ARCHAEOLOGY ON FURLOUGH

ONLINE VOLUNTEER PROJECTS IN THE TIME OF COVID-19

Accessing Archaeological Information Online: A Survey of Volunteers' Experiences

Phoebe Ronn and Rob Wiseman

JUNE 2020

TABLE OF CONTENTS

ABOUT ARCHAEOLOGY ON FURLOUGH

When the spread of coronavirus COVID-19 forced the UK into shutdown in 2020, Archaeology on Furlough was set up to provide volunteer projects for archaeologists unable to work. Around 120 people registered via the website (www.archaeology-on-furlough.com), and most went on to be active participants in the eleven projects on offer. The bulk of participants were professional archaeologists and specialists from commercial units. A number were university students, unable to access laboratories or other research facilities. A few were museum curators or retired archaeologists.

The projects undertaken by the volunteers involved collecting and analysing data on:
- Roman cultivation strips in the East of England
- gravegoods in the Iron Age, Roman and Anglo-Saxon periods (Cambridgeshire and Oxfordshire)
- trauma in skeletons from the in the Iron Age, Roman and Anglo-Saxon periods (Cambridgeshire and Oxfordshire)
- aurochs remains from Scotland, England and Wales
- Saxon houses from Scotland, Northern England and the Midlands
- henges excavated in Scotland, Northern England and the Midlands
- historic sheepfolds of the Lammermuirs Hills (Scotland)
- barrows and other Bronze Age sites on Dartmoor
- temples in Roman Britain
- decoration used in prehistoric Britain
- digital tablets for use in archaeological fieldwork.

Once completed, the grey literature project reports and datasets generated by the project teams will be made available via Cambridge University Library's Apollo repository (https://www.repository.cam.ac.uk/).

Projects were all conducted voluntarily, so that participants could comply with the requirements of the UK Government's Coronavirus Job Retention Scheme.

For many of the volunteers, Archaeology on Furlough was an opportunity to research unfamiliar topics, and discuss them with peers. For some, it was an opportunity to develop new skills, particularly research skills which they did not get the opportunity to use in their regular fieldwork. For a few, Archaeology on Furlough provided their first opportunity to write a report.

Archaeology on Furlough was also intended to provide support for archaeologists out of their normal workplaces. Volunteers used video conferencing and various message boards to keep in touch with one another.

Archaeology on Furlough was developed and coordinated by Rob Wiseman (Cambridge Archaeological Unit, Department of Archaeology, University of Cambridge).

SUMMARY

Archaeology on Furlough (AoF) involved volunteers gathering together and analysing large volumes of archaeological data on a wide range of archaeological themes. Inadvertently, it became a test for how easy it is for archaeologists to obtain information they need. To understand more, AoF volunteers were surveyed about their experiences when locating, obtaining and using archaeological information.

94 AoF volunteers were invited to participate in an online survey: 48% participated.

Respondents' most highly rated ways of locating information were (in order):
- Google and other search engines
- references in archaeological publications
- search facilities in the ADS and commercial units online libraries
- gazetteers and dissertations.

Most valued features for respondents were search engines that allowed searches by keywords or other filters (by period and location when searching for sites; by author and publication year when searching for articles). The main barriers to identifying information resources which respondents noted were:
- poor, incomplete or wrong references
- a lack of tags in resources so that search engines could find resources online
- publication in obscure or inaccessible journals
- inconsistencies between individual HERs as well as other resources (particularly in their terminology or classifications).

The most highly-rated information sources were those which could be downloaded for free from online portals—notably the ADS and online county journals, along with free online resources like Google Books and the Internet Archive. Printed materials were consistently ranked difficult or impossible to access (hardly a surprise given that AoF was conducted in lockdown, but also a problem for commercial archaeologists).

Overwhelmingly the most positive attribute of information sought by respondents was the ability to download it. A much-valued feature was a direct link to online resources—whether from search engines, hyperlinks, or in bibliographies (via DOI). The largest barriers to obtaining archaeological information were resources that were only available in print (not digitally online), and resources behind paywalls.

By far the most useful format for the AoF volunteers were PDFs, whether of grey literature, articles or monographs. This is because they could be searched by machine, and it was also possible to copy text and tables into other digital formats. The least useable resources for respondents were printed articles and monographs.

Respondents' comments point to clear problems with the formatting of many grey literature reports, which makes searching and using them difficult. Documents which were structured inconsistently or idiosyncratically also rated poorly (a problem for monographs in particular). Incomplete, overly-summarised or cherry-picked information could also be a problem, particularly in specialists' reports. Organisations that 'locked' their PDFs were criticised.

Respondents' experiences of obtaining archaeological information overlaps with the challenges faced by others in the sector—particularly commercial archaeologists. This report does not pretend to be a whole-sector review of access to archaeological resources, and acknowledges there are many legitimate reasons why some resources may not be made freely available online. This report concludes with some suggestions which would address respondents' comments. These suggestions do, however, have to work within the political and financial constraints in which archaeological information is produced.

BACKGROUND TO THE SURVEY

Archaeology on Furlough (AoF) was developed in mid-April 2020 for professional archaeologists in the UK who were unable to work because of the coronavirus lockdown. The programme consisted of twelve volunteer projects. 120 people registered via the project website, and the bulk went on to be active volunteers.

Most of the AoF projects involved collating information from excavation reports in order to prepare gazetteers and datasets for analysis. The topics covered were:
- Roman cultivation strips (East of England and East Midlands)
- Grave goods and skeletal trauma for the Iron Age, Roman and Anglo-Saxon periods (Cambridgeshire and Oxfordshire—four teams in total)
- Aurochs remains (England, Scotland and Wales)
- Saxon houses (chiefly northern England and southern Scotland)
- Excavated henges (England, Scotland and Wales)
- Roman-era temples
- Decoration on Neolithic and Bronze Age objects (England, Scotland and Wales).

In the space of six weeks, the volunteers pulled together a huge volume of information, and also sifted through a great deal of other material. The number of records generated gives an impression of the volunteers' industriousness: contents of over 2,500 graves; over 250 excavated Saxon houses; around 110 excavated henges; findspots for c.600 aurochs skeletons or bones; and nearly 400 pieces of prehistoric decoration.

For the most part, the volunteers had to rely on the resources they could access from home for free. A few volunteers had access to university libraries and were able to able to access journal articles behind paywalls. Some volunteers who were members of archaeological units could also access their company's grey literature collections. But the bulk the of material that volunteers used came from open access online sources such as the Archaeology Data Service (ADS), free online journals and monographs, commercial units' online grey literature libraries, Historic Environment Records (HERs), Research Gate, and academia.edu. A few volunteers reached out directly to authors, HERs, and commercial units to access material. Some people also shared scans and photographs of books they had in their personal collections.

It was quickly realised that the problems that volunteers had accessing information matched some of the difficulties faced by commercial archaeological units, which likewise have to rely primarily on open access publications and grey literature. Inadvertently, AoF had set up a large experiment in the challenges of gathering archaeological data. It was decided therefore to poll all of the active volunteers about their experiences, in order to find better ways for the UK's archaeological community to share the fruits of its research.

SURVEY STRUCTURE

The survey was conducted online via Survey Monkey (www.surverymokey.co.uk). Its free version allowed for 10 questions. The first of these was used to ask which project the respondent had participated in; the remaining nine questions were used to explore the stages of information searching and use. To do this, the survey was divided into three parts:

1. how respondents identified the resources they needed for their project (3 questions)
2. how respondents obtained those resources (3 questions)
3. how respondents used those resources to get the information they needed (3 questions).

In each set of questions, respondents were asked to rate a series of options on a five-point scale, then provide free text answers on "the two or three most useful aspects" and the "two or three biggest barriers". We specifically separated the positive and negative aspects, as we wanted to know not just what was wrong, but also what the archaeological community could be doing more of.

In essence, the first question in each set identified *what* was working well or badly, and the remaining two *how* and *why* options worked well or badly.

The survey questionnaire is in the Appendix.

ENGAGING VOLUNTEERS

The survey was only open to people who participated in AoF projects. The Survey Monkey form was open from Tuesday 26 May to 8pm Sunday 31 May.

On 26 May, emails were sent directly to the 94 people understood to be active in AoF, and a post placed on the AoF Facebook page. Only the three volunteers in Project 10 (Review of Digital Tablets for Site Work) were not emailed, as the project had not involved gathering archaeological reports. Two projects using primarily lidar data and aerial photographs (Dartmoor, Lammermuirs) were included in the mailout, in case the teams had had supplemented their analysis by searching for written reports.

A follow-up email was sent on Saturday 30 May, along with a post on the AoF Facebook page, to remind volunteers that the survey would close in 24 hours. A final reminder post was placed to the AoF Facebook page on Sunday morning.

QUESTION 1: NUMBER OF RESPONSES

In total, 45 people responded, from the 94 who were emailed (48%). Given that surveys of the general population typically elicit a 10–15% response rate, we judge this to be a very good response rate. Figure 1 shows the number of people emailed and the number who responded to the survey (Question 1).
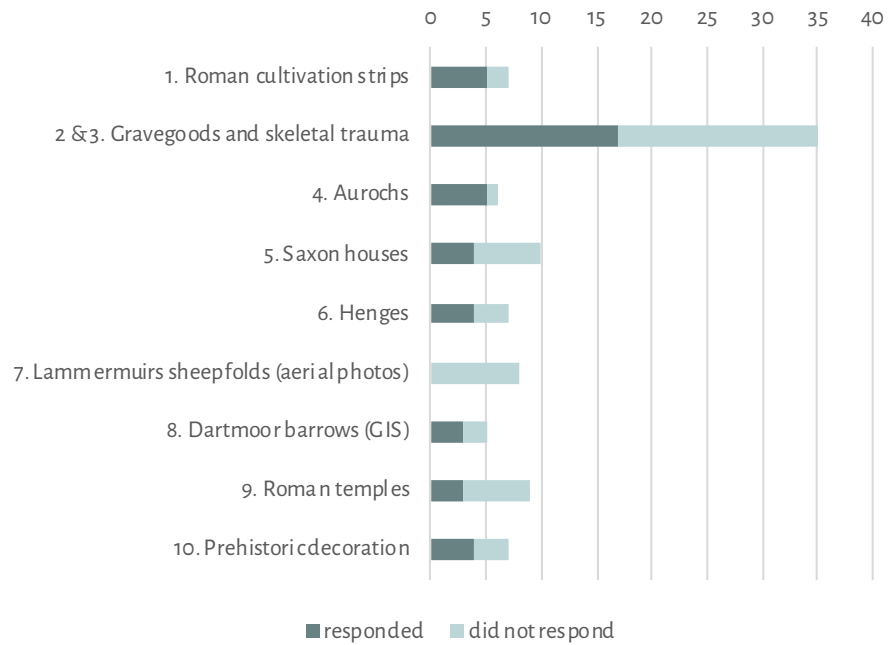

Figure 1: Number of responses and non-responses from each project group.

Note that the lack of responses from the Lammermuirs group is due chiefly to the nature of the project, which was focussed on scanning aerial photographs and old maps, rather than accessing excavation reports.

.

QUESTIONS 2–4: IDENTIFYING RELEVANT INFORMATION RESOURCES

Any search for archaeological information starts with people identifying potential places where they might find the information they are seeking. The first question in this section asked, "how easy it was for you to identify the specific sources of information you needed for your project?" Respondents were asked to consider eleven of the main search routes used during AoF, and rate them on a five-point scale. Results are shown in Figure 2. (For clarity, in all of the graphs that follow, the number of respondents who selected 'not applicable' has been excluded. The two lowest ranked scores have been presented to the right of the main axis, the other three to the right, in order to highlight the use and usefulness of each.)
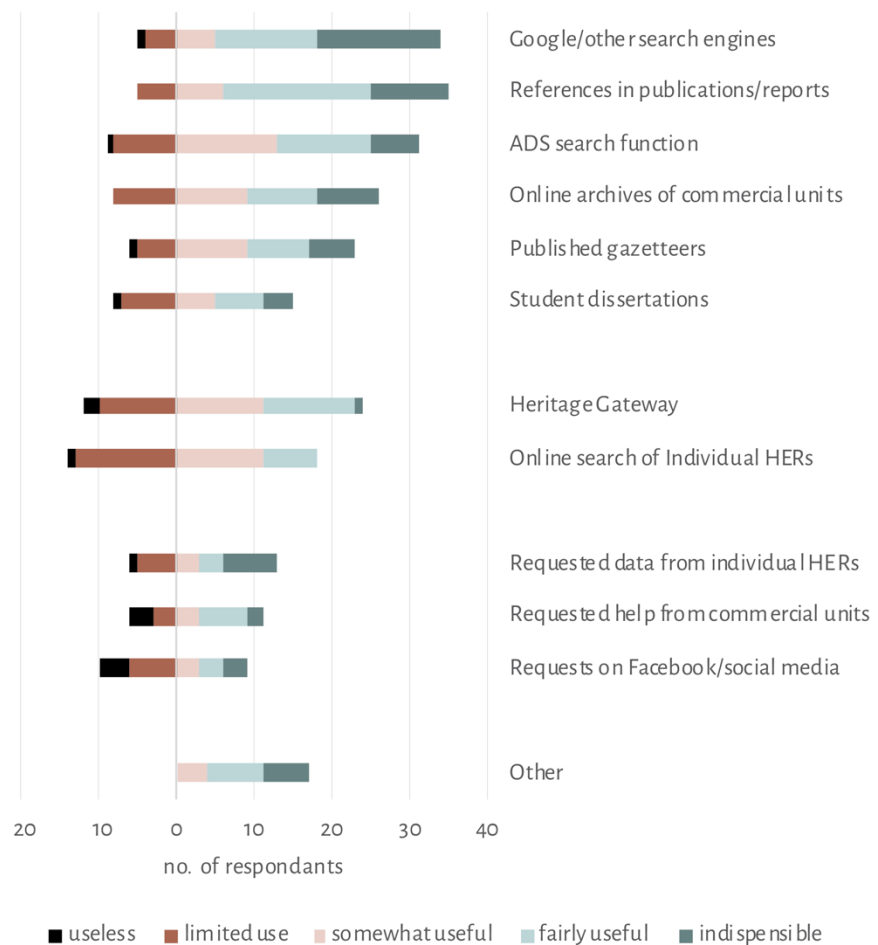
Figure 2: Ease of eleven different methods for identifying potentially useful information resources

As Figure 1 indicates, the ease of identifying sources could be divided into three distinct groups.

First were those widely used, with a high proportion of respondents describing them as 'indispensable' or 'fairly useful'. The most useful resource was not archaeological at all but simply 'Google and other search engines'. No other source could rival online search engines for being indispensable tools for archaeological research. Search facilities in the ADS and commercial units were also commonly used, but rated rather lower for their usefulness—and behind references in published papers. Gazetteers and dissertations were less commonly used or described as 'indispensable'—in part because there were

no relevant resources of this type for many of the AoF projects. However, where they were identified, they could provide a rich source of other references for further searches.

The second group, used by less than half of respondents, and only rarely rated 'indispensable', were 'accumulator' sources, such as Heritage Gateway and HERs. Unlike resources in the first category, archaeological 'accumulator' sites are unavoidably less targeted in their contents.

Only a third of respondents reported making direct requests for information to individual HERs, commercial units or the online community. While HERs were generally well-regarded, commercial units and the general archaeological community were rated rather lower, with a concerning number of respondents regarding them as 'useless'. (In fairness, this may in part be due to units being in lockdown, and unable to respond to requests, and so these figures may not reflect their usefulness in normal times.)

Other resources that respondents described as 'indispensable' or 'fairly useful' were:
- other archaeological databases, such as Archwilio, Canmore, and Coflein
- portals for academic articles such as academia.edu and ResearchGate
- library searches
- online databases (specifically the Biodiversity Library, Past Peoples of Oxford database, and the England's Rock Art and Scotland's Rock Art databases)
- individuals' own personal libraries.

Useful attributes of resources when searching for potential information sources

The most valued resources for locating information (66% of respondents) were search engines which allowed searches by keywords or other filters, in order to narrow down relevant resources. The favoured filters for searching for sites were 'period' and 'location' whereas the preferred filtering options for searching for specific articles were 'author' and 'publishing year'. (Amongst the county journals, special mention needs to go to the website of the Oxfordshire county journal *Oxoniensia*, which was single out for its excellent search facilities, compared with most other similar journals.)

Over a quarter of respondents nominated sources which included direct links to resources they were searching for.

Impediments to searching for potential information sources

Some respondents complained that information that appeared key to their research was published only in "obscure" journals that were frequently not searchable, had no online presence, or did not make back issues available. Being limited to online searches made these sources entirely inaccessible.

Several respondents noted that some HERs were much easier to search than others. This will doubtless be a growing problem, as increasing numbers of HERs establish their own portals, with their own idiosyncratic search methods and data output. Several respondents also noted that HER terminology or classifications were not consistent.

> HER data [is] not compatible from county to county. Some descriptions [are] too vague to be useful, lack of geographic detail, [or] lack of reference to the data sources.

Another commonly reported problem was poor, incomplete, or wrong references, which made hunting for some resources impossible. One respondent complained that "mentions of tantalizing sites were often made without the information to seek them out", which made identifying potentially useful resources very difficult. Missing information—like author, date or publication title—in references can render resources unlocatable. Many HERs, for example, include article or chapter titles in their references, but not the journal or publication they were published in. Failure to note

follow-up excavations was another criticism of some HERs, and led respondents to doubt the completeness of records.

Inconsistency between sources was also an issue. One person reported "significant discrepancies between HERs/Heritage Gateway and PastScape records". A similar problem was inconsistent terminology used across resources with common themes. For example, participants of the Aurochs project found that, in order to retrieve all relevant resources on their topic, they had to search "auroch, aurochs, bos, bos primigenius, wild ox, bison". One member of the Roman Cultivation Strip project reported difficulty finding these features, as there is no standard term for them. Another respondent was frustrated that they had to manually trawl through documents to check if they were relevant to their project, and suggested:

> ...allowing 'tags' to be applied when people upload their report or document. By tagging them—like a hashtag on twitter or Instagram—it will allow the person uploading to choose key words or key topics of note within their report that people can then search for.

The lack of tags in search engines—even very common terms—made locating resources difficult or impossible to locate. One respondent said, "I put in 'Roman' and 'Cambridgeshire' into ADS search engine and it came up with [just] 6 entries".

The British and Irish Archaeological Bibliography (BIAB) housed in the ADS was singled out repeatedly for a lack of links to original source material. (The ADS website reports it has over 300,000 bibliographic references and over 100,000 open access documents.) Another problem mentioned in relation to the ADS was that some forms of its bibliographical information consisted solely of long lists of titles—sometimes numbering in the hundreds—in a format which was not practical to search.

QUESTIONS 5–7: ACCESSING THE INFORMATION RESOURCES IDENTIFIED

Having identified a potential source of information, the second stage in an information search is to obtain the resource. The second group of questions asked respondents to rate how easy is was to obtain the specific resources they wanted from fifteen sources. As for Q2, a five-point scale was used, with an option for 'not applicable'. The results are shown in Figure 3.
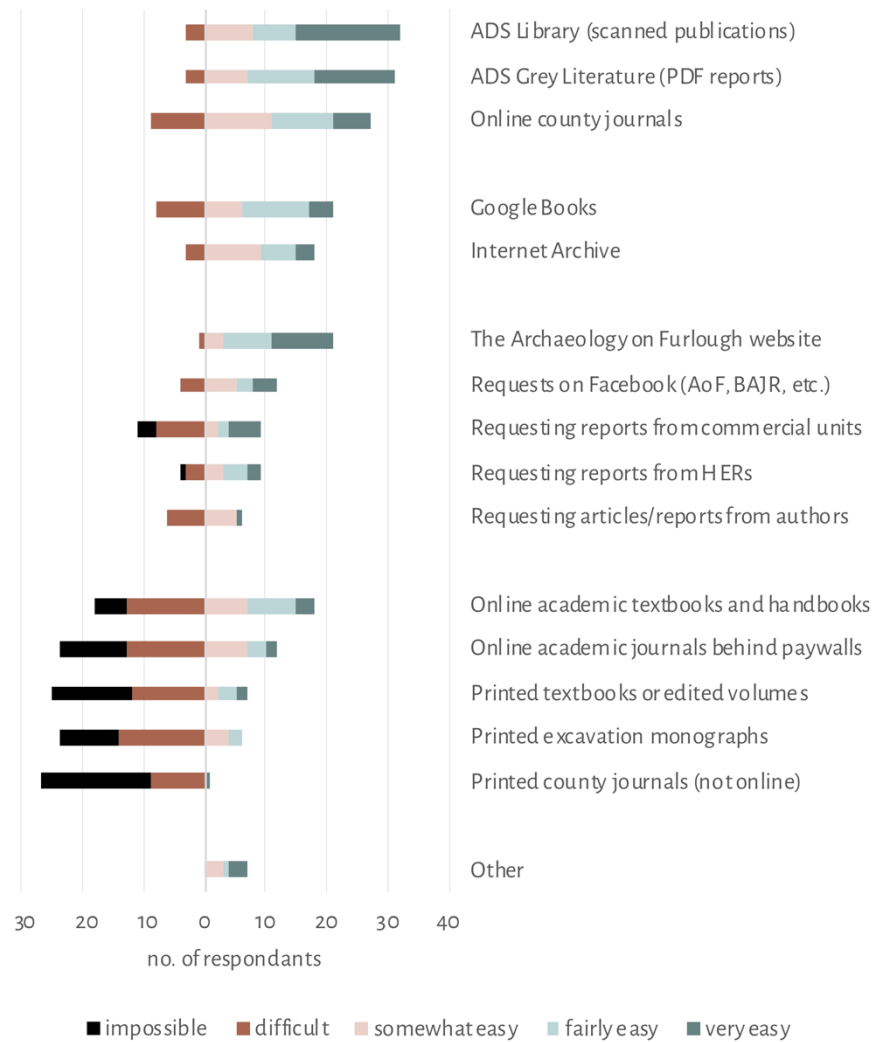

Figure 3: Ease of accessing archaeological information from different sources

By far the most highly-ranked resources for respondents (more than 30) were those which could be downloaded for free from online portals—most notably the ADS but also online county journals)—followed by free online resources like Google Books and the Internet Archive. Printed materials were consistently ranked difficult or impossible to access—hardly a surprise given that AoF was conducted during lockdown. Articles behind paywalls too were overwhelmingly described as difficult or impossible to access.

Around half of respondents had tried requesting materials from authors, commercial units, HERs, and the online community. A fair proportion of these experiences were described as difficult and unsatisfactory. The wide range of ratings given to commercial units—from 'very easy' to 'impossible'—points to the highly variable responsiveness of the sector, although in fairness this may partly be a by-product of lockdown.

Useful features when accessing resources

Overwhelmingly the most positive feature of the information sought by respondents was the ability to download it. One person described it as "probably the most useful feature of a publication". For this reason, the ADS—with over 100,000 open access documents in its library, including 60,000 grey literature reports—and was the most highly ranked source of information amongst respondents. At least two-thirds of English county journals have now been digitised and can be downloaded for free, as can key journals in Scotland and Wales (*Archaeologia Cambrensis*, *Discovery and Excavation in Scotland*, and *Proceedings of the Society of Antiquaries of Scotland*). Although the online archives of commercial units were not on the list of options in this question, they also rated highly in the previous set of questions, and comments from respondents indicated that it was the ability to download their grey literature reports which was especially important.

Another particularly valued feature of search engines was a direct link to publications. One respondent highlighted the usefulness of bibliographic references which contained a digital object identifier (DOI) because it made locating the actual resource easier.

One other highly-rated resource which was used by about half of respondents was the AoF website. This contained a long list of archaeological resources which could be accessed for free. This included all of the open access county journals, along with major academic journals, links to commercial units with online libraries, free textbooks, the ADS library, the database of the Roman Rural Settlement Project, as well as mapping resources such as LiDAR, geology and soils data.

Barriers to accessing resources

Two barriers stood out ahead of all others: price and print—problems for over seventy percent of respondents.

Apart from a few AoF volunteers who had access to university libraries, most participants had no direct access to articles behind paywalls. Shortly after AoF started, participants started emailing friends or posting on Facebook pages to request help obtaining paywall-protected articles. But even despite a network of 100 volunteers, some of them with access to the UK's largest academic libraries, many articles could not be obtained. (This a question about who actually *does* have access to these resources.)

Actual availability of publications was also a distinct issue, as many respondents spent time searching for non-existent digital copies of resources they had seen referenced online. Printed volumes were inaccessible for most participants—despite a large traffic in scanned and photographed texts within the AoF project teams.

Despite the number of articles which have been uploaded into the ADS, respondents noted that many more had not been. And people reported examples where reports which had been uploaded were incomplete—missing appendices or separate tables or plans.

QUESTIONS 8–10: USING THE RESOURCES OBTAINED

Having obtained a resource, the final stage in an information search is to access the relevant information. The final set of questions in the survey asked respondents to rate eight of the most common forms of archaeological output. As in the previous two sections, a five-point scale was used, with an option for 'not applicable'. The results are shown in Figure 4.
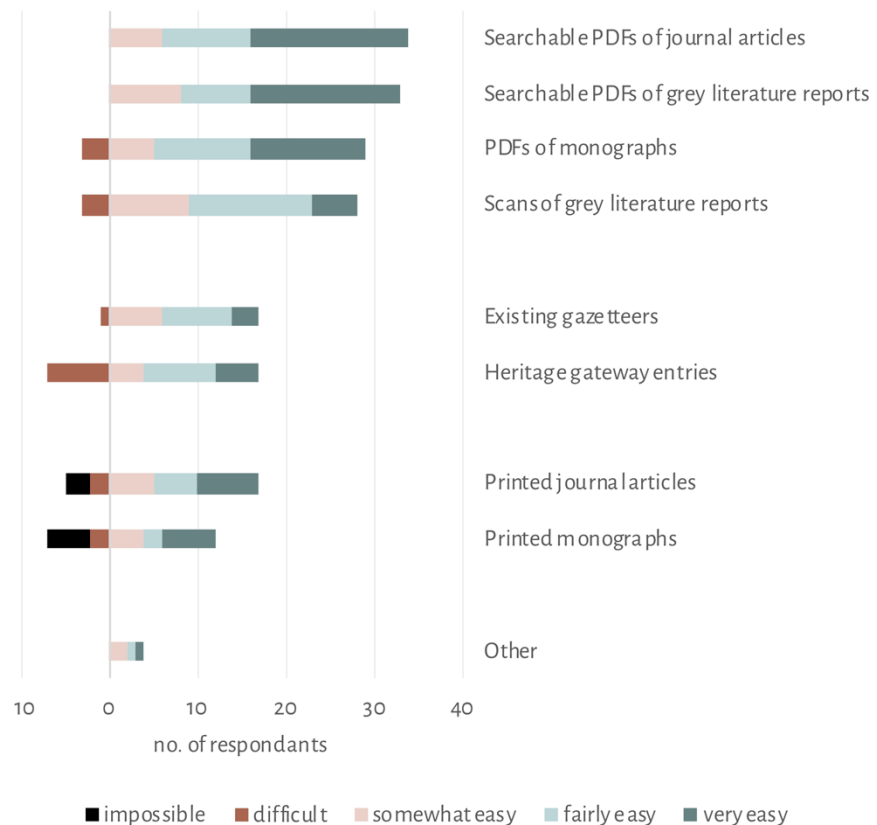

Figure 4: Ease of using various information sources

The ease of using sources to obtain specific information can be divided into three groups, which largely reflect the issues raised in the previous two sections. By far the most useful format for respondents were PDFs of articles, grey literature and monographs. Scanned or photographed text was also generally easy to use, although not as much as machine searchable document. In the second group were sources which had previously compiled information respondents wanted—resources like gazetteers and data in HERs/Heritage Gateway. Although some respondents rated Heritage Gateway as 'very easy' to use, rather more rated it the most difficult to use of the eight sources. The least useable resources were printed articles and monographs.

The ability to search documents was an important factor in ease of use. So unsearchable scans or photographs of journal grey literature were judged more difficult/less easy to use than searchable PDFs of grey literature.

Factors making resources easy to use

Formatting was the most commonly given factor which made resources easy to use. Formatting issues raised included:

- a clear structure
- figures and tables that were easy to understand
- the ability to extract information efficiently.

> Formatting and presentation with easy headings ... meant I could jump straight to that section.

> Clear and well-labelled figures. Clear easy to understand tables.

> Some commercial units had [a] definite consistency in their formatting which meant location of information was good.

> Information was legible and resources were complete. Site plans generally always included.

> Clear delineation between the presentation of data and then the discussion/ interpretation. Clearly presented plans and sections to compare with the descriptions in the text.

Respondents noted in particular the consistent structures used in journals and commercial unit publications, as this made locating relevant information simple. It is noticeable that monographs were rated as less useful than both articles or grey literature (whether in PDF or print). This appears to be chiefly because monographs either omitted information or were arranged idiosyncratically, with information scattered in various places across the text.

> ... most of the commercial archaeology reports ... follow the same standard template so it is easy to follow once you get used to reading them. They are complete and include the appendices, figures and tables within the whole report so it is easy to know what they are referencing.

The ability to machine search a resource for keywords or figures was mentioned by 17 respondents as a valuable feature, as it made identifying relevant data quick and easy. As Figure 4 indicates, scans of documents or print-only versions which had to be read rather than searched by machine were rated as less useful.

As most projects involved collating data from multiple resources into a database, the ability to copy-and-paste information—rather than manually transcribe it—was also rated a particularly useful feature.

Factors making resources hard to use

The themes identified in the negative comments mirrored those in the positive comments section.

Respondents disliked scans or hard copies, compared with machine-readable PDFs, as they were unable to search or copy the information. Formatting of reports was also a common problem:

> Inconsistent formats
> Lack of clearly defined sections or index/contents at beginning [of reports]
> Lack of clear headings

> No standard reporting method across commercial units or in some cases between different parts of the same unit.

> Sometimes no structure [to reports], or not all of the details were provided.

Several respondents noted problems with basic editing and content of grey literature reports. Omissions noted included:

- site location
- north arrows on plans
- lack of scales in inaccurate scales
- lists of context numbers or excavated features
- bibliographies
- incomplete legends or categories that were impossible distinguish visually.

Data not always recorded consistently in grey literature reports, e.g. not always having context tabulation in appendices, not all measurements fully recorded, pottery/environmental reports not always tabulated by individual contexts.

Reports across units, and even within units, are inconsistent in their approach to site reports. Data like measurements of features were missing from reports. Some reports were poorly edited with missing information. Figures [that were] unclear with no titles so I didn't know what they were showing (e.g. one report had figures with no feature or context numbers on [them], so [I] didn't know what feature was what, [and I] couldn't link it to text description). Finds and enviro data was often very separate from the main text and sometimes not even referred to in results section, so [I] had to trawl though the report to find relevant information.

Incomplete, overly-summarised or cherry-picked information could also be a problem, particularly in specialists' reports of pottery, bone and environmental samples.

A few respondents noted that some reports lacked any kind of interpretation.

Lack of mid-level analysis. More recent reports often have a general overview, and a list of raw data, but skip steps [between].... Also, newer reports are often more circumspect, which ... can amount to leaving out this middle range of interpretation.

Some respondents disliked information being reported in text, when it could more usefully have been presented in a table.

As most of the projects in AoF involved collecting data, respondents disliked PDFs where the text had been 'locked' so they had to manually transcribe text. Tables also made copying or transcribing data much easier than presenting it as text.

REFLECTIONS ON RESPONDENTS' INPUT

Despite the wide variety of archaeological topics that the AoF volunteers were exploring, their responses to the survey—particularly in their comments on what made their tasks easy or difficult—were strikingly consistent. They point to a relatively small number of factors which would make archaeological information more accessible and useful, particularly to commercial archaeology in the UK.

Caveats on the applicability of the survey's findings

Before exploring the consequences of respondent's comments, it is appropriate to note the similarities and differences between the AoF volunteers on one hand and archaeologists in commercial units and universities on the other.

First, the AoF respondents worked almost exclusively on projects which involved searching for primary archaeological data—site plans, excavated features, pottery analyses, bone reports, pollen samples, tabulated data, and so forth. They were not generally searching for archaeological theories, interpretations or methods. Nor were they seeking overall syntheses of data (such as the regional research frameworks), other than as sources for their own data gathering.

Second, because the projects were conducted during lockdown and most volunteers were on furlough, they only had access to materials freely available online, plus whatever they could obtain via their social networks. Most of the volunteers did not have their own personal libraries. Most had no access to printed resources while, by contrast, even the smallest units usually have copies of local journals and some monographs, and academics have access to much larger university collections. Only a handful of volunteers had access to paywall-protected articles and books—which in fairness is fairly typical of commercial units in the UK (only eight units appear to have university affiliations: Cambridge, Durham, Leicester, Orkney/UHI, Salford, Staffordshire, UCL/Archaeology South-East, Winchester).

Finally, the AoF volunteers did not have to work within some of the constraints of commercial archaeology: notably budget or time constraints; the need to conform to planning requirements or CIfA standards; or, in the case of university-based units and academics, the need to generate work which would contribute to the Research Excellence Framework (REF).

Notwithstanding all these differences, the experience of AoF volunteers reported in this survey does have relevance for how archaeologists—particularly those in the commercial sector—search for, obtain, and use archaeological information.

Archaeological publishing and accessibility of material

This survey threw into relief some of the ironies of archaeological publishing in the UK. For example, although IT has made significant inroads into archaeological recording, analysis and reporting, the 'final publication' of information routinely sees digital information turned back into print, rather than into a digitally accessible form. Since most print journals and monographs rarely run to more than a few hundred copies— most of which go to university libraries—this really does limit access, particularly for commercial archaeologists.

An even greater irony is the cost of publications. Although commercial units produce the majority of site reports[1], unlike university departments and libraries, they generally lack the resources to acquire more than a tiny fraction of the type of publications they

---

[1]   Between 1990 and 2010, the Archaeological Investigations Project reported that university excavations accounted for 514 records (2.2%) out of a total of 22,800 field investigations (Darvill *et al.* 2018).

themselves produce. While commercial units do try to recover some of the publishing costs via sales, the amounts recouped on the sale of a few hundred monographs are only a few percent of sums units charge their clients for excavations.

While publishing in academic journals performs an important function for academics and postgraduate students, these journals are essentially out of reach for most individual archaeologists and commercial units.

> I wouldn't mind paying for an article if it was reasonable (up to £5 maybe?), but I can't afford to spend £45 for an article on the off-chance that it may have one or two lines that are relevant to my research.

In the face of these costs, archaeologists reach out to colleagues who might have copies they can share. It is hardly a secret that both commercial and academic archaeology run on a large *samizdat* literature—a problem that is largely self-generated. The way that AoF volunteers turned to social media and professional networks is not unlike what commercial archaeologists, specialist and students have to do to obtain information they need (albeit, not as publicly).

Grey literature and monographs

Respondents' comments clearly identified problems with the way some grey literature is structured and formatted. The main problems reported—incomplete or missing data; poorly structured and formatted information; and substandard graphics—were reported by too many people for these to be occasional lapses. There is plainly a wider problem with the organisation and presentation of grey literature by commercial units.

Commercial units which 'locked' their PDFs, so it was impossible to copy the text, were also a bugbear. When respondents were faced with this issue, they merely transcribed the text; the information was replicated despite the 'locking' but this replication was time-consuming, rendering this move pointlessly obstructive.

It is, however, a sign of the problems with archaeological publishing that AoF volunteers rated grey literature *more* useful than monographs—even fully digitised and searchable monographs. The main problems that respondents highlighted with monographs, compared with grey literature and articles were:

- omission of basic data—particularly when it was presented in the site's grey literature (For example, tabulations of skeletal information were often excluded in monographs, which made this media useless to volunteers on projects that involved the collection of grave data.)
- idiosyncratic organisation, which can see the same topic addressed in several locations within the one publication
- over-summarisation of arguments, so it is unclear how key interpretations, calculations and deductions were made (also a problem with some grey literature).

Archaeological reports of all kinds—grey literature, articles, monographs—have to balance to opposing tendencies. The first involves presenting data in detail so that it can be re-used by others; the second involves summarising and synthesising data so that readers can understand key issues. This is a long-standing tension. (Thomas 2004 traced it back to a tension between Pitt-River's emphasis on complete objective presentation of data, and Petrie's emphasis on brief synthesis.) Given that the AoF volunteers were chiefly seeking primary data, it is unsurprising that they had a clear preference for reports which contained complete data, and were critical of over-summarisation.

A second problem with reports, particularly with large monographs, is that they are an 'old technology', and simply cannot do what new digital technology can. Many other products of print technology have been rendered obsolete in the last thirty years—Britannica has been replaced by Wikipedia; the Yellow Pages by Google; the road atlas

by Garmin. Types of functions that respondents wanted to be able to do with reports included:

- word searches
- copying data and pasting it spreadsheets, or copying text into a word processor
- following links to the bibliography or references
- find related material within texts (particularly long monographs).

> [A barrier to using information is] not being able to use 'Ctrl + F' especially if it is a big report and there is no prior knowledge to the report containing what you are looking for. If you are working on a hunch it can mean that looking through the report can be a long process. It may be useful also if, when a report references appendices or figures, they [have] a clickable link within online reports so you can immediately jump to what they are talking about without having to search for it and losing your place.

> The best resources had been at least run through OCR so that they could be searched. All of the reports I have used were complete and legible, but some were not OCR'd and the original report had no page numbers, index, and sometimes unclear organization so I had to flip through every page scanning for the information I was looking for.

Portals and online repositories

Online portals were highly valued by respondents—the ADS in particular was praised by many AoF volunteers, as were commercial units who provided their grey literature reports online. Respondents did point to some limits and problems in using the ADS which, in fairness, are largely out of the ADS's control:

- missing data—because units which do not upload their reports, or provide links to their own online libraries, or Local Planning Authorities have not released the report
- inconsistent or poor tagging of data by units, notably in the OASIS forms from which ADS search terms are derived (although the authors are aware that the ADS is providing training to units in OASIS in 2020)

The volume of material available via the ADS—over 100,000 free documents—creates the challenge for searching and locating information. Many respondents noted the usefulness of the ADS's filtering options (e.g. to filter to geographic areas, archaeological periods). One area that was more of a problem was searching for a specific title. A number of respondents mentioned using Google to search the ADS in preference to the ADS's search tools for that function:

> Googling ... sometimes provided links through to ADS / unit websites or papers which referenced ... articles, but searching through ADS itself often returned nothing.

> ... Google actually found the report in ADS for you, rather than trying to search in ADS and finding nothing.

Historic Environment Records were also an important resource for respondents. As noted above, some challenges people reported were:

- discrepancies between individuals HERs, PastScape and Heritage Gateway
- incompleteness of some HER references (in particular, while article titles were usually supplied, the journal titles or publication names were not)
- the very different local portals offered by individual HERs, with differing options for searching the recording, and the widely varying types of information that could be obtained from them online

SUGGESTIONS FOR IMPROVING ARCHAEOLOGICAL INFORMATION

The purpose of this survey was to gather the experiences of AoF volunteers. We make the following suggestions, based on the comments from respondents, in the hope they might be useful to the wider archaeological community. We freely acknowledge that there may be many perfectly legitimate reasons why they may not be feasible or desirable in particular circumstances—the survey was not intended as a wide-ranging review of archaeological information production, dissemination, access or use. There are many other factors affecting the use of archaeological information beyond those raised in the survey.

Improving the ability to identify resources

To make archaeological information easier to find, producers could:

- place all of their reports and publications in online portals, and encourage gatekeepers like planning archaeologists to clear reports for prompt public release
- include keywords or 'tags' in resources so search engines are more likely to pick them up during searches—more consistent use of OASIS forms could also help
- assigned a DOI (digital object identifier) to all grey literature and articles, so they can be found online.

Bibliographies in all forms of grey literature and formal publications could include as standard DOIs or hyperlinks to online information. Hyperlinks in HERs, the BIAB, and the ADS would also help users locate information they were seeking.

The high ratings given to the Archaeology on Furlough website—which included a large number of links to sites where archaeological information could be downloaded for free—suggests there is a need for a 'signpost' website with links to free online resources. (The BAJR website does contain some of these resources, but we did not investigate its use as part of this survey.)

Improving the ability to obtain archaeological resources

When archaeological resources are uploaded to online portals, they need to be complete. Producers should ensure that all parts of grey literature reports—including plans, bibliographies, appendices, datasheets, and scanned microfiche—are uploaded to online portals, not just the text.

The authors recognise that, while grey literature can be made readily accessible quite simply, formal publications are another issue. The advice of our respondents, we suspect, would be to put articles and monographs online in the same way as grey literature. However, the authors of this report appreciate that there are perfectly legitimate reasons that not all publications can be made open access immediately on release—largely because publishing is in the hands of third parties, not archaeologists themselves. For example, most county journals—which carry the bulk of published articles—are produced by local archaeology or history groups, who charge a journal subscription. They do this, in part, to help secure members to their organisations, because these groups need members to remain operational[2].

Monographs involve more challenges than articles. More than any other form of information produced by archaeologists, major monographs are *political* documents. Their bear the reputations of their authors. They have to respond to the expectations of those who funded them, and work within the constraints of that funding. Monographs arising from commercial projects also need to meet planning conditions and the

---

[2]   Another legitimate challenge for these local resource-poor organisations is the cost and effort involved in digitising their back catalogues. This involves, minimally, scanning every back issue, then preparing machine-readable versions, building a web portal to house them, and the designing a good search engine.

expectations of local government planning archaeologists. Monographs need to reflect the input of all the specialists who contributed to them. They also need to be contextualised geographically and also in the history of previous excavations and thinkers. With some many actors and demands, it can be easy to discount the way that readers want to use the information and the tasks they might want to carry out with it. Making monographs more accessible and useable for readers is a highly political act.

Faced with these constraints then, we suggest the following which might help address issues with accessibility raised by respondents:

- In the best of all worlds, archaeologists publishing excavations would obtain sufficient funding from their clients to publish in Open Access (i.e. involving no cost to the end user). It is entirely feasible to publish both articles and monographs online as Open Access, in tandem with printed copies as well—a number of commercial publishing houses have already adopted this model. As noted above, the cost publishing this way would only be a few percent of what commercial units charge their clients for a major excavation. Arguably, for commercial projects done within the planning system, Open Access publication might also better meet the goal of 'recording and advancing understanding of heritage assets' in the National Planning Policy Framework[3].
- Not all publishers offer Open Access publishing—county journals for example. However, a model being using increasingly is to protect the content of such journals for 1–3 years (usually by providing it as subscriber-only hard-copy), before placing digital versions online for free access.

Improving the usability of information

The formatting of grey literature could be improved by:

- clearer headings and other navigational devices to help readers find information they are looking for within reports
- presenting more data in tables, rather than embedded in text, so readers can readily identify and extract it
- providing grey literature as machine-searchable PDFs (or other searchable formats)
- including more links within related sections of documents (i.e. treating reports primarily as computer-based documents, rather than print-based, and using digital technology to better effect)
- not locking PDFs, so readers can extract data easily for further analysis (as discussed earlier, 'locked' text can easily be circumnavigated by researchers, but at the cost of time and effort).

As discussed above, improving the usability of published reports, particularly monographs, is a more complex issue. As well as transforming an old technology into a new one, change also has to address the political environment in which these documents have to work. There has been discussion on this for over twenty years, with notable discussions in:

- *Internet Archaeology*, Issue 6 (1999). Digital Publication [https://intarch.ac.uk/journal/issue6/index.html]
- *Internet Archaeology*, Issue 15 (2004). Archaeological Infomatics—beyond technology [https://intarch.ac.uk/journal/issue15/index.html]
- *Journal of Field Archaeology* 43, Supplement 1 (2018) [https://www.tandfonline.com/toc/yjfa20/43/sup1]
- Historic England and the Chartered Institute for Archaeologists (2018) *21st century challenges for archaeology*. Workshop 6: Challenges for archaeological publication in a digital age: who are we writing this stuff for anyway? [https://www.archaeologists.net/21st-century-challenges-archaeology]

---

3   Paragraph 199, National Planning Policy Framework

Plainly these issues are far beyond the scope of this survey. We will limit our final suggestions, therefore, to a few suggestions for the organisation and presentation of digital monographs, which could address the main issues raised by respondents.

- Produce digital reports in a form which is machine readable, so readers can search for text easily
- Produce digital reports in a form which allows text and figures to be readily copied to other documents (text to word processors, tables to spreadsheets, vector-based figures to illustration software)
- Use internal hyperlinks to help connect different parts of documents—particularly in the index, table of contents and references
- Include DOIs in bibliographies, so sources can be obtained at a click
- Include links to external data sources in permanent repositories (e.g. ADS or DOI), so readers have access to full datasets and can extract data easily
- Present phase plans in PDF layers, so readers can turn individual phases on and off. Likewise, distribution plans can be layered, so readers can interrogate the combinations that interest them.

REFERENCES

Darvill, T., Barrass, K., Constant, V., Milner, E. and Brown, E. (2018) *Archaeology in the PPG16 Era: Investigations in England 1990–2010*. Oxford: Oxbow Books.

Ministry of Housing, Communities and Local Government (2019) *National Planning Policy Framework*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/810197/NPPF_Feb_2019_revised.pdf

Thomas, J.D. (2004) Online Archives. Internet Archaeology 15. [https://intarch.ac.uk/journal/issue15/richards_index.html]

APPENDIX: SURVEY QUESTIONS

Q1 Which project(s) did you volunteer for?
- Project 1: Roman Cultivation Strips
- Project 2: Violence and the Dead
- Project 3: Roman and Saxon Grave goods
- Project 4: Extinction of the Aurochs
- Project 5: Saxon buildings
- Project 6: Emerging Henges
- Project 7: Sheepfolds in the Lammermuirs
- Project 8: The Guardians of Dartmoor
- Project 9: Coal and Temples in Roman Britain
- Project 10: Decoration and Design in Prehistory
- Project 11: Reconstructing the Wildscape
- Project 12: Rapid review of digital pads

Q2 Rate how useful the following tools and sources were in identifying where you could get information from:

Source	Useless	Limited Use	Somewhat Useful	Fairly useful	Very useful	N/A
ADS Search Function						
Heritage Gateway						
Online Search of individual HERS						
Requested data from individual HERs						
Online archives of commercial units						
Requested help from commercial units						
Requests on Facebook/Social media						
Published gazetteers						
Student dissertations						
References in publications/reports						
Google/other search engines						
Other (please specify below)						

Q3 What were the two or three most useful aspects of these tools which helped you identify the sources of information you needed for your project? Consider:
- Was the website searchable
- Did the results correspond with the keywords in your search?
- Were links to resources included?

Q4 What were the two or three biggest barriers to identifying the sources of information you needed for your project?

Q5 Rate how easy is was to obtain the specific resources you wanted from the following sources?

Source	Impossible	Difficult	Somewhat easy	Fairly easy	Very Easy	N/A
Online county journals						
Printed county journals (not online)						
Online academic journals behind paywalls						
Online academic textbooks and handbooks						
Printed textbooks and handbooks						
Printed textbooks or edited volumes						
Printed excavation monographs						
Downloading from ADS Library (scanned publications)						
Downloading ADS Grey Literature (PDF reports)						
Requesting reports from HERs						
Requesting articles/reports from authors						
Requesting reports from commercial units						
Google Books						
Internet Archive						
The Archaeology on Furlough website						
Requests on Facebook (AoF, BAJR, etc.)						
Other (specify below)						

Q6 What were the two or three most useful features of the sources you used that made it easy to obtain the sources you wanted? Consider:
- Was there a direct link to the resource?
- Were all the tables, figures, and appendices included with the main text or did you have to look for them separately?
- Was the resource downloaded?

Q7 What were the two or three biggest barriers you experienced when trying to obtain the resources you wanted?

Q8 Rate how easy is was to use the following types of information.

Resources	Impossible	Difficult	Somewhat easy	Fairly easy	Very Easy	N/A
Printed monographs						
PDFs of monographs						
Searchable PDFs of grey literature reports						
Scans of grey literature reports						
Printed journal articles						
Searchable PDFs of journal articles						
Heritage gateway entries						
Existing gazetteers						
Other (specify resource below)						

Q9 What were the two or three most useful features of the resources you used that made it easy to extract the information you wanted and transfer it into your research? Consider:

- Was the resource complete?
- Was all information legible?
- Was the data presented/formatted in a way that made it easy to use?
- Were you able to copy the data efficiently?

Q10 What were the two or three biggest barriers to extracting the information you wanted from the resources you used and transferring that information into your research?