

RESEARCH PAPER

Sowing the Seeds of Future Research: Data Sharing, Citation and Reuse in Archaeobotany

Lisa Lodwick

The practices of data sharing, data citation and data reuse are all crucial aspects of the reproducibility of archaeological research. This article builds on the small number of studies reviewing data sharing and citation practices in archaeology, focussing on the data-rich sub-discipline of archaeobotany. Archaeobotany is a sub-discipline built on the time-intensive collection of data on archaeological plant remains, in order to investigate crop choice, crop husbandry, diet, vegetation and a wide range of other past human-plant relationships. Within archaeobotany, the level and form of data sharing is currently unknown. This article first reviews the form of data shared and the method of data sharing in 239 articles across 16 journals which present primary plant macrofossil studies. Second, it assesses data-citation in meta-analysis studies in 107 articles across 20 journals. Third, it assesses data reuse practices in archaeobotany, before exploring how these research practices can be improved to benefit the rigour and reuse of archaeobotanical research.

Keywords: Archaeobotany; Data reuse; Data sharing; Open science

1. Introduction

Archaeology is a discipline built on the production and analysis of quantitative data pertaining to past human behaviour. As each archaeological deposit is a unique occurrence, ensuring that the data resulting from excavation and analysis are preserved and accessible is crucially important. Currently, there is a general perception of a low level of data sharing and reuse. Such a low level of data availability would prevent the assessment of research findings and the reuse of data in meta-analysis (Kansa & Kansa 2013; Moore & Richards 2015). As observed across scientific disciplines, there is a major problem in the reproduction of scientific findings, commonly known as the ‘replication crisis’ (Costello et al. 2013). A range of intersecting debates contribute to this, including access to academic findings (open access), open data, access to software and access to methodologies, which can be broadly grouped as open science practices. Without these, the way that scientific findings can be verified and built upon is impaired. Questions of reproducibility have been raised in recent years in archaeology, with considerations of a range of practices which can improve the reproducibility of findings, and a recent call for the application of open science principles to archaeology (Marwick et al. 2017). Discussion has so far focussed on access to grey literature (Evans 2015), data sharing (Atici et al. 2013), data citation practices (Marwick & Pilaar Birch 2018) and com-

putational reproducibility (Marwick 2017), with a focus on lithics, zooarchaeological evidence, and archaeological site reports.

Quantitative assessments of current levels of data sharing, data citation and reuse remain limited in archaeology. The focus of evaluation has been on the uptake of large-scale digital archives for the preservation and dissemination of digital data, such as the Archaeology Data Service (ADS), utilised by developer-led and research projects, and recommended for use by many research funders in the UK (Richards 2002; Wright and Richards 2018). Much less focus has been paid to the data-sharing practices of individuals or small-groups of university-based researchers who may be disseminating their research largely through journal articles. Recent work on the availability of data on lithics assemblages found a low level of data sharing (Marwick & Pilaar Birch 2018) and there are perceptions of low levels of data reuse (Huggett 2018; Kintigh et al. 2018). Within zooarchaeology numerous studies have explored issues of data sharing and reuse (Kansa & Kansa 2013, 2014), and the sub-discipline is seen as one of the most advanced areas of archaeology in regards to open science (Cooper & Green 2016: 273). Beyond zooarchaeology, however, explicit discussion has remained limited.

This paper assesses data sharing and reuse practices in archaeology through the case study of archaeobotany – a long established sub-discipline within archaeology which has well-established principles of data recording. Archaeobotany is an interesting case study for data sharing in archaeology as it straddles the division of archaeology between scientific and more traditional techniques.

Quantitative data on archaeological plant remains are also of interest to a range of other fields, including ecology, environmental studies, biology and earth sciences. The key issues of data sharing and data reuse (Atici et al. 2013) have been touched upon in archaeobotany over the past decade within broader discussions on data quality (Van der Veen, Livarda & Hill 2007; Van der Veen, Hill & Livarda 2013). These earlier studies focussed on the quality and availability of archaeobotanical data from developer-funded excavations in Britain and Cultural Resource Management in North America (Vanderwarker et al. 2016: 156). However, no discussion of data-sharing and reuse in academic archaeobotany occurred. A recent review of digital methods in archaeobotany is the notable exception, with discussions of the challenges and methods of data sharing (Warinner & d'Alpoim Guedes 2014).

Currently, we have no evidence for the levels of data sharing and reuse within archaeobotany. This article provides the first quantitative assessment of 1) data publication in recent archaeobotanical journal articles 2) data citation in recent archaeobotanical meta-analysis 3) the reuse of archaeobotanical datasets, in order to assess whether practices need to change and how such changes can take place.

2. Data Publication and Re-use Practices in Archaeobotany

2.1. History of data production and publication

Archaeobotanical data falls within the category of observational data in archaeology (Marwick & Pilaar Birch 2018). Archaeobotanical data is considered as the quantitative assessment of plant macrofossils present within a sample from a discrete archaeological context, which can include species identification, plant part, levels of identification (cf. – *confer* or “compares to”), and a range of quantification methods including count, minimum number of individuals, levels of abundance and weight (Popper 1988). Archaeobotanical data is usually entered into a two-way data table organised by sample number. Alongside the counts of individual taxa, other information is also necessary to interpret archaeobotanical data, including sample volume, flot volume, charcoal volume, flot weight, level of preservation, sample number, context number, feature number, feature type and period. Beyond taxonomic identifications, a range of other types of data are increasingly gathered on individual plant macrofossils (morphometric measurements, isotopic values, aDNA).

Archaeobotanical training places a strong emphasis on recording data on a sample-by-sample basis (Jacomet & Kreuz 1999: 138–139; Jones & Charles 2009; Pearsall 2016: 97–107). Time-consuming methodologies utilised in the pursuit of accurate sample-level data recording include sub-sampling and splitting samples into size fractions and counting a statistically useful number of items per sample (Van der Veen & Fieller 1982). The creation of sample-level data means analysis is often undertaken on the basis of individual samples, for instance the assessment of crop-processing stages and weed ecological evidence for crop husbandry practices. The analysis of sample level data also enables archaeobotanical finds to be integrated

alongside contextual evidence from archaeological sites. Requirements for the publication of this data are in place in some archaeological guidelines, for instance current Historic England guidelines for archaeological practice in England (Campbell, Moffett & Straker 2011: 8).

From the earliest archaeobotanical reports, such as Reid's work at Roman Silchester, the sample from which plant remains were recovered was noted (Lodwick 2017a), but often results were reported as a list of taxa, or long catalogues of detailed botanical descriptions with seed counts, such as Knörzer's work at Neuss (Knörzer 1970). Early systematic archaeobotanical reports displayed data within in-text tables, for example Jones's work at Ashville (Jones 1978) and the two-way data table has been the standard form of reporting archaeobotanical data ever since. Often data tables are presented within book chapters or appendices, but the financial, space and time constraints of book publishing are limiting. Furthermore, there is the perception that specialist data was not necessary for publication (Barker 2001). Hence, alternative methods of the dissemination of specialist archaeological data were pursued in the later twentieth century.

From the 1980s, archaeobotanical data tables were often consigned to microfiche following a Council for British Archaeology and Department of Environment report (Moore & Richards 2015: 31), with the example of the excavation of Roman Colchester where the contents of all archaeobotanical samples were available on microfiche (Murphy 1992). An alternative in the 2000s was providing data tables on CD Rom as seen, for instance, in the CD accompanying the study of a Roman farmstead in the Upper Thames Valley (Robinson 2007) or the One Poultry excavations in London (Hill and Rowsome 2011). Meanwhile, the inception of the Archaeology Data Service, a digital repository for heritage data, in 1996 meant archaeological datasets were increasingly digitally archived, for instance the data from the Channel Tunnel Rail Link Project (Foreman 2018) or a recent large-scale research excavation at Silchester (University of Reading 2018). In these cases, archaeobotanical data is available to download as a .csv file.

Whilst the data publication strategy of large excavations was shifting, the availability of data from post-excavation assessment reports has remained challenging. So-called 'grey literature' results from the initial evaluation stage of developer-funded investigations and accompanying post-excavation assessment often contain a semi-quantitative evaluation of archaeobotanical samples on a scale of abundance. Whilst paper reports were initially deposited with county Historic Environment Records, a process of digitisation focussing on the Roman period has meant many pdfs are now available through the ADS (Allen et al. 2018), whilst born-digital reports are now deposited through OASIS (Online Access to the Index of archaeological investigationS), as part of the reporting process (Evans 2015), although the extent to which specialist appendices are included is variable.

These varying 'publication' strategies means archaeobotanical data is often available somewhere for recent developer-funded excavations and large-scale developer-funded

excavations, even if much of this data is as a printed table or .pdf file (Evans 2015; Evans and Moore 2014). However, academic journals are typically perceived as the most high-status publication venue for archaeobotanical data, and a crucial publication venue for academics in order to comply with institutional requirements and the norms of career progression. Aside from the problem of access to pay-walled journals by those without institutional subscriptions to all journals, the publication of primary data alongside research articles faces various problems, from the outright lack of inclusion of data, to problematic curation of supplementary data and a lack of peer review of data (Costello et al. 2013; Warinner and d'Alpoim Guedes 2014: 155; Whitlock, 2011). The extent of these problems for archaeobotany is currently unknown. Given the growth in archaeobotanical data production as methodologies are introduced into many new regions and periods over the last decade, it is vital that we know whether the mass of new data being produced is made available and is being reused.

Recent important advances within archaeobotanical data sharing have focussed on the construction of the ARBODAT database, developed by Angela Kreuz at the Kommission für Archäologische Landesforschung in Hessen. The database is used by a range of researchers in Germany, the Czech Republic, France and England (Kreuz & Schäfer 2002). Data sharing enabled by the use of this database has facilitated research on Neolithic agriculture in Austria, Bulgaria and Germany (Kreuz et al. 2005), and Bronze Age agriculture in Europe (Stika and Heiss 2012). The use of this database makes data integration between specialists easier due to the shared data structure and metadata description, but often the primary archaeobotanical data is not made publicly available.

2.2. Meta-analysis in archaeobotany

Beyond the need to preserve information, a key reason for the formal sharing of archaeobotanical data is in its reuse to facilitate subsequent research. There has been a long-standing concern within archaeobotany with the need to aggregate datasets and identify temporal and spatial patterns. The palaeobotanist Clement Reid maintained his own database of Quaternary plant records in the late nineteenth century (Reid 1899), which formed the foundation of Godwin's Quaternary database (Godwin 1975). Mid-twentieth century studies of prehistoric plant use compiled lists of archaeobotanical materials incorporating full references and the location of the archive (Jessen & Helbaek 1944). The International Work Group for Palaeoethnobotany was itself founded in 1968 in part with the aim to compile archaeobotanical data, first realised through the publication of *Progress in Old World Palaeoethnobotany* (Van Zeist, Wasylikowa & Behre 1991), and subsequently through the publication of annual lists of new records of cultivated plants (Kroll 1997).

To take England as an example, regional reviews produced by state heritage authorities have provided catalogues of archaeobotanical datasets in particular time periods and regions (e.g. Murphy 1998). When one archaeobotanist has undertaken the majority of study within a

region, pieces of synthesis within books have provided a relatively comprehensive review, for instance in the Thames Valley, UK (Lambrick & Robinson 2009). Over the last decade regional synthesis has occurred within several funded reviews which produced catalogues of sites with archaeobotanical data (Lodwick 2014; McKerracher 2018; Parks 2012) and a series of funded projects in France have enabled regional synthesis (Lepetz & Zech-Matterne 2017). However, many of these reviews are not accompanied by an available underlying database, and draw upon reports which are themselves hard to access.

Through the 1990s and 2000s, a series of databases were constructed in order to collate data from sites in a particular region and facilitate synthetic research. However, these databases have all placed the role of data archiving onto later projects specifically funded to collate data, rather than sourcing datasets at the time of publication. Such a model is unsustainable, and is unlikely to result in all available datasets being compiled. The *Archaeobotanical Computer Database* (ABCD), published in 1996 in the first issue of *Internet Archaeology*, contained much of the archaeobotanical data from Britain available at the time of publication, largely at the level of individual samples. The database was compiled between 1989 and 1994 and is still accessible through the accompanying online journal publication (Tomlinson & Hall 1996). The ABCD made major contributions to recent reviews of the Roman and Medieval periods (Van der Veen, Livarda & Hill 2008; Van der Veen, Hill & Livarda 2013). However, the database could only be centrally updated, with the online resource remaining a static version, lacking much of the new data produced subsequent to the implementation of PPG16 in 1990. The ADEMNES database, created through a research project undertaken at the Universities of Freiburg and Tübingen, contains data from 533 eastern Mediterranean and Near Eastern sites (Riehl & Kümmel 2005). Kroll has maintained the *Archaeobotanical Literature Database* to accompany the *Vegetation History and Archaeobotany* articles (Kroll 2005) now accessible as a database (Kirleis & Schmültz 2018). Numerous other databases have collated archaeobotanical studies, including the COMPAG project (Fuller et al. 2015), the *Cultural Evolution of Neolithic Europe* project (Colledge 2016), RADAR in the Netherlands (van Haaster and Brinkkemper 1995), *BRAIN Botanical Records of Archaeobotany Italian Network* (Mercuri et al. 2015) and *CZAD* – Archaeobotanical database of Czech Republic (CZAD 2019).

The majority of databases have a restricted regional coverage, whilst research-project driven period-specific databases provide overlapping content. Whilst there are a wide range of archaeobotanical databases available, few contain primary datasets (other than the ABCD) which can be downloaded as .csv files. Data which is most commonly available are bibliographic references per site, with some indications of mode of preservation, quantity of archaeobotanical data, and sometimes taxa present. The databases do not inter-relate to each other, and function primarily as bibliographic sources enabling researchers to find comparative sites or to identify published datasets which need to be re-tabulated prior to meta-analysis. The

IWGP website curates a list of resources, but otherwise the resources are often disseminated through the [archaeobotany jiscmail list](#).

Beyond the aim of cataloguing archaeobotanical data within a region and period, meta-analysis is often used in archaeobotany to identify spatial and chronological trends in a range of past human activities, for instance crop choice, crop husbandry practices, plant food consumption, the trade in luxury foods or the use of plants in ritual. Meta-analysis can be undertaken on the basis of simple presence/absence data per site, but in order for such analysis to be rigorous and comparable, sample-level data must be utilised. For instance, sample-level data is required for meta-studies, in order to identify high-quality samples of unmixed crops for weed ecology analysis (Bogaard 2004), to assess the importance of context in the evaluation of wild plant foods (Wallace et al. 2019), or to use volumetric measurements as a proxy for scale (Lodwick 2017b). The reuse of archaeobotanical data also extends to include datasets used as “controls” in commonly used forms of statistical analysis, for instance Jones’s weed data from Amorgos, Greece, which is utilised as a control group in discriminant analysis of crop-processing stage (Jones 1984), and ethnographic observations of crop items in different crop-processing stages (Jones 1990).

2.3. Open data principles and solutions

Debates over issues of data publication and meta-analysis have been on-going across scientific disciplines over the last decade (Editors 2009), and have been summarised within principles of open science, as recently set out in relation to archaeology (Marwick et al. 2017). Open Data is one of the three core principles for promoting transparency in social science (Miguel et al. 2014). The FAIR principles, developed by representatives from academia, industry, funding agencies, industry and publishers, provide four principles which data sharing should meet for use by both humans and machines – Findability, Accessibility, Interoperability, and Reusability (Wilkinson et al. 2016). A recent report assessing the adoption and impact of FAIR principles across academia in the UK included archaeology as a case study (Allen and Hartland 2018: 46). It reported how the ADS was often used to archive data, but that “The journal itself provides the “story” about the data, the layer that describes what the data is, how it was collected and what the author thinks it means.” The report also raises the problem that smaller projects may not have the funding to utilise the ADS, meaning that other repositories are utilised. Increasingly, archaeological data is made available through a wide range of data repositories ([OSF](#), [Mendeley Data](#), [Zenodo](#), [Open Context](#)), university data repositories (e.g. [ORA-Data](#)), or social networking sites for academics ([Academia.edu](#), [ResearchGate](#)). More widely in archaeology, some have observed that archaeological data is rarely published (Kintigh et al. 2014), and recent reviews have reported low levels of data sharing (Huggett 2018; Marwick & Pilaar Birch 2018). A closely related issue is that of data reuse. Responsible reuse of primary data encourages the sharing of primary data (Atici et al. 2013), but levels of data reuse in archaeology are thought to remain

low (Huggett 2018). Principles for responsible data citation in archaeology have recently been developed summarising how datasets should be cited (Marwick & Pilaar Birch 2018).

3. Methodology

In order to assess the current status of data sharing, citation and data re-use in archaeobotany, a review was undertaken of the publication of primary data and the publication of meta-analysis in major archaeological journals over the last ten years, building on recent pilot studies within archaeology (Marwick & Pilaar Birch 2018). The review of academic journals provided a contrast to recent assessments of archaeobotanical data deriving from developer-funded archaeology (Lodwick 2017c; Van der Veen, Hill & Livarda 2013). Journal articles have been selected as the focus of this study as the provision of online supplementary materials in the majority of journals and the ability to insert hyperlinks to persistent identifiers (eg a DOI) to link to datasets available elsewhere should not limit the publication of data and references. Much archaeobotanical data is also published elsewhere, especially from projects not based in the university sector, that is commercial or community archaeology in the UK. Archaeobotanical datasets emanating from this research are more commonly published through monographs, county journal articles, and unpublished (or grey literature) reports, but these are beyond the scope of the current review.

All journal articles were included which represent the principle reporting of a new archaeobotanical assemblage. The selected journals fall within three groups. First, what is considered the specialist archaeobotanical journal (*Vegetation History and Archaeobotany* (VHA)). Second, archaeological science journals (*Archaeological and Anthropological Sciences*, *Environmental Archaeology*, *The Holocene*, *Journal of Archaeological Science* (JAS), *Journal of Archaeological Science: Reports* (JASR), *Journal of Ethnobiology*, *Quaternary International*, *Journal of Wetland Archaeology*), which can be considered as specialist sub-disciplinary journals which should be maintaining data-quality. Third, general archaeology journals (*Antiquity*, *Journal of Field Archaeology*, *Oxford Journal of Archaeology*, *Journal of Anthropological Archaeology*, *Journal of World Prehistory*). Finally, the broader cross-disciplinary journals *PLoS One* and *Proceedings of the National Academy of Sciences* (PNAS) were included. Published articles from the past ten years (2009–2018) have been analysed in order to assess the availability of plant macrofossil data. This ten-year period brackets the period where most archaeological journals have moved online and adopted supplementary materials.

Data citation in synthetic studies has been assessed in the same range of publications. The extent of data reuse ranges from the analysis of whole sample data to the presence/absence of individual crops. The location of a data citation has been assessed in the same range of publications, with the addition of journals where occasional research incorporating archaeobotanical data is featured (*Britannia*, *Journal of Archaeological Research*, *Ethnobiology Letters*, *Medieval Archaeology*, *Proceedings of the Prehistoric Society*, *World Archaeology*). The

underlying dataset for the analysis is available in Lodwick 2019.

4. Findings

4.1. Primary data sharing

Here, the location of primary archaeobotanical data, that is sample level counts of macroscopic plant remains, was assessed for 239 journal articles across 16 journals (Lodwick 2019 Table 1). **Figure 1** shows the results grouped by journal. Overall, only 56% of articles shared their primary data. In, *Antiquity*, *JAS*, *JASR*, *PLOS One*, *Quaternary International* and *VHA*, the highest proportion of publications did not include their primary data, that is to say that the sample-by-sample counts of plant macrofossils was not available. This level of data is comparable to the findings of other pilot studies in archaeology. Marwick and Pilaar Birch found a data sharing rate of 53% from 48 articles published in *Journal of Archaeological Science* in Feb – May 2017 (Marwick & Pilaar Birch 2018: 7), and confirm previous assertions that data is often withheld in archaeology (Kansa 2012: 499). This is better than some disciplines, with a 9% data sharing rate on publication found across high impact journal science publications ($n = 500$) (Alsheikh-Ali et al. 2011) and 13% in biology, chemistry, mathematics and physics ($n = 4370$) (Womack 2015), yet still indicates that nearly half of articles did not include primary data. Primary archaeobotanical data is more likely to be shared in archaeobotanical and archaeological science journals than general archaeology journals. However, within the primary archaeobotanical journal, *VHA*, 51% of articles do not include their primary data (**Figure 1**).

Where primary data was not shared, the data which was available ranged from summary statistics, typically

counts or frequencies, reported either by site, site phase, or feature group. **Figure 2** summarises these results by year, showing that there is a gradient within articles not sharing their full 'raw' data, from those only provided sample counts on one aspect of the archaeobotanical assemblage, to those only presenting data graphically or within discussion. Beyond full data, the most common form of data shared is either summary counts per site or summary counts per feature or phase. Whilst this data does enable some level of reuse, the results of any sample-level data analysis presented within an article cannot be verified, and the data cannot be reused for crop-processing or weed ecology analysis which requires sample level data. Furthermore, such data would have been collected on a sample-by-sample basis, but this information is lost from the resulting publication.

The forms in which data are made available vary across journals. The sharing of primary data within an article remains the most common data sharing form in archaeobotany (**Figure 1**). Data tables in text require manual handling to extract data, in journals such as *VHA*, whilst in other journals in-text tables can be downloaded as .csv files. These however would not be citable as a separate dataset. Supplementary datasets are the third most common form of data sharing. Indeed, the use of electronic supplementary material has been advocated recently for by some journals, such as the *Journal of Archaeological Science* (Torrence, Martínón-Torres & Rehren 2015). Microsoft Excel spreadsheets are the most common form of supplementary data, followed by .pdfs and then word documents (**Figure 1**). Both .xlsx and .docx are proprietary file formats, and not recommended for long term archiving or open science principles. There is no indication of

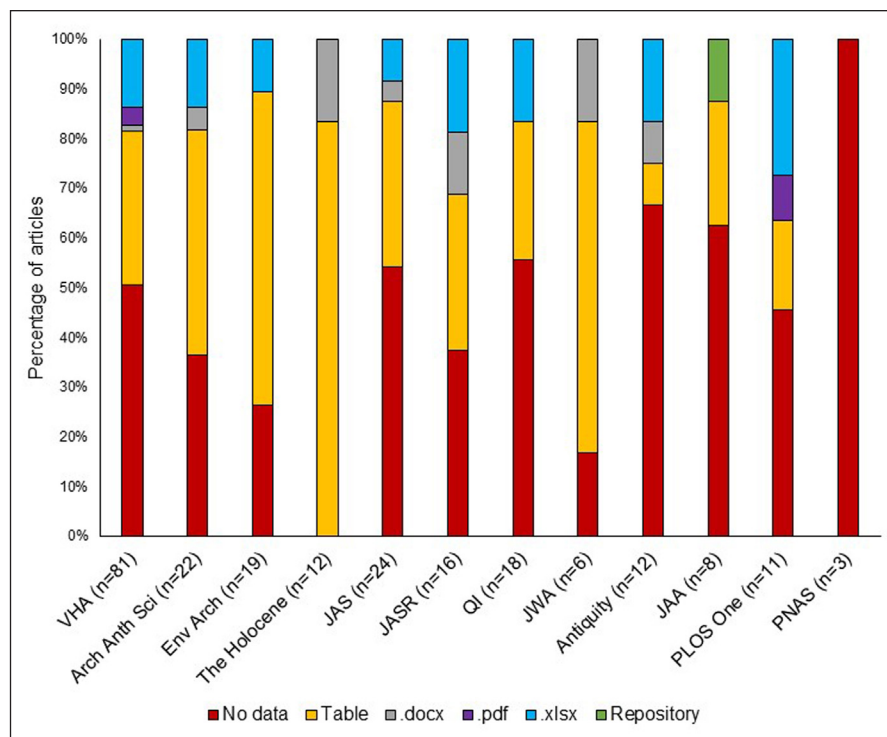


Figure 1: Chart showing the location of primary archaeobotanical data by journal in primary archaeobotanical data publications.

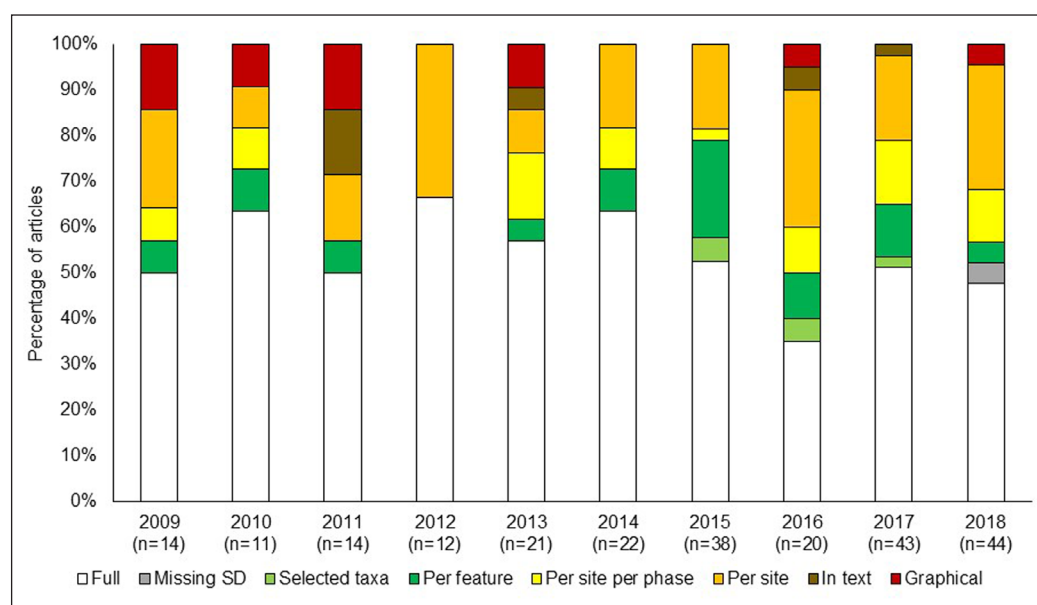


Figure 2: Chart showing the form of archaeobotanical data shared by year in primary archaeobotanical data publications.

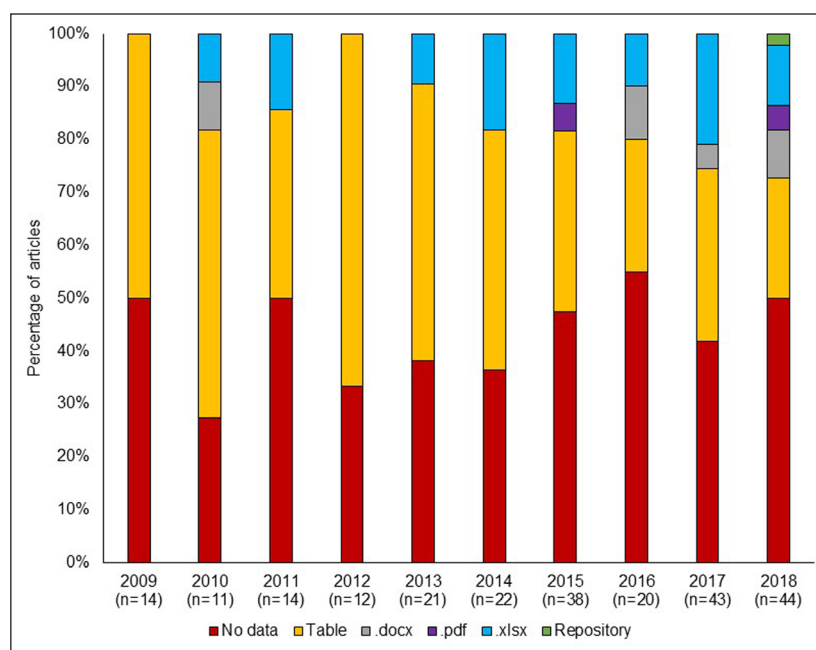


Figure 3: Chart showing the location of archaeobotanical data from 2009–2018 in primary archaeobotanical data publications.

improvement over the last decade in the form of data sharing. In 2018, 50% of articles did not share their primary data, and where the data was shared, it was in proprietary forms (.docx, .xlsx) or those that do not easily facilitate data reuse (.pdf) (Figure 3).

Just one of the articles included in this review incorporated a dataset archived in a repository (Farahani 2018), in contrast to the substantial growth in data repositories across academic disciplines (Marcial & Hemminger 2010). Other examples provide the underlying data for monograph publications, such as that of the archaeobotanical data from Gordion, Turkey (Marston 2017a, 2017b), Silchester, UK (Lodwick 2018; University of

Reading 2018) and Vaihingen, Germany (Bogaard 2011a; Bogaard, 2011b).

Several of the journals that have been assessed have research data policies. In the case of *Vegetation History and Archaeobotany*, sufficient papers have been surveyed to assess the impact of the research data policy on the availability of data. Figure 4 show the proportion of data sharing formats through time just for *VHA* (note the small sample size). The introduction of a research data policy in 2016 encouraging data sharing in repositories has not resulted in any datasets being shared in that format. Of the 10 articles published in *PLOS One* after the introduction of a clear research data policy in 2014, 4 did not

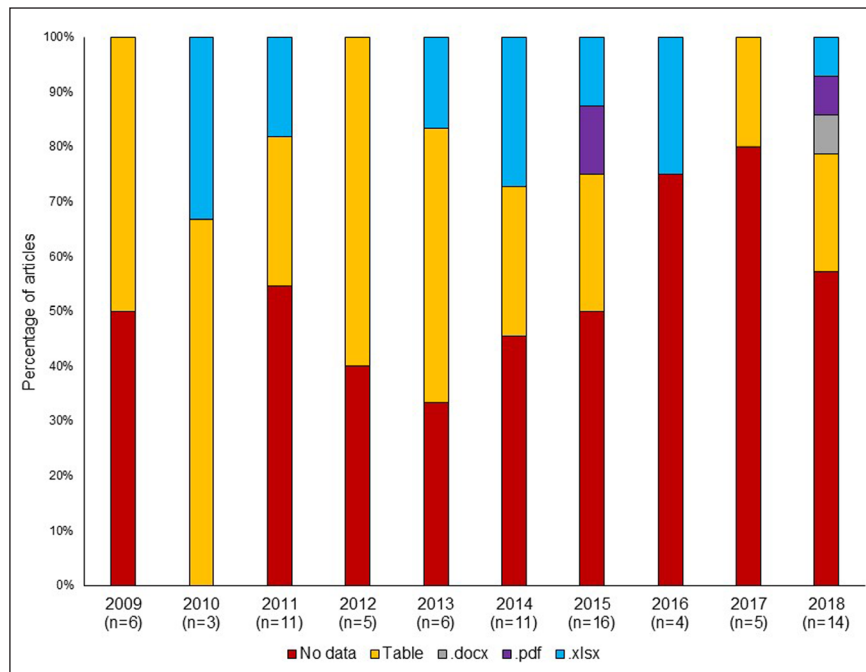


Figure 4: Chart showing the location of primary archaeobotanical data in *Vegetation History and Archaeobotany*.

contain primary data. However, elsewhere, journals with no research data policy, such as *Antiquity*, has one of the lower levels of data sharing (Figure 1).

There are various reasons for why a primary dataset may be lacking. The option of providing supplementary datasets has been available in many of the journals here since before the start of the surveyed period (e.g. *Vegetation History and Archaeobotany* in 2004), and so cannot be a reason for the absence of data publication in this journal while it may be a reason in other journals. Reasons suggested for a lack of data sharing within archaeology include technological limitations, and resistance amongst some archaeologists to making their data available due to cautions of exposing data to scrutiny, lost opportunities of analysis before others use it and loss of ‘capital’ of data (Moore & Richards 2015: 34–35). Furthermore, control over how data tables is presented (taxa ordering, summary data presented) may also contribute to the preferential publishing of data within journal articles. Another factor to consider is the emphasis on the creation of new data through archaeological research (Huvila 2016). The creation of a new archaeobotanical dataset through primary analysis is a key form of training in archaeobotany, and the perception of the value of the reuse of other previously published archaeobotanical journals may be low, hence not encouraging the sharing of well-documented datasets. Excellent exams of data reuse have resulted in influential studies (Bogaard 2004; Riehl 2008; Wallace et al. 2019), and would hopefully encourage further data sharing in the future.

Given that there are numerous examples of meta-analysis which do take place in archaeobotany, it seems likely that the prevalent form of data sharing is through informal data sharing between individual specialists. However, this does not improve access to data in the long term, and is inefficient and time consuming, with large potential for data errors (Kansa & Kansa 2013), and relies on personal

networks, which are likely to exclude some researchers. The absence of primary data in many archaeobotanical publications thus inhibits the verification of patterns observed within a dataset, and strongly limits the re-use potential of a dataset.

4.2. Data citation

One of the common arguments for increasing data sharing is an associated increase in the citation of the articles which have data available. Here, the data citation practices of meta-analyses of plant macrofossil data undertaken over the last decade have been reviewed. 20 journals were consulted, including a wider range of period-specific journals, and 107 articles were assessed (Lodwick 2019 Table 2). Data citation was assessed as ‘in text’ or ‘in table’ to refer to when the citation and the bibliographic reference were within the article, as ‘in supplementary data’ when the citation and reference were within the supplementary materials, and as ‘no citation’ when no citation and reference was provided.

21% of articles (n = 22) did not contain any citations to the underlying studies. 16% (n = 17) contained citations within supplementary data files. 50% of articles (n = 53) contained a citation within a table within the main article, and 14% (n = 15) contained citations within the main text. For the 21% of articles without data citations, the results of these studies could not be reproduced without consulting individual authors. The papers supplying the underlying data also received no credit for producing these datasets. Where articles contain citations within the main article (in text or table), full credit is provided to the underlying studies, a citation link is created through systems such as google scholar, and the study can be easily built upon in the future. Where the citation is provided within supplementary data, the original studies do receive attribution, but are not linked to so easily.

Through time, there is a steady decrease in the proportion of studies without citations to the underlying data, whereby of the 17 meta-analysis articles published in 2018, only one had no data citations. In comparison, in 2009, 3 out of 8 meta-analysis articles contained no data citation (Figure 6). Overall this is a more positive outlook on the reuse of published data, but the consistent presence of articles lacking data citation indicates that improvements are needed. Reasons for a lack of data citation may include restrictions on word counts imposed by journals, a lack of technical knowledge in making large databases available,

or the wish to hold on to a dataset to optimise usage. Considering the type of journal (Figure 5), levels of data citation are worse in general archaeology journals, with sub-disciplinary journals showing slightly better levels of data citation. In particular *VHA* has a lack of consistency in where data citations are located.

4.3. Reuse of archived archaeobotanical datasets

The majority of data citations assessed in the previous section are to articles or book chapters rather than datasets. The ADS currently hosts 66 data archives which have

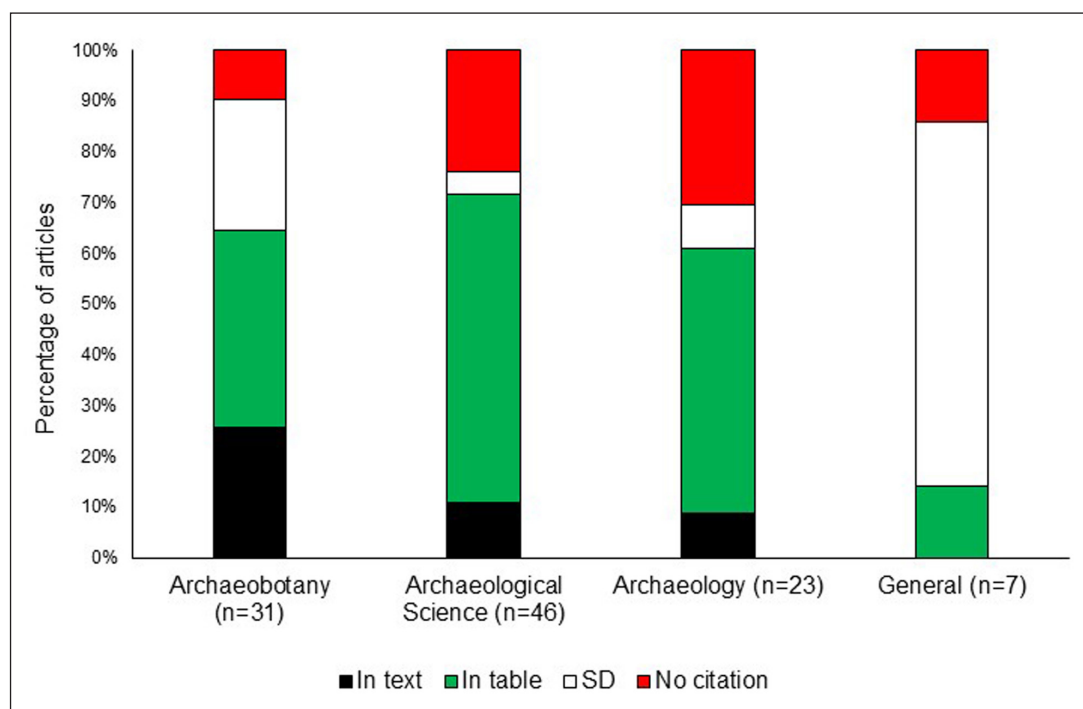


Figure 5: Chart showing the location of data citations in meta-analysis journal articles by journal type.

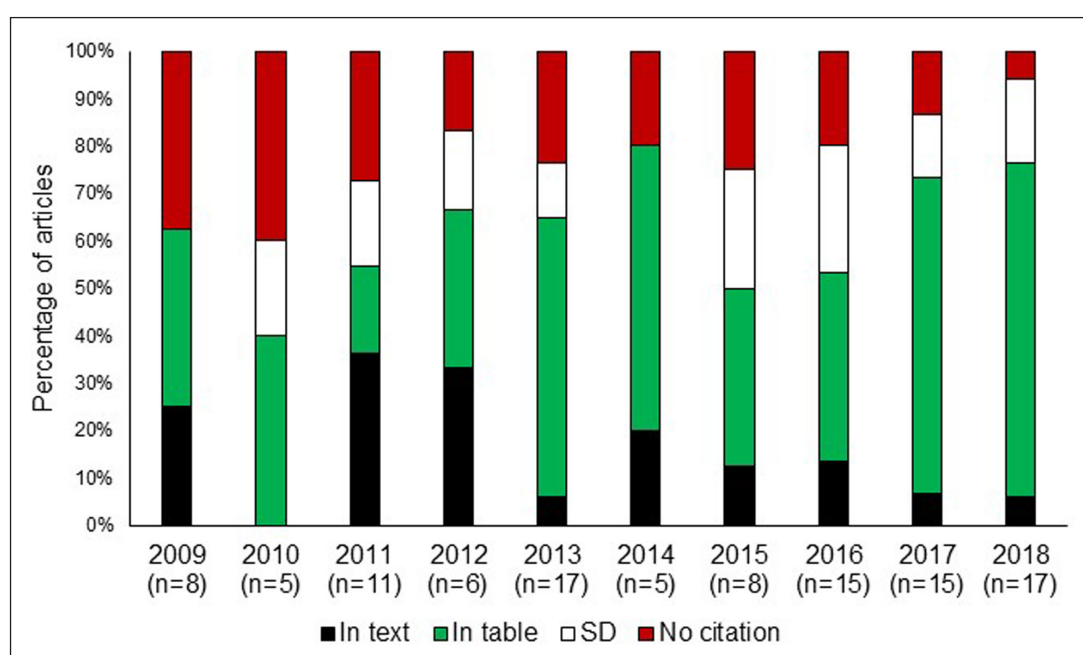


Figure 6: Chart showing the location of data citations in meta-analysis journal articles from 2009–2018.

been tagged as containing plant macro data, deriving mainly from developer-funded excavations but also some research excavations. However, in some of these the plant macro data is contained within a pdf. As, the archiving of archaeobotanical datasets in data repositories is still at an early stage, the reuse of these datasets is assessed here on a case-by-case basis. The archaeobotanical dataset from the Neolithic site of Vaihingen, Germany (Bogaard 2011b) has not been cited on google scholar. Metrics are provided through the ADS, showing this dataset has been downloaded 56 times with 477 individual visits (as of 25/2/19). The archaeobotanical dataset from Gordion by Marston has no citations on Google Scholar (Marston 2017b), neither does the Giza botanical database (Malleson & Miracle 2018), but these are both very recently archived datasets. In contrast, the Roman Rural Settlement Project dataset, which includes site-level archaeobotanical data, has received greater levels of use, with 12 citations in Google Scholar, over 40,000 file downloads, and over 35,000 visits (Allen et al. 2018) and the archaeobotanical computer database (Tomlinson & Hall 1996) has been cited 44 times, and is the major dataset underpinning other highly-cited studies (Van der Veen, Livarda & Hill 2008; Van der Veen, Hill & Livarda 2013). Whilst there is clearly precedence for the reuse of archaeobotanical databases, current data citation practices within archaeobotany do not yet appear to be formally citing individual datasets, meaning an assessment of the reuse of archived archaeobotanical datasets is challenging.

5. Steps Forward

This review of data sharing, citation, and reuse practices in archaeobotany has found medium levels of data sharing, good levels of data citation, but so far limited levels of reuse of archived data sets. This picture is similar across archaeology, in part attributed to the status of archaeology as a small-science, where data-sharing takes place ad-hoc (Marwick & Pilaar Birch 2018). Here, recommendations are discussed for improving these data practices within archaeobotany, of applicability more widely in archaeology.

Clearly an important step is improving the sharing of plant macrofossil data. Given the reasonable small size of most archaeobotanical datasets (a .csv file < 1mb), and a lack of ethical conflicts, there seems to be few reasons why the majority of archaeobotanical data couldn't be shared. In the case of developer-funded derived data, issues of commercial confidentiality could limit the sharing of data. A key stage is establishing why levels of data sharing are not higher. Issues within archaeobotany may include the conflict between having to publish results within excavation monographs, which may take some time to be published, and have limited visibility due to high purchase costs and no digital access, and the need to publish journal articles for career progression within academia. The production of an archaeobotanical dataset is very time-consuming, and interim publication on notable aspects of an assemblage may be considered as a necessary publication strategy. More broadly, one important aspect is issues of equity in access to digital archiving resources (Wright &

Richards 2018), such as differential access to funds, training and knowledge. A recent study in Sweden found that we need to know concerns, needs, and wishes of archaeologists in order to improve preservation of archaeological data (Huvila 2016), especially when control of ones data may be linked to perceptions of job security. In order to make improvements in data sharing and reuse across archaeology, we need improved training in data sharing and the reuse of data in higher education (Touchon & McCoy 2016; Cook et al. 2018), improved training in data management (Faniel et al. 2018), and crucially, the necessary software skills to make the reuse of archived datasets attainable (Kansa & Kansa 2014: 91). Examples of good practice in archaeobotany are the Vaihingen and Gordion datasets which demonstrate how datasets can be archived in data repositories to accompany a monograph (Bogaard 2011b; Marston 2017b), whilst Farahani (2018) provides an excellent example of a journal article, where the primary data is supplied as a .csv in a cited data repository along with the R script for the analysis.

In tandem with the need to encourage authors to share their data, is the need for journals to create and implement research data policies. Given the existence of research data policies in many of the journals included here, this reflects other findings of the poor enforcement of data policies by journals (Marwick & Pilaar Birch 2018), supporting arguments that journals should not be relied upon to make data accessible, and data should instead be deposited in digital repositories. In order to implement change in data sharing, there is a role to play for learned societies and academic organisation in lobbying funding bodies, prioritising data sharing in research projects. A key step is through journal editorial boards, and the enforcement of any pre-existing research data policies (Nosek et al. 2015). Reviewers should be required to comment on the availability and format of data, and editorial board members should take this into account when making decisions on papers. In North America the recent formation of the Society of American Archaeology Open Science Interest Group seeks to galvanise open research practices in archaeology (Marwick et al. 2017).

Beyond moves towards the sharing of data, we also need improvements in how data is shared in archaeology. Whilst supplementary data tables are a common option for sharing data (**Figure 1**), these contain problems of a lack of peer-review, lack of curation compared to the main article, and link decay (Warinner & d'Alpoim Guedes 2014: 155; Whitlock 2011; Costello et al. 2013). Several articles reviewed here had missing supplementary data files which should have contained the raw data. However, the use of supplementary data files is often promoted by journal editors to save the word count of articles. Furthermore, publishing of data within journal articles is not an ideal solution. This often places data behind a paywall, inaccessible to those not working for institutions (typically universities) which subscribe to journals, does not guarantee the long-term preservation of data (Smith 2008), and does not enable content searches just for datasets.

Another route to sharing comes through formal data publication (Kansa & Kansa 2013; Moore & Richards

2015: 38). Whereby data sets are published as data papers and assigned a DOI, for example the *Journal of Open Archaeology Data* provides a venue for the publication of datasets archived in repositories with examples of the publication of archaeobotanical datasets from Neolithic Europe (Colledge 2016) and Bronze Age Ireland (Johnston 2014). This practice ensures that datasets are peer-reviewed, quality checks are made, authors receive increased credit, and the survivability of data is increased (Costello et al. 2013; Kansa & Kansa 2013; Tennant et al. 2016). A number of data archives are available for archaeological datasets which incorporate peer review, such as tDAR (The Digital Archaeological Record), Open Context and ADS. Other options available are PANGAEA or Dryad (Warinner & d'Alpoim Guedes 2014). Many academic institutions also provide data repositories, but these would not facilitate specialist peer review of the dataset prior to making it publicly available. It should be remembered that the sharing of data is usually not sufficient to enable reuse, and meta-data must be provided (Atici et al. 2013). Metadata specific to archaeobotanical data would be the taxonomy used, context and sample number, sample volume, float and mesh size, flotation technique, sub-sampling specifications, along with site meta-data. Whilst this paper has considered archaeobotanical datasets in isolation, all datasets were born from an entire archaeological excavation – and need to be integrated alongside other specialist's data (Faniel et al. 2018: 114). A drawback of sharing archaeobotanical datasets separately from information on the rest of the archaeological site is that they may lack sufficient information about the site stratigraphy and chronology to enable accurate reuse.

Work is also needed in agreeing standardised formats for referring to archaeobotanical taxa, which would make reuse easier and quicker. Currently, archaeobotanical datasets use different names depending on which flora is being referred to, making the larger-syntheses of datasets challenging. Currently, ArboDat users produce datasets with the same plant codes or P codes – a seven letter code referring to plant family, genus or species based on *Flora Europaea* (Kreuz and Schäfer 2002; Halliday and Beadle 1982). However, the database is mainly designed for use in central and western Europe, relies on central agreement for the addition of new species, and is also requires the use of proprietary software (Microsoft Access). A further advance would be the adoption of Linked Open Data concepts in archaeobotany, a step which has been taken in zooarchaeology (Kansa 2015; LeFebvre et al. 2019). Linked Open Data would not enforce the universal adoption of the same flora, but through referring to persistent taxonomic identifiers, such as the *Encyclopedia of Life*, would enable the connection of archaeobotanical and biological data. However, the use of standardised data archiving formats and naming systems must come through community agreement.

Improvements in data citation within journal articles must come from journal editors. Growing numbers of journals have Research Data Policies which encourage the citation of sources of underlying data. Again, the implementation of these lies with the journal editorial boards. Further steps towards reproducibility in archaeobotany is

the use of code driven analysis rather than point-and-click. It has been recently argued that a tool-driven scientific revolution in the use and sharing of code is emerging in archaeology (Marwick & Schmidt 2019). Two examples of archaeobotanical papers which have shared their code are d'Alpoim Guedes et al. (2015) and Sinensky and Farahani (2018). Such work in standard archaeobotany is limited, but related examples include the creation of a R package for morphometric analysis (Bonhomme et al. 2014), and the use and sharing of code to analyse climate archaeology (d'Alpoim Guedes, Jin & Bocinsky 2015) and stable isotope analysis results from charred plant remains (Styring et al. 2017).

The benefits of improved data sharing practices are wide-ranging. Improvements to data citation practices would ensure that authors gain citation credit for increasing the reusability of their data and encourage more formal data sharing of archaeobotanical data in the future (Kansa & Kansa 2014; Marwick & Pilaar Birch 2018). Beyond the practicalities of ensuring one's own data is not lost, and future meta-analysis is more successful, improving data practices in archaeobotany would stop the fragmentation of the field by ensuring reuse of archaeobotanical data by others within and beyond archaeobotany can take place (Marston, Warinner & D'Alpoim Guedes 2014: 12). Furthermore, there is hope that access to primary data would encourage debate within the discipline, advance intellectual understanding and have economic benefits (Moore & Richards 2015: 30). There are numerous benefits for increasing openness of archaeobotanical data such as increased potential for new scientific findings, increased professionalism and increased scientific rigour (Moore & Richards 2015: 34–35). However, there are currently insufficient reward mechanisms to motivate data sharing and other open research practices within archaeology. The Open Science Framework has created badges which can be assigned to journal articles which meet open science practices, but the uptake in archaeology journals is currently limited to *Internet Archaeology* (Nosek et al. 2015).

A recent FAIR report has stated that standards in data sharing in archaeology are maintained through the publication of good practice articles (Allen and Hartland 2018: 23). However, it is worth considering whether a more top-down approach through disciplinary institutions is needed to encourage better data sharing and reuse practices, especially given the lack of better data sharing practices in sub-discipline specific journals (**Figures 1 and 4**) and the absence of improvements in data sharing through time (**Figure 3**). Given the advances made in recent years by zooarchaeology, a field in many ways similar to archaeobotany in term of its position in archaeology more broadly and in terms of the excavation process, there is much scope for improvement.

6. Conclusions

In undertaking one of the largest assessments of data sharing and citation practices in archaeological research published within peer-reviewed journals, this article has found that levels of data sharing of archaeobotanical data

require improvement, with only 56% of 239 papers sharing their primary data. The data that is being shared is not in forms amenable to reuse. This picture is likely to be similar across archaeology. Data citation practices are better, with 79% of 107 articles citing their data sources, albeit with exceptions in major journals. The formal publication of datasets and reuse of them is so far limited within archaeobotany. This low level of data sharing and lack of data citation prohibits future work in archaeobotany, increases errors in meta-analysis studies, and prohibits the reuse of archaeobotanical data by other disciplines. A number of pathways to promote the sharing and reuse of archaeobotanical data have been discussed, including improved author guidelines in journals, and increased rewards for data sharing. Improvements in data citation standards would assist with promoting data sharing, along with the creation of meta-data standards, whilst the sharing and reuse of archaeobotanical data in other publication formats is required. Finally, it is vital that such issues are discussed within the archaeobotanical and broader archaeological community.

Data Accessibility

The data that support the findings of this study are openly available in Oxford University Research Archive at <http://doi.org/10.5287/bodleian:7QkmEEp9Y>.

Acknowledgements

I am grateful to Jasmine Nirody, David Addison and Elizabeth Stroud for providing comments on the text, and for the very helpful comments from the peer reviewers.

Funding Information

This research was completed during a research fellowship at All Souls College, University of Oxford.

Competing Interests

The author has no competing interests to declare.

References

- Allen, M.**, et al. 2018. *The Rural Settlement of Roman Britain: an Online Resource [data-set]*. York: Archaeology Data Service [distributor]. DOI: <https://doi.org/10.5284/1030449>
- Allen, R** and **Hartland, D.** 2018. FAIR in Practice – JISC Report on the Findable Accessible Interoperable and Reuseable Data Principles (Version 1). *Zenodo*. DOI: <https://doi.org/10.5281/zenodo.1245568>
- Alsheikh-Ali, AA, Qureshi, W, Al-Mallah, MH and Ioannidis, JPA.** 2011. Public availability of published research data in high-impact journals. *PLOS ONE*, 6(9): e24357. DOI: <https://doi.org/10.1371/journal.pone.0024357>
- Atici, L.**, et al. 2013. Other people's data: a demonstration of the imperative of publishing primary data. *Journal of Archaeological Method and Theory*, 20(4): 663–681. DOI: <https://doi.org/10.1007/s10816-012-9132-9>
- Barker, G.** 2001. Agendas for environmental archaeology. In: Albarella, U (ed.), *Environmental Archaeology: Meaning and Purpose*. Dordrecht: Springer Netherlands. 305–313. DOI: https://doi.org/10.1007/978-94-015-9652-7_24
- Bogaard, A.** 2004. *Neolithic Farming in Central Europe: an Archaeobotanical Study of Crop Husbandry Practices*. London: Routledge. DOI: <https://doi.org/10.4324/9780203358009>
- Bogaard, A.** 2011a. *Plant Use and Crop Husbandry in an Early Neolithic Village: Vaihingen an der Enz, Baden-Württemberg*. Bonn: Habelt.
- Bogaard, A.** 2011b. *The Archaeobotany of Vaihingen an der Enz, Baden-Württemberg, Germany [data-set]*. York: Archaeology Data Service [distributor]. DOI: <https://doi.org/10.5284/1011849>
- Bonhomme, V.**, et al. 2014. Momocs: outline analysis using R. *Journal of Statistical Software*, 56(13): 1–24. DOI: <https://doi.org/10.18637/jss.v056.i13>
- Campbell, G, Moffett, L and Straker, V.** 2011. *Environmental Archaeology: A Guide to the Theory and Practice of Methods, from Sampling and Recovery to Post-excavation (second edition)*. Portsmouth: English Heritage.
- Colledge, S.** 2016. The cultural evolution of Neolithic Europe. EUROEVOL Dataset 3: Archaeobotanical Data. *Journal of Open Archaeology Data*, 5. DOI: <https://doi.org/10.5334/joad.42>
- Cooper, A and Green, C.** 2016. Embracing the complexities of “Big Data” in archaeology: the case of the English Landscape and identities project. *Journal of Archaeological Method and Theory*, 23: 271–304. DOI: <https://doi.org/10.1007/s10816-015-9240-4>
- Costello, MJ.**, et al. 2013. Biodiversity data should be published, cited, and peer reviewed. *Trends in Ecology & Evolution*, 28(8): 454–461. DOI: <https://doi.org/10.1016/j.tree.2013.05.002>
- CZAD.** 2019. Archaeobotanical database of Czech Republic. <http://www.arup.cas.cz/czad/index.php?l=en>. Last accessed 18/03/2019.
- d'Alpoim Guedes, J, Jin, G and Bocinsky, RK.** 2015. The Impact of climate on the spread of rice to north-eastern China: a new look at the data from Shandong province. *PLOS ONE*, 10(6): e0130430. DOI: <https://doi.org/10.1371/journal.pone.0130430>
- Editors.** 2009. Data's shameful neglect. *Nature*, 461: 145. DOI: <https://doi.org/10.1038/461145a>
- Evans, TNL.** 2015. A reassessment of archaeological grey literature: semantics and paradoxes. *Internet Archaeology*, 40. DOI: <https://doi.org/10.11141/ia.40.6>
- Evans, TNL and Moore, RH.** 2014. The use of PDF/A in digital archives: a case study from archaeology. *International Journal of Digital Curation*, 9(2): 123–138. DOI: <https://doi.org/10.2218/ijdc.v9i2.267>
- Faniel, I.**, et al. 2018. Beyond the archive: bridging data creation and reuse in archaeology. *Advances in Archaeological Practice*, 6(2): 105–116. DOI: <https://doi.org/10.1017/aap.2018.2>
- Farahani, A.** 2018. A 2500-Year historical ecology of agricultural production under Empire in Dhiban,

- Jordan. *Journal of Anthropological Archaeology*, 52: 137–55. DOI: <https://doi.org/10.1016/j.jaa.2018.09.006>
- Foreman, S.** 2018. *Channel Tunnel Rail Link Section 1 [data-set]*. York: Archaeology Data Service [distributor]. DOI: <https://doi.org/10.5284/1000230>
- Fuller, DQ**, et al. 2015. Comparing pathways to agriculture. *Archaeology International*, 18: 61–66. DOI: <https://doi.org/10.5334/ai.1808>
- Godwin, H.** 1975. *The History of the British Flora: a Factual Basis for Phytogeography*. Cambridge: Cambridge University Press.
- Halliday, G** and **Beadle, M.** 1982. *Consolidated Index to Flora Europaea*. Cambridge: Cambridge University Press.
- Hill, J** and **Rowsome, P.** 2011. *Roman London and the Walbrook Stream Crossing: Excavations at 1 Poultry and Vicinity, City of London Part 2*. London: MoLA Monograph 37.
- Huggett, J.** 2018. Reuse remix recycle: repurposing archaeological digital data. *Advances in Archaeological Practice*, 6(2): 93–104. DOI: <https://doi.org/10.1017/aap.2018.1>
- Huvila, I.** 2016. 'If we just knew who should do it', or The social organization of the archiving of archaeology in Sweden. *Information Research*, 21(2).
- Jacomet, S** and **Kreuz, A.** 1999. *Archäobotanik: Aufgaben, Methoden und Ergebnisse Vegetations- und Agrarcheologischer Forschung*. Stuttgart: Utb GmbH.
- Jessen, K** and **Helbaek, H.** 1944. *Cereals in Great Britain and Ireland in Prehistoric and Early Historic Times*. København: E. Munksgaard.
- Johnston, P.** 2014. Archaeobotanical data from two Middle and Later Bronze Age round house sites in Cork, Ireland. *Journal of Open Archaeology Data*, 3. DOI: <https://doi.org/10.5334/joad.ac>
- Jones, G.** 1984. Interpretation of archaeological plant remains: ethnographic models from Greece. In: Casparie, W and Van Zeist, W (eds.), *Plants and Ancient Man: Studies in Palaeoethnobotany*. Rotterdam: A. A. Balkema. 43–61.
- Jones, G.** 1990. The application of present-day cereal processing studies to charred archaeobotanical Remains. *Circaea*, 6(2): 91–96.
- Jones, G** and **Charles, M.** 2009. Charred seeds: quantification. In: Charles, M, et al. (eds.), *Archaeobotanical Online Tutorial* <<http://archaeobotany.dept.shef.ac.uk/>>.
- Jones, M.** 1978. The plant remains. In: Parrington, M (ed.), *The Excavation of an Iron Age Settlement, Bronze Age Ring-Ditches and Roman Features at Ashville Trading Estate, Abingdon (Oxfordshire) 1974–76*. Oxford: Oxford Archaeological Unit: 93–110.
- Kansa, E.** 2012. Openness and archaeology's information ecosystem. *World Archaeology*, 44(4): 498–520. DOI: <https://doi.org/10.1080/00438243.2012.737575>
- Kansa, EC** and **Kansa, SW.** 2013. Open Archaeology: We all know that a 14 is a sheep: data publication and professionalism in archaeological communication. *Journal of Eastern Mediterranean Archaeology & Heritage Studies*, 1(1): 88–97. DOI: <https://doi.org/10.5325/jeasmedarcherstu.1.1.0088>
- Kansa, SW.** 2015. Using Linked Open Data to improve data reuse in zooarchaeology. *Ethnobiology Letters*, 6(2): 224–231. DOI: <https://doi.org/10.14237/ebl.6.2.2015.467>
- Kansa, SW** and **Kansa, E.** 2014. Data publishing and Archaeology's information ecosystem. *Near Eastern Archaeology (NEA)*, 77(3): 223–227. DOI: <https://doi.org/10.5615/neareastarch.77.3.0223>
- Kintigh, KW**, et al. 2014. Grand challenges for archaeology. *PNAS*, 111(3): 879–880. DOI: <https://doi.org/10.7183/0002-7316.79.1.5>
- Kintigh, KW**, et al. 2018. Data integration in the service of synthetic research. *Advances in Archaeological Practice*, 6(1): 30–41. DOI: <https://doi.org/10.1017/aap.2017.33>
- Kirleis, W** and **Schmütz, K.** 2018. ArchbotLit CAU Kiel, <https://www.wikis.uni-kiel.de/archbotlit/>. [Last accessed 18/03/2019].
- Knörzer, K-H.** 1970. *Novaesium IV: Römerzeitliche Pflanzenfunde aus Neuss*. Berlin: Mann. Limesforschungen 10.
- Kreuz, A**, et al. 2005. A comparison of Early Neolithic crop and weed assemblages from the Linearbandkeramik and the Bulgarian Neolithic cultures: differences and similarities. *Vegetation History and Archaeobotany*, 14(4): 237–258. DOI: <https://doi.org/10.1007/s00334-005-0080-0>
- Kreuz, A** and **Schäfer, E.** 2002. A new archaeobotanical database program. *Vegetation History and Archaeobotany*, 11: 177–180. DOI: <https://doi.org/10.1007/s003340200019>
- Kroll, H.** 1997. Literature on archaeological remains of cultivated plants (1995/1996). *Vegetation History and Archaeobotany*, 6(1): 25–67. DOI: <https://doi.org/10.1007/BF01145883>
- Kroll, H.** 2005. *Literature on archaeological remains of cultivated plants 1981–2004*. Available at: <http://www.archaeobotany.de/database.html>.
- Lambrick, G** and **Robinson, M.** 2009. *The Thames Through Time: The Archaeology of the Gravel Terraces of the Upper and Middle Thames: The Thames Valley in Late Prehistory, 1500 BC-AD 50*. Oxford: Oxford University School of Archaeology. Thames Valley Landscapes Monograph No. 29.
- LeFebvre, MJ**, **Brenskelle, L**, **Wieczorek, J**, **Kansa, SW**, **Kansa, EC**, et al. 2019. ZooArchNet: Connecting zooarchaeological specimens to the biodiversity and archaeology data networks. *PLOS ONE*, 14(4): e0215369. DOI: <https://doi.org/10.1371/journal.pone.0215369>
- Lepetz, S** and **Zech-Matterne, V.** 2017. *Productions agropastorales, pratiques culturelles et élevage dans le nord de la gaule du deuxième siècle avant J.-C à la fin de la période romaine*. Quint-Fonsegrives: Editions Mergoil.
- Lodwick, L.** 2014. *An Archaeobotanical Analysis of Silchester and the Wider Region*. Unpublished DPhil

- thesis, University of Oxford. Available at: <https://ora.ox.ac.uk/objects/ora:11845>.
- Lodwick, LA.** 2017a. "The debatable territory where geology and archaeology meet": reassessing the early archaeobotanical work of Clement Reid and Arthur Lyell at Roman Silchester. *Environmental Archaeology*, 22(1): 56–78. DOI: <https://doi.org/10.1080/14614103.2015.1116218>
- Lodwick, L.** 2017b. Arable farming, plant foods, and resources. In: Allen, M. et al. (eds) *The Rural Economy of Roman Britain: Farming, Industry, Transport and Markets*. London: Society for the Promotion of Roman Studies. 11–84.
- Lodwick, L.** 2017c. *Approaches to the investigation, analysis and dissemination of work on Romano-British rural settlements and landscapes: a review. Paper 7: The recovery, reporting and analysis of plant macrofossils*. Available at: <http://cotswoldarchaeology.co.uk/wp-content/uploads/2017/08/RRS-Methodology-Paper-7-plants.pdf>.
- Lodwick, L.** 2018. The charred and waterlogged plant remains. In Fulford, M. et al. *Late Iron Age Calleva: The Pre-Conquest Occupation At Silchester Insula IX. Silchester Roman Town: The Insula IX Town Life Project: Volume 3*. London: Society for the Promotion of Roman Studies. 285–313.
- Lodwick, L.** 2019. Data sharing and citation in archaeobotany. University of Oxford. DOI: <https://doi.org/10.5287/bodleian:7QkmEEp9Y>
- Malleson, C and Miracle, R.** 2018. *Giza Botanical Database, Open Context*. DOI: <https://doi.org/10.6078/M7JH3J99>
- Marcial, LH and Hemminger, BM.** 2010. Scientific data repositories on the Web: An initial survey. *Journal of the American Society for Information Science and Technology*, 61(10): 2029–2048. DOI: <https://doi.org/10.1002/asi.21339>
- Marston, JM.** 2017a. *Agricultural Sustainability and Environmental Change at Ancient Gordion*. Philadelphia: University of Pennsylvania Museum Press. DOI: <https://doi.org/10.9783/9781934536926>
- Marston, JM.** 2017b. *Agricultural Sustainability and Environmental Change at Ancient Gordion (tDAR id: 427491)*, tDAR. DOI: <https://doi.org/10.9783/9781934536926>
- Marston, JM, Warinner, C and d'Alpoim Guedes, J.** 2014. Paleoethnobotanical method and theory in the twenty-first century. In: Marston, JM, d'Alpoim Guedes, J and Warinner, C (eds.), *Method and Theory in Paleoethnobotany*. Boulder: University Press of Colorado. 1–15. DOI: <https://doi.org/10.5876/9781607323167.c001>
- Marwick, B.** 2017. Computational reproducibility in archaeological research: basic principles and a case study of their implementation. *Journal of Archaeological Method and Theory*, 24: 424–450. DOI: <https://doi.org/10.1007/s10816-015-9272-9>
- Marwick, B, et al.** 2017. Open science in archaeology. *SAA Archaeological Record*, 17(4): 8–14. DOI: <https://doi.org/10.31235/osf.io/72n8g>
- Marwick, B and Pilaar Birch, SE.** 2018. A standard for the scholarly citation of archaeological data as an incentive to data sharing. *Advances in Archaeological Practice*, 6(2): 125–143. DOI: <https://doi.org/10.1017/aap.2018.3>
- Marwick, B and Schmidt, S.** 2019. Tool-driven revolutions in archaeological science. *SocArXiv*. DOI: <https://doi.org/10.31235/osf.io/4nkvx>
- McKerracher, M.** 2018. *Farming Transformed in Anglo-Saxon England: Agriculture in the Long Eighth Century*. Oxford: Windgather Press.
- Mercuri, AM, et al.** 2015. Pollen and macroremains from Holocene archaeological sites: A dataset for the understanding of the bio-cultural diversity of the Italian landscape. *Review of palaeobotany and Palynology*, 218: 250–266. DOI: <https://doi.org/10.1016/j.revpalbo.2014.05.010>
- Miguel, E, et al.** 2014. Promoting transparency in social science research. *Science*, 343: 30–31. DOI: <https://doi.org/10.1126/science.1245317>
- Moore, R and Richards, J.** 2015. Here today, gone tomorrow: open access, open data and digital preservation. In: Wilson, AT and Edwards, B (eds.), *Open Source Archaeology: Ethics and Practice*. Berlin: De Gruyter Open. 30–43. DOI: <https://doi.org/10.1515/9783110440171-004>
- Murphy, P.** 1992. *Environmental Studies: Culver Street*. In: Crummy, P (ed.), *Colchester Archaeological Report 6: Excavations at Culver Street, the Gilbert School, and Other Sites 1971–85*. Colchester: Colchester Archaeological Trust. 273–287.
- Murphy, P.** 1998. *A Review of Plant Macrofossils from Archaeological Sites in the Eastern Counties*. Centre of East Anglian Studies, University of East Anglia.
- Nosek, BA, et al.** 2015. Promoting an open research culture: author guidelines for journals could help to promote transparency, openness, and reproducibility. *Science*, 348: 1422–5. DOI: <https://doi.org/10.1126/science.aab2374>
- Parks, K.** 2012. *Iron Age and Roman Arable Practice in the East of England*. University of Leicester. Unpublished PhD thesis.
- Pearsall, DM.** 2016. *Paleoethnobotany: a Handbook of Procedures*. London: Routledge.
- Popper, V.** 1988. Selecting quantitative measurements in paleoethnobotany. In: Hastorf, C and Popper, V (eds.), *Current Paleoethnobotany. Analytical Methods and Cultural Interpretations of Archaeological Plant Remains*. Chicago: The University of Chicago Press. 53–71.
- Reid, C.** 1899. *The Origin of the British Flora*. London: Dulau. DOI: <https://doi.org/10.5962/bhl.title.24157>
- Richards, JD.** 2002. Digital preservation and access. *European Journal of Archaeology*, 5(3): 343–366. DOI: <https://doi.org/10.1179/eja.2002.5.3.343>
- Riehl, S and Kümmel, C.** 2005. *Archaeobotanical database of Eastern Mediterranean and Near Eastern sites*. Available at: <http://www.ademnes.de/database.php>.

- Robinson, M.** 2007. The environmental archaeology of the Cotswold Water Park. In: Miles, D, et al. (eds.), *Iron Age and Roman Settlement in the Upper Thames Valley: Excavations at Claydon Pike and Other Sites within the Cotswold Water Park*. Oxford: Oxford University School of Archaeology. 355–364.
- Sinensky, RJ and Farahani, A.** 2018. Diversity-disturbance relationships in the Late Archaic Southwest: implications for farmer-forager foodways. *American Antiquity*, 83(2): 281–301. DOI: <https://doi.org/10.1017/aaq.2017.74>
- Smith, K.** 2008. Institutional repositories and e-journal archiving: What are we learning? *The Journal of Electronic Publishing*, 11(1). DOI: <https://doi.org/10.3998/3336451.0011.107>
- Stika, H-P and Heiss, AG.** 2012. Bronzezeitliche landwirtschaft in Europa – der versuch einer gesamt-darstellung des forschungsstandes. *Studien zur nordeuropäischen Bronzezeit*. Neumünster: Wachholtz, 1: 189–222.
- Styring, AK, et al.** 2017. Isotope evidence for agricultural extensification reveals how the world's first cities were fed. *Nature Plants*, 3: 17076. DOI: <https://doi.org/10.1038/nplants.2017.76>
- Tennant, JP, et al.** 2016. The academic, economic and societal impacts of Open Access: an evidence-based review. *F1000Research*, 5: 632. DOI: <https://doi.org/10.12688/f1000research.8460.2>
- Tomlinson, P and Hall, A.** 1996. A review of archaeological evidence for food plants from the British Isles: an example of the use of the Archaeobotanical Computer Database (ABCD). *Internet Archaeology*, 1. DOI: <https://doi.org/10.11141/ia.1.5>
- Torrence, R, Martín-Torres, M and Rehren, T.** 2015. Forty years and still growing: Journal of Archaeological Science looks to the future. *Journal of Archaeological Science*, 56: 1–8. DOI: <https://doi.org/10.1016/j.jas.2015.03.001>
- Touchon, JC and McCoy, MW.** 2016. The mismatch between current statistical practice and doctoral training in ecology. *Ecosphere*, 7(8): e01394. DOI: <https://doi.org/10.1002/ecs2.1394>
- University of Reading.** 2018. *Late Iron Age Calleva: The Pre-Conquest Occupation at Silchester Insula IX [data-set]*. York: Archaeology Data Service. DOI: <https://doi.org/10.5284/1046734>
- Van der Veen, M and Fieller, N.** 1982. Sampling seeds. *Journal of Archaeological Science*, 9(3): 287–298. DOI: [https://doi.org/10.1016/0305-4403\(82\)90024-3](https://doi.org/10.1016/0305-4403(82)90024-3)
- Van der Veen, M, Hill, A and Livarda, A.** 2013. The archaeobotany of Medieval Britain (c AD 450–1500): identifying research priorities for the 21st century. *Medieval Archaeology*, 57: 151–182. DOI: <https://doi.org/10.1179/0076609713Z.00000000018n/a>
- Van der Veen, M, Livarda, A and Hill, A.** 2007. The archaeobotany of Roman Britain: current state and identification of research priorities. *Britannia*, 38: 181–210. DOI: <https://doi.org/10.3815/000000007784016557>
- Van der Veen, M, Livarda, A and Hill, A.** 2008. New plant foods in Roman Britain – dispersal and social access. *Environmental Archaeology*, 13(1): 11–36. DOI: <https://doi.org/10.1179/174963108X279193>
- Vanderwarker, AM, et al.** 2016. New World paleoethnobotany in the new millennium (2000–2013). *Journal of Archaeological Research*, 24: 125–177. DOI: <https://doi.org/10.1007/s10814-015-9089-9>
- van Haaster, H and Brinkkemper, O.** 1995. RADAR, a relational archaeobotanical database for advanced research. *Vegetation History and Archaeobotany*, 4(2): 117–125. DOI: <https://doi.org/10.1007/BF00206920>
- Van Zeist, W, Wasylikowa, K and Behre, K-E.** 1991. Preface. In: Van Zeist, W, Wasylikowa, K and Behre, K-E (eds.), *Progress in Old World Palaeoethnobotany*. Rotterdam: A. A. Balkema. vii–ix.
- Wallace, M, et al.** 2019. Re-analysis of archaeobotanical remains from pre- and early agricultural sites provides no evidence for a narrowing of the wild plant food spectrum during the origins of agriculture in southwest Asia. *Vegetation History and Archaeobotany*, 28(4): 449–463. DOI: <https://doi.org/10.1007/s00334-018-0702-y>
- Warinner, C and d'Alpoim Guedes, J.** 2014. Digitizing the archaeobotanical record. In: Marston, JM, D'Alpoim Guedes, J and Warinner, C (eds.), *Method and Theory in Paleoethnobotany*. Boulder: University Press of Colorado. 147–160. DOI: <https://doi.org/10.5876/9781607323167.c008>
- Whitlock, MC.** 2011. Data archiving in ecology and evolution: best practices. *Trends in Ecology & Evolution*, 26(2): 61–65. DOI: <https://doi.org/10.1016/j.tree.2010.11.006>
- Wilkinson, MD, et al.** 2016. The FAIR guiding principles for scientific data management and stewardship. *Scientific Data*, 3: 160018. DOI: <https://doi.org/10.1038/sdata.2016.18>
- Womack, RP.** 2015. Research data in core journals in biology, chemistry, mathematics, and physics. *PLOS ONE*, 10(12): e0143460. DOI: <https://doi.org/10.1371/journal.pone.0143460>
- Wright, H and Richards, JD.** 2018. Reflections on collaborative archaeology and large-scale online research infrastructures. *Journal of Field Archaeology*, 43(sup1): S60–S67. DOI: <https://doi.org/10.1080/00934690.2018.1511960>

How to cite this article: Lodwick, L. 2019. Sowing the Seeds of Future Research: Data Sharing, Citation and Reuse in Archaeobotany. *Open Quaternary*, 5: 7, pp. 1–15. DOI: <https://doi.org/10.5334/oq.62>

Submitted: 25 March 2019 **Accepted:** 29 May 2019 **Published:** 15 July 2019

Copyright: © 2019 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

]u[*Open Quaternary* is a peer-reviewed open access journal published by Ubiquity Press.

OPEN ACCESS 