seeking and Reading Patterns. *Aslib Proceedings:
    New Information Perspectives*, 61(1), DOI:
    http://dx.doi.org/10.1108/00012530910932267.
UCL 2010. UCL Publications Policy 2010.
    http://www.ucl.ac.uk/library/publications-policy.shtml (accessed 16 August
    2011).

Van de Sompel, H. and Lagoze, C. 2009. All Aboard: Toward a Machine-Friendly Scholarly Communication System. In: Hey, T., Tansley, S and Tolle, K. *The Fourth Paradigm: Data-Intensive Scientific Discovery.* Redmond: Microsoft Research, http://research.microsoft.com/en-us/collaboration/fourthparadigm/4th_paradigm_book_part4_sompel_lagoze.pdf (accessed 15 October 2011).

Vision, T. J. 2010. Open Data and the Social Contract of Scientific Publishing. *BioScience* 60(5):330-331, DOI: http://dx.doi.org/10.1525/bio.2010.60.5.2.

WAC 2010. About WAC. http://www.worldarchaeologicalcongress.org/about-wac/about-wac (accessed 17 August 2011).

Ware, M. and Mabe, M. 2009. *The STM Report: An Overview of Scientific and Scholarly Journal Publishing.* Oxford: STM: International Association of Scientific, Technical and Medical Publishers, www.stm-assoc.org/2009_10_13_MWC_STM_Report.pdf (accessed 15 October 2011).

Whitfield, J. 2008. Group Theory. *Nature* 455: 720-723, DOI: http://dx.doi.org/10.1038/455720a.

Zhang, J. and Chen, C. 2010. Collaboration in an Open Data eScience: A Case Study of Sloan Digital Sky Survey. *ArXiv:* http://arxiv.org/abs/1001.3663v1.