



This PDF is a simplified version of the original article published in *Internet Archaeology*. Enlarged images, videos and appendices which support this publication can be found in the original version online. All links also go to the online version.

Please cite this as: May, K., Taylor, J.S. and Binding, C. 2023 *Stratigraphic Analysis and The Matrix: connecting and reusing digital records and archives of archaeological investigations*, *Internet Archaeology* 61. <https://doi.org/10.11141/ia.61.2>

Stratigraphic Analysis and The Matrix: connecting and reusing digital records and archives of archaeological investigations

Keith May, James S. Taylor and Ceri Binding

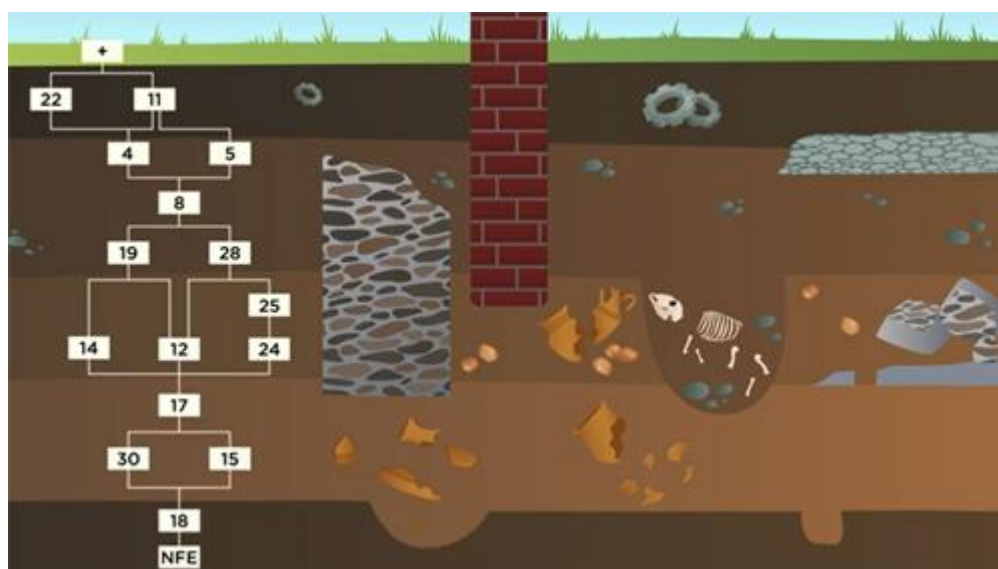


Illustration of archaeological stratigraphic laws and principles. Taken from *What is The Matrix? How do archaeologists use stratigraphy?* [YouTube](#)

Stratigraphic data and relationships form the backbone of all the related archaeological records from each excavated site and, along with the phasing and interpretive information derived through stratigraphic analysis, are essential for chronological modelling, broader synthesis of inter-site phases and periods and, we argue in this paper and elsewhere, stratigraphic data should be a required component in digital archives of the growing body of archaeological information and reports generated through the commercial archaeological sector in the UK and internationally.

Not every site has complex stratigraphy, but understanding the nature of the stratigraphy, be that deep or shallow, complex or otherwise, enables researchers to



piece together the underlying details of how the excavator(s) arrived at the interpretations they have made about the site. The stratigraphic record, including associated relationships and data, which in the case of complex stratigraphy are usually visualized in the form of a stratigraphic matrix diagram, acts as a primary, if not the *primary evidence* for how, and in what order, the site was excavated. As such the stratigraphic data can be the key mechanism that enables anyone less familiar with the site, to re-visit and re-use the excavation records; understand what data is most relevant for addressing certain research questions; or grasp the nature of the chronological sequence encountered; and piece together the underlying details of how the excavator(s) arrived at their interpretations. However such records are often only held on paper or as scanned image copies (as PDFs) of matrix diagrams that cannot easily be re-used with all the associated data. This article presents outcomes from The Matrix project (AHRC AH/T002093/1) that address the current problems caused by the lack of standardized approaches to digital archiving of archaeological data using the case study of stratigraphic and phasing data.

1. Background and Introduction to The Matrix Project

This article presents outcomes from [The Matrix](#) project (AHRC AH/T002093/1) that address the current problems caused by the lack of standardised approaches to digital archiving of archaeological data, using the case study of stratigraphic and phasing data. Stratigraphic data and relationships form the backbone of all the related archaeological records from each excavated site. Along with the phasing and interpretative information derived through stratigraphic analysis, they are the essential evidence underpinning integrated chronological analysis, wider synthesis of inter-site phases and periods, and thereby semantically-rich, interoperable (and Findable, Accessible, Interoperable and Reusable: FAIR) archiving of the growing body of archaeological data and reports generated through the commercial archaeological sector in the UK and internationally.

The stratigraphic record acts as a primary, if not the primary 'evidence' for how, and in what order, a site was excavated. By stratigraphic record, we mean all the recorded data about stratigraphic relationships (above/below/equals), stratigraphic units (contexts) and, when complexity demands, often illustrated in the form of a stratigraphic matrix diagram. Not every site has complex stratigraphy, but understanding the nature of the stratigraphy, be that deep, shallow, complex or otherwise, enables researchers to piece together the underlying details of how the excavator(s) arrived at the interpretations they have made about the site.

Stratigraphic analysis is the term used to collectively describe the work undertaken to check and validate the depositional sequence, to provide interpretative grouping of



stratigraphic units, and to derive functional interpretation and spatiotemporal phasing of the primary record of stratigraphy from an excavation (Roskams [2001](#), 239-66). As such, stratigraphic analysis is usually, but not necessarily, undertaken by archaeologists during the post-excavation analysis. If the stratigraphy is complex, a grouped and phased stratigraphic matrix diagram can be the key mechanism that enables anyone less familiar with the site to re-visualise, re-visit and reuse the excavation records, to understand what data are most relevant for addressing certain research questions, problems encountered, and how interpretations are arrived at.

However, currently parts of these records are often only held on paper or as scanned image copies (PDFs) of matrix diagrams that cannot easily be reused with all the associated analytical data that gets incorporated into a matrix diagram during post-excavation analysis (e.g. phase lines, group matrices or the interpreted temporal relationships that are based upon often limited, and sometimes uncertain, dating evidence from associated finds objects). Quite often, the key underlying interpretative phasing data from analysis (which would be the supportive evidence for the conclusions published as text in an excavation report, and on which broader comparative synthetic publications may be based) are not consistently archived, if at all. Yet it is these types of data and related interpretations that underpin conclusions in publications of multi-phase projects and correlations and synthesis of inter-site phasing (Bradley [2006](#)) - let alone any further work on Bayesian chronological modelling (Buck *et al.* [1996](#)) or semantic cross-search (Tudhope *et al.* [2011](#)). This results in key records being unsearchable or remaining unconnected (un-interoperable) with other data and, at best, usually requires lengthy and wasteful re-keying if anyone wishes to work with and reuse the archives from such sites (un-FAIR). The focus of digital repositories and museums is now switching from simply enabling better online access to digital archives, to answering questions of how users in commercial units, curatorial organisations and academia (and the general public) are going to make best (re)use of this growing body of digital information and data.

1.1 The research questions and project methods

The Matrix project has investigated how digital data from archaeological excavations can be made more useful and interesting to a range of users and audiences. The project had four main areas where this investigation was focused:

1. Digital data reuse – especially data from analysis processes
2. Heritage Data characterisation - especially data from analysis of stratigraphic data
3. Stratigraphy standards
4. Research tools, especially for stratigraphic analysis

The project and this article are addressing two topics from the [Historic England Research Agenda \(2017\)](#). The first is to encourage better sharing, reuse and interoperability of archaeological data and information derived from the commercial sector. The second is ensuring the consistent development, application and enforcement of technical information and data standards. This article is part of publicising that developing plan and methods to get such data more consistently



recorded, analysed, disseminated and archived in a way that is FAIR. The overall aim is to maximise public value and enhance the research potential of the archaeological data being recorded and preserved.

The project's original plan was to carry out a review of what standard methods and processes are more commonly applied at different stages of stratigraphic analysis, moving through the current processes from:

1. checking recorded stratigraphic relationships;
2. identifying collections of stratigraphic units for interpretative Grouping as identifiable archaeological activities or features;
3. site-wide phasing of the stratigraphic sequences.

Visits were planned to around ten of the major archaeological contracting organisations in the UK to gather documentation relating to the post-excavation analysis processes that were undertaken in each organisation. The intention was then to model these documented post-excavation processes in a way that could identify and collate the most common methods used by archaeologists undertaking post-excavation stratigraphic analysis, and the data generated. A number of initial visits to [MOLA](#), [PCA](#), [Oxford Archaeology](#) and [Red River Archaeology](#) took place, but as the COVID-19 pandemic unfurled in the early months of the project in 2020 and Government restrictions on travel were put in place, much of the original work envisaged ended up being undertaken via online (Zoom) meetings and sharing of digital resources. In the course of these research activities, a series of associated issues were explored in more detail, leading to many of the [conclusions and recommendations](#) presented in this article.

1.2 What is the problem? Requirements and (re)uses for FAIR stratigraphic (and other) archaeological records

Why do we need to reuse our stratigraphic data? What is the impetus for developing and using software tools for stratigraphic analysis at all and storing our stratigraphic data in a format that is FAIR? The findings of this project suggest that FAIR archaeological archiving isn't currently the norm in commercial archaeological practice and Research Data Management best practice is still more of an aspiration than a routine in both academic and commercial archaeological practice around the world (Richards *et al.* [2021](#)). However, this situation may be changing as digital archiving of commercial archaeological data in the public domain is finally becoming more commonplace. It is worth noting here that even if FAIR access to (meta)data is achieved in a digital archive, that does not necessarily result in freely available and open digital data for reuse (Higman *et al.* [2019](#)). This article will argue that the 'R' for Reuse in [FAIR](#) should reflect truly useful human reuse, rather than a certain level of 'machine-readable' reuse. Machine-readable reuse simply makes a dataset more retrievable, while FAIR does not necessarily always mean freely available, so it can sometimes mean little more than make certain archived metadata more useful to machines. There is an argument that truly human reusable data means openly



reusable data (Costa *et al.* [2013](#)). In that sense, to make data humanly reusable, it really needs to be available for FAIR + Open (FAIRO?) reuse.

The following sections will review a number of common use cases where stratigraphic data might need to be queried, revisited, updated, integrated into new datasets or otherwise remixed and reused (Huggett [2018](#)) and ideally be FAIRO.

1.2.1 Transparency in interpretation and evidence in data: an ethical imperative

Archaeological knowledge is complicated. Producing archaeological knowledge is difficult work that involves a lot of potential (and imperfect) data constrained by a wide variety of parameters and constraints. Where do we stop digging? What do we retrieve (or throw away)? What should we record? Should we sample? Sub-sample? The most common output of the excavation process and representation of our knowledge is the written narrative (whether that be a complete synthesised report with full analysis and a large collection of data, or a more technical grey literature assessment report). Huvila *et al.* ([2021](#), 16) highlight that the process of report writing in archaeology is in reality 'messy and disjointed', despite our best efforts to smooth it over with a carefully crafted narrative. As we will show, the analysis of the stratigraphy (correlation, grouping and phasing) forms a key part of that process which is in itself 'messy and disjointed', but nevertheless both underpins and scaffolds those same narrative outputs - even if that stratigraphic data and interpretation rarely makes its way into the final (digital) archive.

There has been a strong focus in recent decades upon the standardisation of the metadata of our digital objects (see e.g. Archaeology Data Service and Digital Antiquity [2009](#)) and recently there has been a growing discussion around the importance of paradata as a mechanism for understanding the context for the creation of knowledge (see e.g. the discussion put forward in Börjesson *et al.* [2020](#)). Elsewhere, Huvila also stresses the importance of paradata arguing that

'without proper documentation of the human processes of creating, understanding and interpreting data objects, there is a risk of creating and archiving large collections of data that are incapable of supporting research and other types of reuse' (Huvila [2022](#), 41)

Accepting the limitations of most paradata and lack of any standard practice here (cf. for example the production of metadata), Börjesson *et al.* ([2020](#), 192) argue that 'well-structured process descriptions are vital to maintaining insight into the production of visualizations'. If we accept that our chronologies and narratives are a key visual output, increasingly underpinned by a wide range of digital (as well as more traditional analogue) data, then understanding the context of that data becomes important in the interests of transparency of knowledge creation.

Digital archaeologists often talk about the need to replicate the analytical process, what Huvila *et al.* ([2021](#), 16) describes as a tendency to see archaeological documentation as a 'means to redo an earlier excavation', whether or not this is a conscious intention of the archaeologist. While the capacity of future researchers being able to re-do excavations from the archive is largely a fallacy (they are a



necessarily reductive and selective medium of representation in their own right), it does suggest that many archaeologists at least want to be transparent in their process in case others wish to revisit their interpretative process. This is perhaps more in line with reflexive post-processual thinking, which has long called for innovation and democratisation within the interpretative process (see e.g. Shanks and Tilley [1992](#); Shanks and McGuire [1996](#); Hodder [1997](#); [1998](#); [1999](#); [2000](#); Chadwick [2001](#); Berggren and Hodder [2003](#)) and perhaps is in accord with more recent calls for slower, more reflexive, digital workflows and digital systems that embrace the complexity of the archaeological record (see Perry and Taylor [2018](#)). In this sense, transparency of process can be seen as an ethical responsibility of the archaeologist and given how pivotal the stratigraphy is to the interpretation of single context excavation data, one can make the case that there is an ethical imperative for the 'raw' stratigraphic data to be available, and a scientific responsibility (Popper [2002](#)) for the 'analysis' data in archives to be made more 'FAIR' and open (FAIRO) in order to lay bare, and evidence, our thinking.

1.2.2 Reinterpretation: revisiting a particular dataset

So what about revisiting the data? We have argued that being able to *re-do* the excavation from the archive is a fallacy, but this doesn't mean that future archaeologists might not need to revisit our work in order to understand our interpretative process (and perhaps re-interpret our data). Within the internal structure of most research projects, this happens all the time (as areas are reopened season after season, and last season's interpretations are re-evaluated and modified in the light of new data). There are legitimate reasons why data might need to be re-evaluated and re-interpreted in this way within the commercial sector.

There is a need for 'joined up data' at a landscape scale (Doneus *et al.* [2022](#)), particularly when excavations are carried out at different times in adjacent areas, where the earlier intervention has already been archaeologically understood to some extent (i.e. open quarries and gravel extraction sites, expansion of housing developments). The increase in large scale infrastructure projects (Aitchison *et al.* [2021](#)), often with associated digital project management systems specifically designed to 'join up the data' (e.g. HS2, Crossrail, etc.), further highlights this need.

1.2.3 Deposit modelling: interoperable stratigraphy

If synthesising excavation data in this way to create regional or urban narratives represents a *post-hoc* approach to linking and reusing stratigraphic data and associated archives, then this might be compared with the more pre-emptive practice of deposit modelling in archaeology. Deposit modelling in archaeology also has a long pedigree, with key early examples including Biddle *et al.*'s ([1973](#)) *The Future of London's Past* and Arup *et al.*'s ([1991](#)) *York Development and Archaeology Study*. Carey *et al.* ([2018](#), 4) define deposit models in their simplest form as 'visual representations of the spatial and stratigraphic relationships between sediments, archaeological and palaeoenvironmental remains in areas preserving both vertically and laterally accreting sediment sequences'. They have been closely linked to broader approaches to modelling and landscape characterisation (Carey *et al.* [2018](#)).



By definition therefore, deposit models rely heavily upon a cross-understanding of the stratigraphy, chronology and phasing of sites across a region or area. Crucially, deposit modelling (particularly in the urban environment) has effectively become a requirement within planning policy guidance and frameworks for strategic environmental assessment, to mitigate and monitor the effects of large-scale developments and public sector plans and strategies. At a more focused or local level, these practices can be seen to extend to the production of the ubiquitous Desk-Based Assessment, where researchers regularly need to drill down into the stratigraphy of adjacent sites in order to assess the potential for archaeological remains within the agreed study area of the development.

1.2.4 Bayesian chronological modelling: vive la révolution

It has been argued that the so-called Bayesian Revolution (Naylor and Smith [1988](#); Buck *et al.* [1996](#); Bronk Ramsey [2008](#); Bayliss [2009](#)) has had, and continues to have, a profound impact on the discipline of archaeology (see Griffiths [2017](#)). This is exemplified by a number of synthetic outputs that have significantly changed our story of prehistory (e.g. Whittle *et al.* [2011](#) or Whitehouse *et al.* [2014](#)). It is hard to quantify the lasting extent to which these approaches will impact our overall understanding of the past, but one thing is clear: Bayesian chronologists rely upon archaeological stratigraphy as a 'prior belief' framework for informing the 'standardised likelihoods' (dates) that inform their Bayesian Models (see Bayliss [2009](#) 127-32). Indeed Bayliss stresses that:

'it is absolutely critical to emphasize the fundamental importance of the taphonomy of the dated samples. How did the datable material get into the deposit from which it was recovered?'

stating that in order to do this

'all samples must be presumed to be residual (older than the deposit in which they were found) unless there is evidence, or at least a convincing argument, that they were fresh when deposited' (Bayliss [2009](#), 129).

Simply put, this line of reasoning means that for Bayesian chronologists, revisiting and understanding the stratigraphy is critical for the contextualisation and understanding of the datable material that underpins their models.

When dealing with specific sites, Bayesian chronologists regularly seek to drill down into legacy site stratigraphic data, so access to that data is particularly important to them (Moody *et al.* [2021](#)). Other research projects have attempted to draw together archaeological archives from different excavation teams to analyse the temporal sequences and use the stratigraphic relationships recorded on site to cross-search using semantic technologies for artefacts and structures from related phases (Tudhope *et al.* [2011](#)). But a lack of consistent practice in digital deposition of such records has placed considerable limitations on the amounts of archaeological records available for such analyses.



For chronological modellers of archaeological data, these problems are exacerbated by a lack of standardised approaches to the archiving of stratigraphic data, often held in hard-copy matrix diagrams or inconsistently structured database tables. The Matrix project is helping to inform decisions on digital archiving standards and best practice for stratigraphic data deposition and reuse, so that such digital data can be held and reused in the most suitable form for input into Bayesian calibration software such as BCal, OxCal, or Chronomodel. The use of Bayesian chronological modelling techniques has become critical in the more accurate dating of archaeological sites and phases over the last ten years, but the way such information is analysed is quite painstaking and often involves many hours of laborious manual data preparation for key staff involved (Dye and Buck [2015](#)).

2. Stratigraphic Laws and Principles for Stratigraphic Analysis

Different types of archaeology can have very different stratigraphy, but having a clear record from the excavator of the nature of the stratigraphy should enable other researchers to understand the underlying evidence for how the excavator(s) arrived at the interpretations they have made about the site. On many archaeological sites, the archaeological stratigraphy forms the backbone of all the related archaeological records and while we acknowledge that not all sites will require intensive stratigraphic analysis, all archaeological excavations will have stratigraphy to a greater or lesser degree. We would argue therefore that the data resulting from stratigraphic excavation records are essential for integrated analysis, intelligible synthetic publication and accessible and reusable archiving of complex or inter-related sites. Stratigraphic recording in archaeology is based upon a number of fundamental laws and principles set out most clearly and seminaly in the work of Harris ([1979](#); [1989](#)) but these were built upon a succession of shoulders of geological and archaeological 'giants' (e.g. Steno, Hutton, Smith, Lyell, Pitt-Rivers, Wheeler, Kenyon). Rather than attempting another overview of these stratigraphic principles, a very simple exposition of the basic principles of archaeological stratigraphy are illustrated by the animation in Figure 1.

[\[Watch on YouTube\]](#)

Figure 1: Animated illustration of archaeological stratigraphic laws and principles. This video has sound.

In the following sections, we will delve deeper into aspects of these stratigraphic principles and methods that are key to the sorts of stratigraphic analysis commonly undertaken as part of post-excavation practice and which have been examined by the Matrix project. Particular focus has been placed on the specification of requirements for the stratigraphic matrix analysis prototype software tool that has been entitled Phaser (see [Section 7](#)).



2.1 Stratigraphic units as units of spacetime

The principles laid down by Harris are considered sound and the four laws of archaeological stratigraphy have been widely adopted (Barker [1977](#); Spence [1990](#); Carver [2009](#); Roskams [2001](#)), although not necessarily universally, among archaeologists (Pavel [2010](#)). In defining the essential stratigraphic relationship between two archaeological stratigraphic units, the principles of archaeological stratigraphy set out the relative positions of two related stratigraphic units in spacetime. The Law of Superposition and Law of Stratigraphic Succession, in separate ways, define a stratigraphic unit as a unit of spacetime.

The 'life-span' of a stratigraphic unit (SU) can be understood as the persistence of the spatio-temporal extents of that stratigraphic unit as a four-dimensional entity, or 'chunk of spacetime', from its first deposition through its use or reuse, and its continued existence up to the present if the SU is preserved *in situ* as part of an archaeological display (see e.g. discussion in Lucas [2001](#), 160-62, and Taylor [2016](#)). The stratigraphic matrix allows us to visualise, analyse and interpret, the spatio-temporal inter-relationships of the various chunks of spacetime that have persisted, and how those chunks of spacetime fitted together through previous activities over time up to the point of destruction by excavation (see [Recommendation 4.2](#)).

2.2 Groups, sub-groups, phase and period – terminology used in stratigraphic software development

In addition to incorporating the four basic laws and the established stratigraphic principles in the prototype software (see [Section 7](#)), we have tried to adopt and enable use of the most common approaches by archaeological stratigraphers to stratigraphic analysis. Most of the archaeologists we spoke to carried out some form of singular or iterative grouping and phasing process as part of their stratigraphic work. A previous article (May [2020](#)) sets out the general definitions of group, sub-group and phase that are adopted in the Matrix project work. Some additional commonly encountered 'rules' of stratigraphic analysis were also used in the prototype - see [Appendix A](#) and [Appendix B](#) for data validation rules used.

Higher order synthetic groupings, such as land-use diagrams (see e.g. Spence [1993](#); Westman and Shepherd [1992](#); Steane [1993](#)), require a completed Harris Matrix to create. Our investigations revealed that land-use processes are usually recorded in data records from London (especially by MOLA), but land-use diagrams are not created as a matter of course, other than on sites with complex stratigraphy, and they are not considered an integral part of the 'Harris matrix' methodology. Anecdotally, even with the assistance of the archivists, we could find no examples included in archives of stratigraphic data held by the Archaeology Data Service (ADS). As such, land-use processes were considered beyond the scope of this project's process modelling for stratigraphic analysis and we have not included them as a requirement in the development of the matrix analysis software prototype.



Dye and Buck have also pointed out the ambiguities in the usage of the terms such as 'period' and 'phase' between archaeologists and chronological modellers.

'The terms "period" and "phase" are defined variously and sometimes interchangeably by archaeologists ... Because "phase" is also used to describe Bayesian chronological models, here we use the term "stratigraphic phase" to refer to a group of contexts, and the term 'chronological phase' to refer to a time period in a chronological model' (Dye and Buck [2015](#) 87).

We will consider the possible derivation of some of these ambiguities and some approaches to dealing with them in the following section.

2.3 The use of the matrix diagram for phasing visualisation and interpretation

The value of the matrix diagram is that it enables archaeologists to record and visualise the 4-D spatio-temporal entities and the relationships between a sequence of Stratigraphic Units, both during the recording and the subsequent stratigraphic analysis of those Stratigraphic Units (Balm [2015](#), 112-15). As Harris puts it 'A stratigraphic sequence is a diagram of relative time: it shows all four dimensions of the stratigraphic accumulation of a site, unlike the two-dimensional image of the physical world of stratified deposits seen in a section' (Harris *et al.* [1993](#), 18)

In this way, the matrix diagram also allows the archaeologist to de-construct and re-configure the way in which the 4-D segments of an archaeological site have been accumulated and consequently de-constructed by excavation. The stratigraphic matrix also records the persistence of a Stratigraphic Unit through spacetime and enables the visualisation of how that Stratigraphic Unit fits together with all the other different segments of spacetime across the site, like one of a series of building blocks. The matrix allows us to see, and represent, the inter-relationships of the various 'chunks of spacetime' that have continued to exist into the present, and how they fit together and have persisted until they are excavated.

3. Matrix Project Activities - Initial Visits and Consultation with Organisations

In order to address some of the identified issues and consider how we might improve best practice in the digital archiving of stratigraphic data, it was important to communicate with practitioners in the organisations to establish what were current practices. The approach taken was to hold informal consultations with representatives of key organisations (see Table 1), followed by a pair of structured workshops throughout the project, designed to explore organisational approaches to post-excavation work and stratigraphic analysis. Initial contact was made around January 2020 by email to heads of archaeological organisations (CEOs or Directors)



setting out the broad nature and aims of the Matrix research project and asking for help in identifying the most appropriate members of staff to contact for telephone consultations and potential visits.

Table 1: List of organisations consulted by the Matrix project

Organisations contacted in the UK

Federation of Archaeological Managers and Employers (FAME)

Museum of London Archaeology (MOLA)

Archaeology South East (ASE) UCL

York Archaeological Trust

Pre-Construct Archaeology (PCA)

MOLA-Headland

Oxford Archaeology

Wessex Archaeology

Cotswold Archaeology

Red River Archaeology (Rubicon)

L-P Archaeology

DigVentures

Landscape Research Centre



3.1 Related issues with digital data archives of stratigraphic records

In addition to seeking the documentation of post-excavation procedures, a search was undertaken for any suitable datasets from projects that had already been archived at the UK's main digital repository, the ADS at the University of York. This would further help to identify what were the most common approaches to digital outputs from post-excavation stratigraphic analysis practice and isolate the most common data to help inform a 'to be' model for a consistent data package that would enable the finding and reuse of relevant data (i.e. a data package for stratigraphic and chronological data). This work would be used to inform what kind of tool would be most useful to aid those archive and reuse processes.

The project was looking for examples of data from sites that had completed excavation through the analysis of finds dating and included the full phasing (periodisation) of data records so that they could provide a testbed to be used in the design and testing of the prototype software. However, the results of that search on ADS revealed a distinct lack of comprehensive archives from commercial archaeological investigations that contained suitable digital data. To be clearer here, there were a good number of well-archived projects, some of them large scale (e.g. [Elms Farm](#), [T5](#), [CTRL](#), [Crossrail](#)), that included the primary stratigraphic relationships (e.g. 'Above' and 'Below' relations between individual stratigraphic units). In a very few cases there were even examples of matrix diagrams archived, although these were images (.SVG) that did not link easily with the associated database records, e.g. [Silchester LEAP example](#) (Clarke *et al.* 2007). Much more commonly lacking was archive materials from any subsequent analysis and publication-orientated work such as the interpretative Grouping and Phasing of the 'raw' excavation stratigraphic data. Of all the archives searched on ADS, the [XSM10 Crossrail archive](#) was the only site that had the full range of stratigraphic data needed for our project (see [Section 4.6](#)).

3.2 Stratigraphy (or lack of) in digital archive contents

These findings relate to parallel work by Bryony Moody that was undertaken as part of a related Collaborative Doctoral Partnership studentship, to examine OASIS reports and digital archives within ADS for data relevant to her work on Bayesian chronological modelling (Moody in [prep.](#)). Moody was searching for examples of datasets containing one (ideally all) of the following:

- Stratigraphic primary data
- Stratigraphic diagrams
- Harris matrix
- Radiocarbon dating results

Moody *et al.* ([2021](#)) were attempting to find archive data from investigations that could be used to inform and test software development of a new graph database



platform and user interface for an updated version of B-Cal software to aid Bayesian chronological modelling (Buck *et al.* 1996). However, they found similar evidence that very few commercial archaeological projects have reached full analysis and deposited the results of their phasing data in a way that can be linked back to the primary stratigraphic records of contexts in the stratigraphic sequence.

Moody assessed both the deposited ADS digital archives and a large sample of the available OASIS reports in the ADS library for examples of relative and 'absolute' chronological dating evidence. Out of 37,320 OASIS reports, although over 10,000 made mention of stratigraphy, only 358 reports had any record of a matrix diagram, and of those, nearly all only had the diagram in a PDF format (see Figure 2), meaning that the data was not readily reusable or would require complete re-entry of all the data by hand. Moody also noted that OASIS Reports 'often stated that the stratigraphy and phasing information was contained in the stratigraphic archive, with no suggestion as to where we might find this' (Moody 2019). This may be due to the fact that large numbers of OASIS reports are produced and submitted at a stage before the full analysis of dating evidence, stratigraphic analysis and phasing of the archaeology are carried out. Until relatively recently, the so-called 'developer report' was the only digital output required by the Local Authority to be archived. This practice is gradually changing with the increased emphasis on increasing the public benefit from the work undertaken through the planning system.

Research to inform Bayesian Chronological Modelling

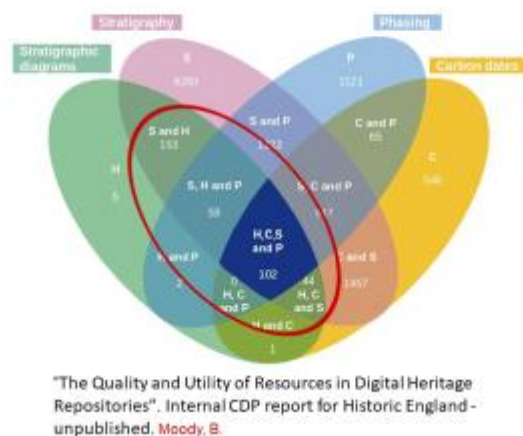


Figure 2: Venn diagram showing the total numbers of documents from a sample of 10,000+ OASIS reports containing each combination of chronological dating evidence types as defined by each segment label. (Each total represents the number of documents containing that particular combination of dating evidence type only and none of the others) (Moody 2019)

It is worth emphasising here that it is not that Moody *et al.* (2021) did not find any stratigraphic data. Indeed there is a considerable amount of 'raw' stratigraphic data present in some of the ADS archives and Moody was able to make use of this with direction from archaeologists. However, the issue is that without this 'insider knowledge', the problems of retrieval of any suitable archives for reuse by the chronological modellers would have proved insurmountable.



A further issue was that when stratigraphic data was identified, it was not deposited in a consistent structure and therefore required additional processing to enable reuse of the primary data. For the purposes of the Bayesian chronological modelling work, Moody *et al.* required 'A table containing mutually consistent pairwise statements of the stratigraphic relationships between contexts (stratigraphic units) as they were observed in the field' ([2021](#), 3). In practice, the stratigraphic data that was accessible in the archives came in various 'flavours' or combinations, reflecting the fact that each archaeological organisation deposited stratigraphic records as they had been held in their own database systems. We also encountered this issue and a certain amount of work was needed to 'transform' data from e.g. .LST format to a more usable CSV structure.

During consultations conducted as part of this project, it became apparent that there is little standardisation in the way that organisations deal with the stratigraphic components of their archaeological archives. It seems likely that the *ad hoc* approach to what software is used in post-excavation analysis (which varies project by project depending upon a variety of factors, such as location, complexity of the depositional sequence, post-ex culture, or indeed individual working practices), combined with a limited amount of documentation setting out what outputs might be expected as part of the stratigraphic analysis process, seems to be a major factor in explaining the lack of identifiable stratigraphic data in the digital archives on ADS, or elsewhere. When questioned about this, an issue raised by several consultees was the anticipated costs of archiving multiple digital files. Most bluntly, it was expressed as 'ADS costs too much'. Whether true or not, the perception of the relative cost of digital archiving for commercial organisations is a challenge.

These issues are compounded by a lack of consensus in practice over what is expected to be in the digital archive of an archaeological site, and what archiving costs are appropriate to pass on to developers who fund projects. This means that many contractors are reluctant to add costs for viable digital archiving, because of concerns that others can undercut by not including adequate costs for digital archiving. Although there are valuable recent initiatives that have begun to address what is included in the digital archive from excavations (see ClfA's '[Dig Digital](#)' project), the actual processes used in the analysis stage of projects still varies quite markedly, and therefore the by-products from that stage of the process, regardless of whether excavation data have been deposited, are far less consistent in the resulting digital archives. In this way, agreeing the process(es) for post-excavation practice across organisations should help set a more level playing field for commercial archaeological practice, including charges for archiving the products of post-excavation. It is recommended ([Recommendation 1.2](#)) that work is undertaken collaboratively across the sector to better quantify and express the cost-benefits of adequate digital archiving and better manage the risks involved with charging developers for archiving. This should result in more consistent approaches for creating the (digital) outputs from stratigraphic analysis using existing stratigraphic principles and recognised standards, and especially for identifying the common stratigraphic data required as part of a consistent, complete digital archive deposition.



3.3 Data gathering issues for the Matrix project

Telephone calls were made to all the identified organisations (Table 1) using a list of questions so that similar information was obtained from each. Not intended to be a formal survey questionnaire, it was rather to provide representatives with an introduction to the main research areas of the project, to gather information about process documentation and what data might exist, as well as the degree of cooperation they were able to offer to the project. Separate consultations were also conducted with members of the Historic England Scientific Dating Team.

Initial discussions sought to establish whether there were any commonly used manuals in use for post-excavation analysis procedures, the organisational approaches to stratigraphic analysis and any common processes for phasing in post-excavation and digital archiving. Because of the lack of documentation on the subject, we feel a need to set out what we mean here by 'post-excavation' processes. Post-excavation is a general term used by archaeologists to encompass the various processes of finds analysis and a series of related activities relating to pre-publication analysis work. This includes any specialist work on analysing finds artefacts, palaeoenvironmental analysis of ecofacts, scientific dating and other dating of objects on stylistic grounds, as well as the stratigraphic analysis that draws all of these together to make interpretative assertions about the types of features, structures, activities and phases of land-use and activity that went on during the lifetime of a site. The precise definition is made more difficult because the actual processes of analysis that might be involved in post-excavation cannot simply be pre-determined, as they depend largely upon what types of archaeology, and archaeological objects, are encountered and discovered (often unexpectedly) during excavation. So only if waterlogged layers that preserve organic materials like leather are encountered, is it likely that any leather objects will be recovered from an excavation, and that analysis of such leather objects would feature in that site's post-excavation analysis process.

The enquiries seeking post-excavation manuals were subsequently followed-up initially with visits to various organisations or (during the pandemic restrictions) remotely via email. In the end, with one exception for an infrastructure project, none of the organisations consulted responded by providing formally documented procedures for post-excavation, particularly relating to the stratigraphic analysis and phasing processes (although one 'slightly out of date' document was provided to us by a individual late on in the project). Only MOLA-Headland Infrastructure was able to provide a manual produced relatively recently for staff working on the post-excavation programme for the A14 road infrastructure project (see Figure 3).



Figure 3a: Example of post-excavation analysis documentation from MOLA-Headland Infrastructure A14 project (Image: MOLA-Headland Infrastructure and National Highways)

[\[ONLINE ONLY\]](#)

Figure 3b: Animated graphic of post-excavation analysis documentation from MOLA-Headland Infrastructure A14 project

Curiously in most cases, in lieu of specific post-excavation documentation, organisations tended to provide copies of excavation recording pro-forma as an alternative, to help indicate where standard procedures for recording practice were being followed. This is interesting in its own right, because the implication here is that if one adopts a standard, pro-forma approach to recording on site, that there is a standard approach to dealing with that data off site (i.e. post-ex); something that did not prove to be the case when discussed openly in the follow-up workshops.

3.4 Workshop inputs and feedback

A number of informal consultations, formal workshops and research seminars were held at which feedback was gathered on the process modelling (see [Section 4](#)) and the questions that had arisen about the different organisational approaches to post-excavation documentation.

There was noticeable variation in the outputs of analysis from different organisations that seemed to reflect differences in organisational practice and especially different software. One major factor noted was how the nature and origin of the funding for archaeological projects can heavily influence the sort of outputs that reach the digital archive. Projects that receive research-based funding are generally more likely to have enough funding to take their outputs right through to publication and archive



(Davies [2017](#)). It is noticeable how a much greater proportion (over 95%) of completed projects with archives on ADS are the result of research funding. This may highlight a greater degree of enforcement by research funding bodies such as UKRI and Historic England to require deposition of digital data with an accredited digital repository. Projects that are based on development control related funding have, to date, much less frequently managed to deposit fully completed digital archives. Most often, commercial projects only deposit an investigation report through the [OASIS system](#), but without any accompanying fully analysed digital archive (Tsang [2021](#)), although it must be remembered that the scale and research potential of the archaeology recorded on different types of commercial projects will vary significantly.

However, archives from what can be characterised as 'Infrastructure' projects usually do have resources to complete the analysis and archiving, but because these large projects are most often undertaken by a number of different organisations, the archive outputs often seem to be produced by different analytical methods and do not always form a consistent set of archives with equivalent datasets (e.g. [Crossrail archive](#)).

Additional feedback concerned the differences in the stratigraphic record and resulting records that derive from sites with deep and complex stratigraphy (most commonly found in urban situations), compared to more shallow stratigraphy (often characterised as 'rural' sites).

3.5 Data gathering issues and reflections on standards for stratigraphic data

Current practice for archiving of stratigraphic data from excavations is very variable, and often lacking, particularly for commercially excavated sites; 'at best, 2-3% of all commercial projects have been digitally archived with the ADS' (Tsang [2021](#)). The project investigations and the workshops in particular confirmed that there is no commonly accepted standard to ensure that the primary stratigraphic data from excavations is included in the digital archive (see [Section 3.2](#)). A second major issue already mentioned is that stratigraphic matrix diagrams tend not to be included in the digital archive and, in many cases, grouping and phasing information from the analysis of stratigraphy is not associated with the stratigraphic data (Moody *et al.* [2021](#); Figure 1).

A more positive observation is that primary (i.e. pre-analysis) stratigraphic relationships may be recorded as separate columns in spreadsheets and archived in Comma Separated Value (CSV) format, which at least is useful for preservation and reuse purposes. In other cases, the stratigraphic relationships will be held as part of the site database and archived in a format that such database software enables to be digitally preserved and migrated (again most commonly in .csv format). Even so, and in either case, that does not necessarily guarantee commonality in how the stratigraphic information within the data are represented and preserved. Whatever the format, where the primary stratigraphic relationships are included in the archive, it is much less common for them to be associated with the analytical information to



explain how the stratigraphic units are further defined into interpretative groups, sub-groups and phases. We also saw that site-wide interpretative information based upon dating evidence, including information about land use processes and interpretative land use diagrams (Westman and Shepherd [1992](#), 441, fig.4), are also not archived consistently.

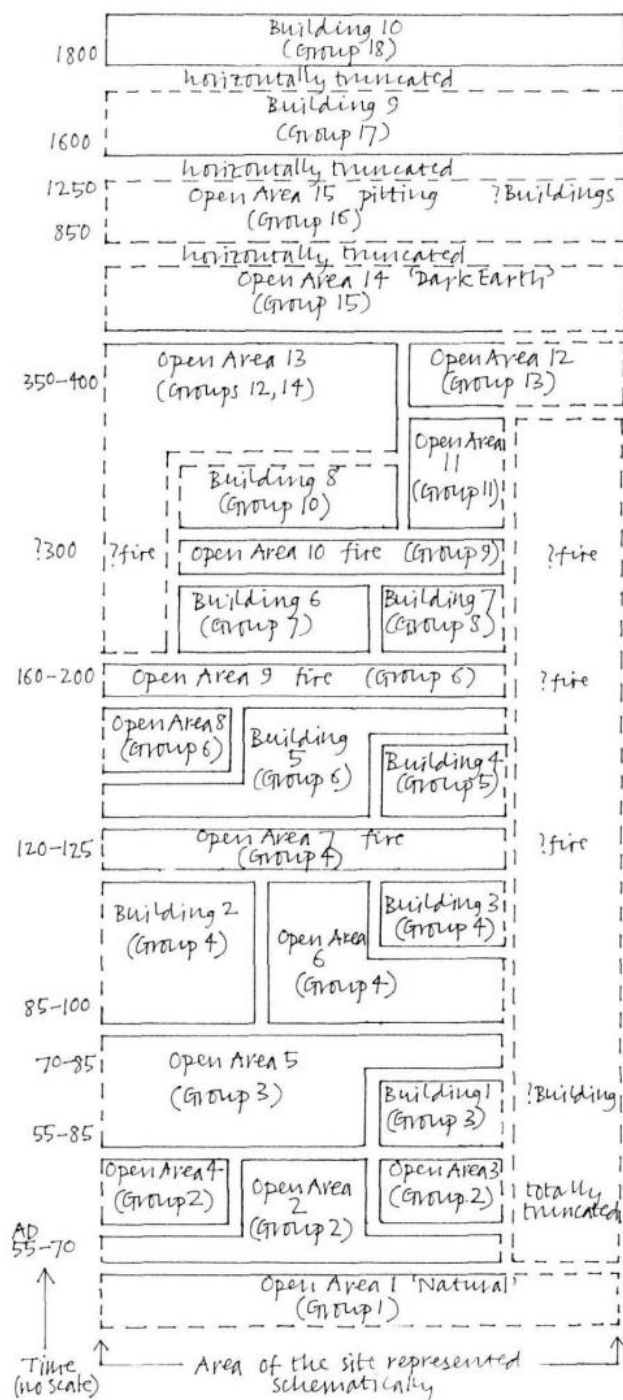


Figure 4: Typical interpretive land use diagram, with dates added (after Westman and Shepherd [1992](#), 441, fig. 4).



3.6 Summary of post-excavation issues encountered

The final digital archive content would appear to be determined as much by the scale and nature of the funding for the archaeological project as it is by the significance of the archaeological material found during an investigation (see [Section 8.1](#)). The funding trajectory of a project, be it 'development funded' (Rocks-Macqueen and Lewis [2019](#)) or 'research funded' can strongly determine the trajectory for the analysis process and subsequent publication and archive outputs. Differences in the scale of funding are also factors in both the quality and quantity of information that reaches the archive. These differing funding trajectories and their outputs are summarised in [Figure 9](#).

There are several likely causes of the inconsistencies found between different post-excavation products deposited in digital archives, and in our experience, is a result of differing combinations of the following factors (not exhaustive):

- Different funding sources (Commercial, Research, Large-scale Infrastructure)
- Different organisational policies
- Differing recording methodologies (particularly an issue in work across countries and/or beyond the UK)
- Lack of awareness, resources and training for digital archiving good practice
- Lack of enforcement of standards for digital archiving good practice
- Differing post-excavation practice
- Lack of systematic manuals for post-excavation process
 - Issues within organisations
 - Lack of consistency across organisations
- Differing archaeology (e.g. deep 'urban' stratigraphy or shallow 'rural' stratigraphy)
- Different software choices, availability and outputs
- Different publication targets and media (resulting in multiple digital 'by-products')
- Interim site reports without archived stratigraphic data

The resulting discrepancies and the fragmentation of the associated data makes reuse of data across different sites much more difficult and impractical, making archived data non-interoperable.

4. Current Practice in Stratigraphic Recording and Analysis - a UK Perspective

The post-excavation stage of an archaeological project encompasses a wide range of quite broad activities, including data validation and consolidation (analogue and digital), interpretation and analysis by various specialists of both field data and material culture, as well as the creation and management of corresponding archives, collections and written outputs for reporting and dissemination of the results in various media to various audiences (see e.g. Roskams [2001](#), 239-66). This process



is complex and multifaceted, and most of those consulted as part of this project suggested, or agreed, that the degree of complexity in the archaeology encountered during excavation on any site will be the major determinant of the complexity of the resulting analysis undertaken. Therefore, the nature and scale of resulting analysis greatly depends on, and reflects, the methodology (processes) used in excavation.

4.1 Post-excavation practices

Despite a long history of thought linked to archaeological methods (see Roskams [2001](#); Carver [2009](#); Lucas [2001](#); Trigger [2006](#) etc.), there is a disciplinary dearth of coherent literature about post-excavation processes and practice. Indeed, in this project's initial scoping exercise, the authors could only find one clear post-excavation manual in the public domain (Hammer [2002](#)). Aside from this, across the sector there are at most a handful of unpublished documents outlining in-house approaches to post-excavation and report writing. Compare this to the huge range of organisational and project-specific variations on field work manuals, the most well known perhaps being the Museum of London's 'Archaeological Site Manual' (Spence [1990](#)).

The Matrix project consultations suggest that within separate archaeological organisations, practitioners agree and implement a shared data-gathering methodology for working on site. This is because the excavation process and practice for on-site activity by its nature involves a team of different excavators in fieldwork recording, and needs to be more standardised if the data records are to be consistent. Hence, pro-forma recording sheets or data-entry forms for database entry have been developed over many years of practice and are commonly (almost universally) used (Spence [1990](#)). Generally these pro-forma recording systems have been developed out of, or reflect the particular mode of, open area excavation commonly referred to as 'single context' stratigraphic excavation. This 'single context system' is based upon a methodology developed in the 1970s, and first implemented in British commercial archaeology by the Department of Urban Archaeology (DUA) of the Museum of London in 1977 (Hammer [2002](#), 640; Thorpe [2012](#), 38), and which has come to function as an informal standard practice in contract archaeology in the United Kingdom (Hodder [2005](#), 3).

Early on, in setting out the rationale for the design of the single context recording system, Westman and Shepherd noted that the approach prompts excavators 'to record their own wider interpretations on the same pro forma sheets as they record more factual observations, while scrupulously distinguishing these from each other', as they suggest that 'this system generates comprehensive records whose orderly form opens them easily to interrogation at any later time' (Westman and Shepherd [1992](#), 436). They go on to note that:

'a practical effect is that archaeologists need do on site only what they must do there and can defer until later other processes of analysis and interpretation, with an efficiency that has become an absolute necessity in the physical conditions of modern redevelopment and the modern building industry' (Westman and Shepherd [1992](#), 436).



This is interesting, as it suggests that there has been a disjunction between field recording and post-excavation activities since the literal invention of the modern industry standard for recording excavation data. It is also worth noting that a great deal of emphasis is placed by many practitioners on efficacy in recording and post-excavation analysis, which Watson ([2019](#), 1646) argues to this day fosters persistent 'formalised division of description and interpretation' in an effort to ensure 'that programme and budgetary constraints are met'. Much of this can be linked to the professionalisation of the discipline in the last half of the 20th century, and the standardisation of project management practice as a result of the introduction of 'developer-led' archaeology in the early 1990s (epitomized by central management guidance documents like MAP2 (English Heritage [1991](#)) and MoRPHE (Historic England [2015](#))). However, critiques of these systems often highlight the fact that this approach to analysis renders the excavators 'invisible' in the final interpretation and analysis of the material they excavated (Lucas [2001](#), 13).

Our own consultations support these arguments, as they suggest that post-excavation analysis practice still mostly happens off site. As such, the process for working on the post-ex of each site is usually carried out, or at least largely managed, by one individual and there is greater latitude for individual working practice to be used. Davies ([2017](#), 35) characterises current post-excavation processes as covering four broad stages of activity:

1. Checking the archaeological records and producing a possible interpretation.
2. Identifying the finds and producing specialist finds assessments reports.
3. Integrating the specialist finds assessments reports with the possible interpretation.
4. Full analysis and the completion of a synthesis report.

However, there is considerable latitude in the way these stages are implemented throughout the sector by different organisations. This was immediately apparent in the Matrix project consultations. A serious lack of standardised and interoperable digitally archived datasets resulting from commercial archaeological investigations has been identified. There is a need for further work to be undertaken with stakeholders across the sector, and particularly with the major contracting archaeology organisations, to develop shared good practice documentation for post-excavation and resulting archive material, which in turn should help improve the sharing and interoperability for reuse (FAIRness) of data deposited in archaeological archives.

Our consultation revealed that the by-products resulting from this analysis work are often quite individualistic and variable, dependent upon the scale of site involved and the software skills and availability of the individual project supervisors who get to 'write up' a site. This is especially compounded when digital data are supplied by specialists who are often external to the contracting archaeological organisation. Further work should be undertaken with stakeholders across the sector, and particularly the major contracting archaeology organisations, to develop shared good practice documentation for analysis, possibly in the form of an online handbook ([Recommendation 1.1](#)).



4.2 Common approaches and variations to grouping based upon stratigraphic complexity and methodological approaches

Despite the lack of consistency in post-excavation practice and outputs, there are a number of *de facto* standard commonalities in the way that practitioners manipulate and analyse complex stratigraphic data (specifically the grouping and phasing of stratigraphic contexts), even if procedures and standard practice continue to vary (often from individual to individual).

Shortly after the Harris matrix was conceived, two key approaches began to emerge for the grouping of stratigraphic contexts into rationalised higher order interpretative units (which commonly include: pits, walls, graves, etc.). Historically, these can be traced back to two schools of practice based respectively upon the Department of Urban Archaeology's 'single context recording', and the Central Excavation Unit's 'feature-group' approach respectively, both developed in the 1970s with the professionalisation of archaeology during this period (see Hammer [2002](#); Roskams [2001](#); Thorpe [2012](#)). The latter formed the basis for Carver's ([1979](#); [1990](#), 132) development of the 'feature sequence diagram' (Carver [1990](#), 132) that sought to incorporate higher order interpretative groups of strata *in the field*. Carver's approach allocates features which are grouped with their own numbering system and stand alone from the stratigraphic unit. The DUA approach essentially takes place post-excavation, clustering strings of related numbers within the matrix itself and does not stand alone from the main sequence but rather sits on top of it. The rationale for doing this is because the piecemeal archaeological interventions (separate small trenches, excavated at different times in the project lifecycle) and the disruption of modern truncations (building foundations, piles, services) meant that it was often difficult or impossible to correlate and group stratigraphic contexts in the field, without the holistic overview afforded when the archive is gathered together afterwards.

For a long time there was some debate over whether it is appropriate to perform this higher level grouping on-site, or as a part of the post-excavation process (Carver [1987](#); Hammer [2002](#); Roskams [2001](#) 244-46; Thorpe [2012](#), 36-40; Roskams [2013](#), 38-45). However what is clear from our consultations is that most archaeologists seem to be aware of conventional systems of higher order stratigraphic grouping associated with single context recording at least (in the tradition of the DUA) such as those outlined by Roskams ([2001](#), 257-61), even if they do not consistently deploy them on less complex sites.

Where a moderate to large amount of stratigraphy was encountered then, regardless of the relative complexity of the stratigraphy there was consensus in the consultations that the most common practice in analysis involved a succession of iterative steps of grouping the individual contexts identified and recorded on site into other interpretative, higher order, units based on related data from analysis of finds, samples and other excavated materials. In simple terms, a typical approach to stratigraphic analysis is to identify and group combinations of stratigraphic units that



go together to represent distinguishable activities (or related features?) at separate phases of time during the archaeological duration of the site.

Several consultees noted that the approach to grouping depended upon the perceived complexity of the stratigraphy encountered. In deeply stratified (urban) excavations, interpretative units known as sub-groups are ascribed as standard practice to identify and describe what are interpreted as separate single activities within the stratigraphic record that are most like to contain coherent dating evidence (e.g. a succession of floor layers within a room, or different infills of pits). Westman and Shepherd (1992, 439) note that 'contexts are usually placed in a subgroup if they are directly related to each other stratigraphically and if they are interpreted as representing a single phase of activity'. But in less well-stratified archaeology (e.g. open area settlement or shallow stratigraphic sequence in a rural setting) quite often there is limited complexity in the stratigraphic record and a single interpretative step of grouping stratigraphic units together to identify major features or structures is a more efficient first step. Only where there is more complex stratigraphy within an identified group might the use of further divisions of stratigraphy into sub-groups be needed (see [Figure 3a](#)).

In this way, the nature of the stratigraphic record can be seen as a by-product and to some extent a quantification or measure of the scale of archaeological complexity on a site. This is not to suggest that less stratigraphy necessarily makes for an easier interpretation, but it does suggest that the amount of associated data available for archive and reuse on less stratified sites is likely to be smaller. It means that on sites with less stratigraphic complexity, there may be less need for archaeologists to produce a stratigraphic matrix diagram in order to understand the stratigraphic relationships and interpret the phasing of the main features represented.

4.3 Common approaches and variations to phasing

Phasing is another higher order interpretative process commonly undertaken as part the stratigraphic analysis. Phases form one of the key outputs of many excavations, often serving to structure the reports and narratives associated with this sort of fieldwork and very clearly evidenced through the production of 'phased plans'. They are generally distinct from 'periods' in that they are local to the site, even if they reflect broader regional periodisation in their definition. Conventionally, they are defined into existence through a process of 'detailed examination of stratigraphic relationships and their formation processes, usually in relation to the material culture and environmental evidence which they contextualise' (Taylor 2016, 179), allowing 'strands' of the matrix to be 'drawn up and down' (both conceptually and on paper) until they are 'in phase' and therefore considered to share the same band of temporality' (Taylor 2016; see also Roskams 2001; Hammer 2002; Farid 2014, 91-92).

Phasing is always an 'interpretative negotiation' and deciding which units belong to which phase is a matter of reasoning on the part of the archaeological stratigrapher. It is perhaps worth noting that there are a number of possible approaches to phasing



stratigraphy (see Roskams [2001](#); Lucas [2001](#); Pearson and Williams [1993](#)). Crucially, all these approaches share a common purpose: to divide the vertical sequence horizontally in order to group stratigraphic units and groups into bands that are related spatiotemporally.

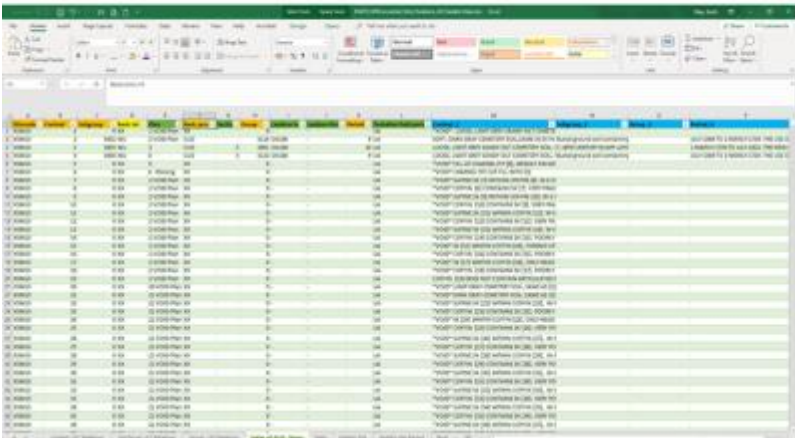
4.4 Standard Outputs - a consistent phasing index?

In the search for good practice approaches to the archiving of stratigraphic analysis and phasing data, the closest to a common reusable format that might be used as a common basis for data sharing was encountered in the digital archives deposited by MOLA. The MOLA IAA.CSV file contains the key fields needed to 're-construct' (or re-interpret) the main steps and interpretative decisions taken by the stratigraphic analyst in the grouping and phasing of the data. From discussion with MOLA archaeologists, this was identified as an output from the MOLA Oracle database, rather than a file that was worked on by the archaeologists during post-excavation, and that the IAA.CSV (Figure 5) had only more recently become a standardised output for the digital archiving of MOLA data.

'it's [the IAA.CSV] an output from our Oracle database, created for the archive... during post-excavation we wouldn't work with this file, but directly with the database and various Excel reports generated from it which are linked with our Intrasite GIS. Until recently, our Oracle database has not contained stratigraphic relationships, which we usually enter into Bonn (or sometimes ArcEd). For deeply-stratified urban sites, the rigid and unwieldy nature of these stratigraphic matrices and our reliance on them for post-excavation work is a big challenge for us.' Louise Fowler, Post-Excavation Manager, MOLA (Museum of London Archaeology).

It seems that this is becoming more typical practice within many larger archaeological organisations, where they work on data within their own database systems and then the data are exported to archive, not necessarily by the post-excavation archaeologists but possibly by a database manager or the archivist, and only at the end of a project when it is considered 'archive ready'.

Phasing Index data - e.g. Index of Archaeological Association (IAA.xlsx)
 – Museum of London Archaeology (MOLA) Digital Archive File
 XSM10 Crossrail archive downloaded from ADS – IAA file used for recent MOLA archives on



IAA Queries

1. MOLA only?
2. Do other people use similar "Phasing Indexes"?

Is use of IAA.csv file –

1. Helping re-use?
2. Is it used consistently as a digital standard across sites?

3. Note - **no primary Stratigraphic relationships** in this index. Product of MOLA archive process.



Figure 5: Phasing Index example - Index of Archaeological Association. Museum of London Archaeology (MOLA) Digital Archive File (XSM10 – IAA.csv downloaded from ADS and converted to IAA.xlsx)

Adoption of a more standardised approach to the outputs of post-excavation analysis would greatly help with the deposition of more easily reusable data in digital archives. It is recommended that the key data held in the MOLA Index of Archaeological Association (IAA) is used as a starting point for such a baseline of analysis data that could be supplemented with further data indices, depending upon the scale and complexity of the archaeological site stratigraphy (see Figure 5). To follow this through, it is proposed that a specific working group on stratigraphic standards be created as part of the work to investigate the best practice documentation for post-excavation analysis practice. Research funds have been obtained from AHRC for exploring the set-up of this group in the UK and further more sustainable funding can be investigated by that group to build a pan-global International Convention. An online forum for the former 'Interpreting Stratigraphy' mini-conference is one possible route for taking such initiatives forward, including a related online Community of Practice (CoP). Such a CoP could re-visit, re-affirm and refresh as necessary the existing 'Principles of Stratigraphy' ([Harris 1989](#)) along with developing a federated online system, using the online handbook and tools, for promoting best practice and minimum requirements for phasing and stratigraphic analysis procedures across the UK and internationally as part of an International Convention on Archaeological Stratigraphic and Chronological Methods and Data ([Recommendation 4.3](#)).

4.5 Lack of post excavation analysis 'manuals'

Within individual archaeological organisations, practitioners generally agree and implement shared data gathering methods for working on site, because the excavation process and practice for on-site activity by its nature needs to be supervised and regulated to achieve standardised, relatively consistent and interchangeable records of the individual stratigraphic units excavated. Hence pro-forma recording sheets or data-entry forms for database entry are commonly used (Figure 6). This is reflected in the huge range of recording manuals and methodologies that are available, often as unpublished project or organisational documentation, exemplified by the MoLA Archaeological Site Manual (Spence [1990](#); but also see Joukowsky [1980](#) or the range of recording systems discussed in Pavel [2010](#) and Masur *et al.* [2013](#)).



Figure 6: Examples of commonly used excavation recording sheets

By contrast, in off-site post-excavation practice, there is very little comparative documentation (with the notable exception of Hammer [2002](#)). Our consultation further suggested that during post-excavation, while the outcomes of the work need to be agreed across a team, the process for working on the post-excavation of each site is usually carried out, or at least largely managed, by one individual. Although one person may lead the process and take responsibility for 'writing up the site', usually no single person carries out all the processes involved in that analysis, so data management practices may be fragmented, or more likely rather siloed. This may be an issue for data reuse if we are trying to match the outputs of several peoples' work to the needs of a single end reuser, instead of considering packages of data for multiple, but known and sign-posted, reuse scenarios.

Our consultations also noted that variations in staff post-excavation practices often depend upon the archaeological experience and individual background and writing styles of staff. This in turn validates further variation (and idiosyncrasies) in the methods used to analyse the site, perhaps accounting for the inconsistencies in the nature of the data outputs. This general trend is highlighted by Davies in his [2017](#) study:

'Post-excavation projects therefore gradually changed as individual archaeological organisations developed their own internal processes and procedures to cope with the unpredictable commercial environment and the needs of both specific developers and specific post-excavation projects (Participant Interview 13: 55.00). In practice this involved a move towards a more ad hoc form of post-excavation process, so although the basic structure of a post-excavation project remained roughly the same what occurred within each stage would depend upon individual Project Officers, individual Project Managers, individual circumstances, specific deadlines and the knowledge that most post-excavation projects could also come to a sudden halt at any point for any one of a number of different reasons' (Davies [2017](#), 179).

The lack of post-excavation documentation, even for the more commonly occurring activities such as stratigraphic analysis, came as something of a surprise and leads to our [Recommendation 1.3](#).



Interestingly, from the UK perspective, there is no formal definition of post-excavation work (beyond Hammer's [2002](#) attempt, in her - now defunct - online *Post-Excavation Manual*) until after the introduction of PPG16 around 1992, when it became embodied through the MAP2 guidelines and when a number of post-excavation projects were instigated to address a need for synthesis of large 'backlog' excavation archives in several major urban centres. All the early players (Wheeler, Kenyon, Woolley, Piggott) make no mention of 'post-ex' as part of the process, preferring instead to talk in general terms about how archaeologists need to deal with finds and then publishing. The term post-excavation is most often used with reference to project management stages and given as a job title e.g. 'Post Excavation Manager'. As such, the post-excavation stage is probably derived from project management terminology (English Heritage [1991](#) MAP2; Historic England MoRPHE [2015](#)) as a broad term for a whole range of possible, but not necessarily required, analytical steps taken in order to move a project's outcomes from the initial excavation records created on a site, through the various analytical processes enabling interpretation and synthesis of different data discovered during excavation, to the eventual reporting and publication of the overall site narrative and interpretative phasing.

4.6 Case study of XSM10 data – a best practice example

Even with the assistance of ADS staff, there were very few digital archives available for reuse with the full range of primary stratigraphic data from excavation along with grouping and phasing from analysis. The [XSM10 Crossrail archive](#) was the only site that had primary stratigraphic data (i.e. above/below/equals relationships matrix data) and a moderately complex range of different stratigraphic phases, saved in the Stratify [LST format](#) (which still required a degree of processing to convert to an equivalent CSV format, using an old copy of Stratify running on Windows 7). The primary stratigraphic data was archived with a file containing the [Grouping and Phasing data](#) that could be related to the primary stratigraphic data. The XSM10 site was one of 30 sites excavated by MOLA and Oxford Archaeology undertaken for the 'Crossrail' railway/underground project in London. For the record, [XRW10 Limmo Peninsula](#), also deposited by MOLA, contained the necessary datasets, but this predominantly post-medieval site did not have such a complex stratigraphic depth of multi-phase stratigraphic sequences that were most suitable for the purposes of our stratigraphic analysis and software testing.

The XSM10 fieldwork comprised a series of watching briefs, evaluations and excavations, undertaken between 2011 and 2015. The site was excavated to a depth of 25m below street level for what became the Elizabeth Line underground station and platforms at Liverpool Street station in London. The site lay 120m north of the Roman London town boundary on the east bank of a tributary of the Walbrook stream. The earliest Roman activity focused on draining the site sufficiently to allow burial and road building in the area. Extensive remains of an early 2nd- to 3rd-century AD west–east metalled road were traced across the site, along with several phases of roadside ditches. Further reclamation of the marshy ground took place in the Medieval period. The burial ground documented as the 'New Churchyard' (also



known variously as the Old Bedlam or Bethlem burial/burying ground/place), was in use between 1569–1739. The archaeological investigations involved the excavation of c.3750 skeletons, as well as boundary walls and burial structures associated with the burial ground.

The stratigraphic data in the ADS archive was split between the post-medieval cemetery data and the pre-medieval, late Iron-Age and Roman data, mostly associated with the Roman roadside activity. For the purposes of our project, we did not use the post-medieval cemetery data because it was predominantly a mass of grave cuts but without the depth of complexity of stratigraphy to use for matrix and phasing test purposes.

5. Process Modelling of Post-Excavation and Stratigraphic Analysis Processes

The Matrix project always planned to model the post-excavation analysis process by collating and synthesising various organisations' documentation and guidelines relating to post-excavation management and 'produce a high-level process model and diagram for stratigraphic data analysis'. But because the project found such a lack of up-to-date written disciplinary guidance for the post-excavation analysis process (or indeed even a clear definition of the processes involved), it made the process modelling a more exploratory exercise than might have been expected. Nevertheless the 'To Be' process model was produced by the project (Figure 8) and shared with consultees and workshop participants in order to understand and inform design decisions for the prototype tool. This process modelling work ultimately identified five common steps in the process of stratigraphic analysis, as follows:

1. Excavation data checking, including any existing matrix diagrams, and construction of a consistent stratigraphic matrix diagram from the primary excavation data.
2. Sub-group data creation and interpretation (when complexity of stratigraphy required).
3. Group data creation, interpretation and construction of a group matrix.
4. Incorporate chronological data (from material culture assemblages, samples and spot dates).
5. Phase (periodise) the stratigraphic data and create a phasing index and phased matrix diagram. This can include creating/checking phase plans (often in GIS) for illustrative purposes.

An initial layout for the stratigraphic matrix is based upon excavation data relationships. During the analysis work the Grouping and Phasing stages can be characterised as an iterative process, whereby stratigraphic units are clustered together as entities based upon function and activities. These are adjusted and updated according to dating evidence from finds and samples, understanding of the site formation processes, and with cross-reference to spatial information often from the drawn records (plan and sections), which may or may not be digitised into a site



GIS. An overview of this process is shown in Figure 7 and illustrated in more detail in the swim-lane process modelling diagram (Figure 8).



Figure 7: Overview of work flow processes for stratigraphic analysis of complex stratigraphy (N.B. In steps 2 and 3 the degree of grouping or sub-grouping used is largely dependent upon the nature and complexity of the stratigraphy encountered on each independent site)

5.1 Data fields to be used in stratigraphic analysis and inform prototype software

As part of modelling the data required for stratigraphic analysis, a list of minimum required data fields was derived, based on feedback from the consultations with archaeologists and the experience of the project team. This data field list was incorporated into a model (Figure 8) and shows which data elements (data fields derived from archaeological datasets or entered as part of analysis) were used to inform the design of the prototype stratigraphic analysis software tool.

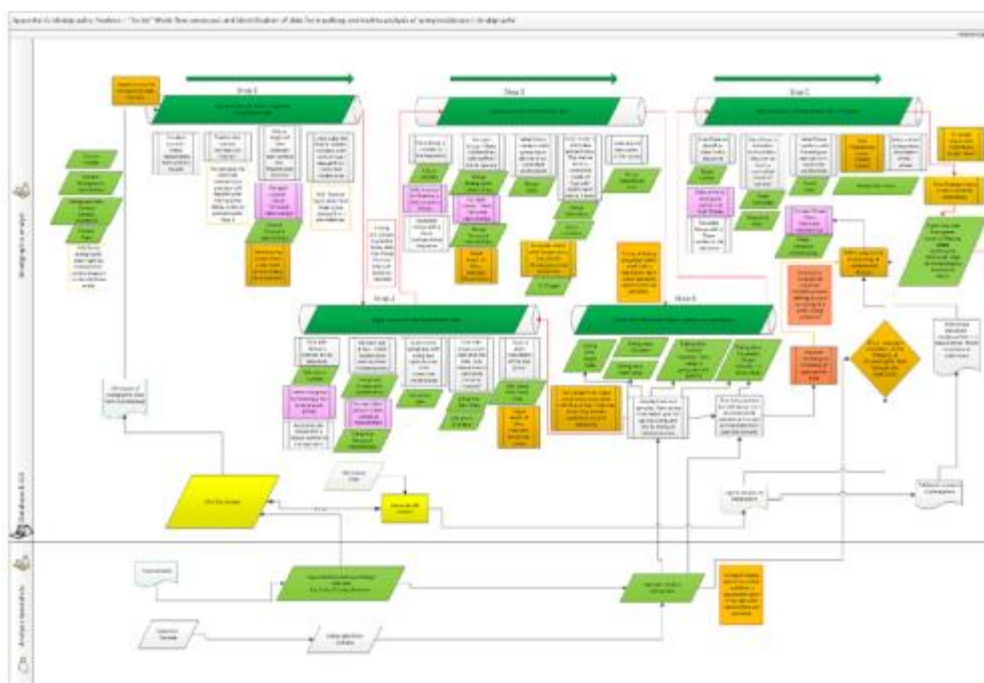


Figure 8: Process modelling diagram ('To Be' model) developed to inform development of the prototype software

The initial fields used in Step 1 are derived from fieldwork records of basic stratigraphic relationships between individual context records. Further records are then entered as part of the analysis process (Figure 8, steps 1.1 onwards). This would include data derived from grouping and sub-grouping analysis activities, along with dating evidence from specialists. This modelling was used to inform the specification of the initial software requirements that led to the development of the Phaser prototype software ([Section 7](#)).

For initial prototype testing purposes, the test datasets, sourced from archive records, needed to have the 'analysis' data available from a completed analysis stage in the digital archive records so that they could be entered to test the software. In practice, when it came to testing the software, finding digitally archived datasets from completed projects that had comprehensive and readily reusable data covering all the five different steps identified above proved more difficult than initially expected.

5.2 How excavation practices determine process - deep 'urban' stratigraphy vs rural and infrastructure

Deeper stratified archaeological deposits, and therefore generally more complex stratigraphic sequences, are most often associated with the excavation of sites in 'urban locations' (seen explicitly in the MOLA guidance and in Hammer [2002](#)) - although an urban site may be so damaged by modern basements that remains are heavily truncated, and therefore stratigraphically unconnected, with features cut



straight into 'natural'. By their very nature, urban excavations generally contain more concentrated build-up and sequences of stratigraphic deposits, contrasting with excavation on rural or non-urban sites, which do not tend to produce such deep sequences. Consultation responses showed that even where individual context stratigraphic relationships were recorded in rural sites, there was no need to create Harris matrix diagrams in order to understand the overall stratigraphy. A matrix diagram was considered extra work and an unnecessary overhead when up against tight project management deadlines.

Further expansion on this idea also led to consideration of 'infrastructure' sites, as a sub-type of the 'rural site'. This included projects on road, motorway or by-pass schemes or railway lines such as the Channel Tunnel rail link, HS2, and Heathrow Terminal 5. Here too it was felt that evaluation trenches all with fewer than ten contexts did not require (or merit) the time that would be needed to draw out a Harris matrix diagram. It may be that the difference here is more about the funding trajectory than the approach to stratigraphic methods. On larger infrastructure projects, the amount of stratigraphy dealt with may vary a lot along the whole route of a road or rail corridor as different sites are excavated with differing degrees of stratigraphic preservation encountered (Figure 9).

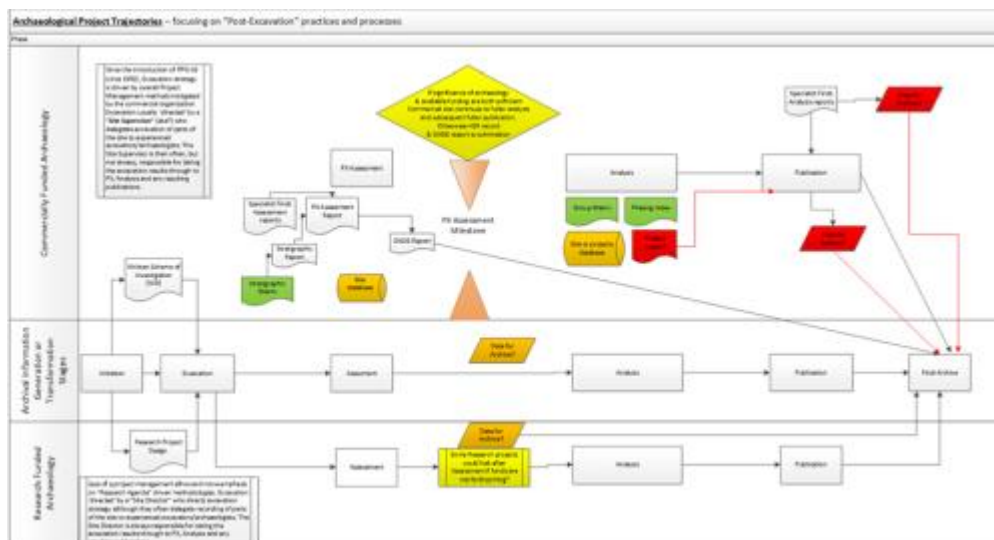


Figure 9: Different archaeological trajectories followed by projects depending on funding sources, resulting in different digital outputs (to archive) at different stages in a project's process

In consultations, this led to suggestions that there was a need to identify different methodological approaches for recording stratigraphic data. For instance, the sub-grouping of uncomplex stratigraphy might be an unnecessary overhead when using GIS to manage the spatial data. This may in turn help sign-post the type of stratigraphic record that would be expected to result from excavations following certain processes, something that could be represented as a decision tree e.g.

1. Scale and nature of funding
2. Type of site – depth of stratigraphic sequence
 - i. Are there multiple intercutting threads of stratigraphy
3. Scale of archaeology encountered



- i. Is the site part of a larger scheme of investigation (e.g. infrastructure project)?
- 4. Degree of dating evidence encountered in the stratigraphic units
 - i. Is there any scientific dating evidence?
 - ii. Has any sampling for dating purposes been undertaken?

5.3 Dating evidence – five main formats of date evidence?

In gathering examples of dating evidence to enable the prototype construction and testing, the project team identified and characterised five main forms for archaeological dating evidence. These five are shown in Table 2 with some of the examples used for the prototype development.

Table 2: Five common forms of archaeological dating evidence

Dating form	Note	Example data
Single date	Single dates are actually relatively rare in archaeology	203 AD after c. 1760
Date range	Most dates will be in this form	117-138 (coin)
Period date	Could be as broad as 'Roman', but could potentially carry an associated date range too	Early Roman SABA - Reece Period 6
Probability date range	Dating data: Probability Range (Usually to a year with an associated \pm error value)	c. AD 99-134 Dendro 95% confidence
Duration	<p>This is most likely a date range but without necessarily fixed start or end dates. Examples will most likely occur during grouping and phasing</p> <p>. This may be especially relevant to:</p> <ul style="list-style-type: none"> Construction (e.g. <1yr) Use – duration (>1yr) Disuse (<1yr) 	The building could only have been in <i>use</i> for at most 50 years (i.e. duration = min 0 – max 50)

An additional practical factor in the choice of approach taken in post-excavation analysis may also be a matter of when the 'spot dates' (e.g. coins or pot dates) for assemblages become available. Larger organisations may have some specialists who can do certain spot-dates in-house but others may not.



Our prototype has been designed to accommodate some of the most commonly encountered differences in approach to recording dating evidence. These differences are reflected in Roskams' recent article on variations in the post-excavation practices carried out on the Heslington excavation records.

'Although superficially similar, the distinct approaches of each organisation express fundamentally different approaches to integrating finds dating and stratigraphy, from forming groups independent of recovered assemblages at one end of the spectrum to using detailed finds dating to create initial groups at the other.' Roskams [2020](#)

We have applied the Allen operators to the date ranges on the objects that are associated with the stratigraphic units.

Many archaeological datasets express dating evidence most commonly in the form of date ranges on finds, although there can still be some variation in how date ranges are expressed in data fields (Binding [2010](#)), including the degree of certainty in such dates, e.g. as shown in Table 3

Table 3: Examples of date ranges encountered for different types of finds objects

coins	Earliest date = 50 BCE – latest date = 200 BCE
pottery	Earliest date – latest date
glass vessels	Earliest date – latest date
architectural items	Second half 12th century, i.e. 1150-1200
various objects or features	'Roman' i.e. 43-410 (UK)

5.4 Use of periods in prototype software development

Archaeological periods have a tendency to be indeterminate or could be described as 'fuzzy' chunks of spacetime, i.e. both the spatial and temporal boundaries of any identified and named period could vary in space over an indeterminate time, e.g. the Roman period (see also the [CIDOC CRM](#) overview of spatiotemporal modelling summarised below). In Phaser, we have therefore chosen to enable users to reference the labels (names) of 'Periods' according to the geolocation from which their site data derive, i.e. we use the [Perio.do](#) Linked Open Data (LOD) terminologies that are appropriate to the geolocation of the origin of the dataset.



The [CRMgeo](#) spatiotemporal modelling standard (Hiebel et al. [2015](#)) introduces the notion of 'Spacetime Volumes' for semantic modelling of GIS data:

An analysis of Open Geospatial Consortium (OGC) concepts and CIDOC CRM concepts revealed that in order to integrate these models the explicit differentiation between spatiotemporal properties of real world phenomena (phenomenal Spacetime Volumes) and human assumptions about these (declarative Spacetime Volumes) is required. Phenomenal Spacetime Volumes derive their identity from phenomena defined as classes in the CIDOC CRM family of models like events or persistent items and are fuzzy due to the nature of the phenomenon...The differentiation between phenomenal and declarative is applied to CIDOC CRM Places and Time-Spans as well, due to their definition as spatial and temporal projections of a Spacetime Volume' (Hiebel *et al.* [2015](#)).

The 'Fuzzy' spatiotemporal operators used in the core CIDOC CRM are intended to be used for phenomena (which could in some cases be periods) with unknown start and end dates (Papadakis *et al.* [2014](#)). Otherwise where start and end dates are available, the Allen operators can be used, as we have implemented in Phaser when using date ranges from dating evidence such as archaeological finds like pottery and coins.

As Papadakis puts it, 'Information about the relevant topology of precise time intervals can be stated using Allen's operators. In cases of imprecise information, the temporal association of fuzzy intervals can be approximated by a set of Allen's operators that hold between the possible endpoints of the imprecise intervals' (Papadakis *et al.* [2014](#), section 3.3).

6. Impacts of Digital Technologies upon Stratigraphic Analysis and Matrix Diagrams

Recent decades have seen increasing sophistication of software and computing technology, which has allowed the collation and storage of more archaeological recording and spatial data than ever before. This section will briefly consider the use of computers to manage spatial and stratigraphic data in archaeology as context for the design decisions that informed the prototype development discussed in [Section 7](#).

6.1 The use of software for stratigraphic analysis and matrix diagrams

It is worth summarising some of the main previous or existing uses of software to see how stratigraphic data are being analysed, and to help understand why there is such variability in what ends up as outputs in the digital archive records of stratigraphic, phasing and matrix information.



At the first project workshop (15 July 2021) in response to the question 'What other Matrix software applications are used across the UK?', the following software packages were identified as the most likely to have been used by archaeological stratigraphic practitioners.

- **ArchEd** (MoLA, Wessex) - Analytical
- **Bonn** (MoLA) - Analytical
- **ARK Matrix** (L-P) - Analytical? (or archive only?)
- **Stratify** (Wessex) - Analytical
- **Intrasis 3** (HE) - Analytical
- **GSYS Matrix Manager** (LRC) - Visual and Analytical
- **IADB** (YAT) - Visual and Analytical
- **ASEbase Bespoke Software** (ASE) - Visual and Analytical
- **Harris Composer** - Analytical
- **AutoCAD** (MoLA, Northampton?, Wessex?) - Visual
- **Excel** (PCA, RRA, and others) - Visual

Probably the most widely mentioned matrix construction 'solution' encountered during our enquiries was a version of MS Excel to set out a matrix diagram, without any database connectivity to enable data cross-referencing or analysis. Sometimes these diagrams (drawings) are archived as .XLS or PDF files, but very few (if any) examples of completed diagrams were encountered in the ADS archives we investigated.

In addition, the following were also contacted directly for more information during the course of the project. Key members of these initiatives took part in the Matrix project workshops, symposia and discussions. In addition, specific staff at L-P Archaeology and ASE were contacted directly for more information during the project. John Layt and Guy Hopkinson took part in the Matrix project workshops and discussions.

- [ARK Matrix](#) - developed by John Layt at L-P Archaeology as part of the ARK software
- [ASEbase](#) - developed by Guy Hopkinson (UCL-ASE) sophisticated matrix analysis tool incorporating GIS functionality (see Figure 10)

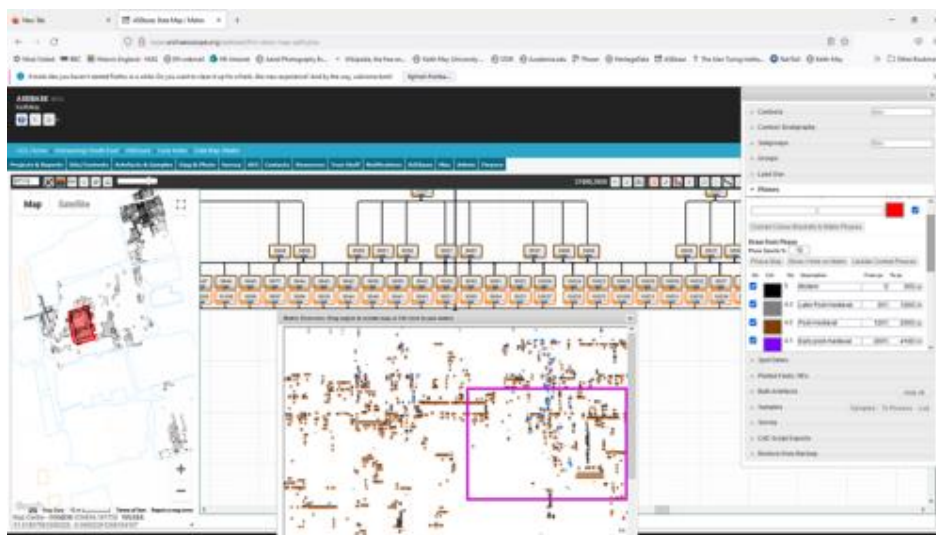


Figure 10: Screen shot of ASEbase 'Data Map/Matrix' analysis screen. Courtesy of Guy Hopkinson, ASE



The ASEbase system has integrated a matrix data entry and analysis tool with a GIS (map) viewer enabling the stratigraphic analyst to browse and navigate between and around stratigraphic context level data and the digital site plans. In effect this enables the digital overlaying of stratigraphic plans to check the stratigraphic relationships in both the database and drawn records. This is similar in approach to the GSYS Matrix Manager tool developed by The Landscape Research Centre (May 2020, section 6.2). It also enables plans of groups and phases to be identified and output for checking and publication where suitable. Because of awareness of such GIS-linked matrix software development, a decision was made that the Matrix project research and, more specifically, the Phaser tool would focus on functionality to support analysis of the temporal and spatiotemporal relationships and did not attempt to integrate with any GIS analysis functionality that could be undertaken in tools such as ASEbase.

Even more recently we have become aware of further work on 'Harris Matrix Composer+' (Doneus *et al.* 2022) <https://www.youtube.com/watch?v=rXaruMkg81Y&t=8s>. However, no official updated release has appeared on the HM Composer website as yet at <https://harrismatrixcomposer.com/#/>.

What emerged from the consultations with contracting practitioners was a range of similar issues around use of matrix software that seemed, rather surprisingly, to reflect limited consistency in the approaches taken to post-excavation analysis even within the same organisations, let alone across separate organisations. A typical experience is summarised in the following extract from an online post:

A few years back, I had used [Stratify](#)...I had found it relatively easy to use, and free — which is always a plus. But once the matrix gets very complicated, there is no way of taking control of how the matrix is presented, and it ended up — in my view — rather un-aesthetic, with interconnecting lines all over the places. Moreover, the software has not been updated in 10 years, now, and the code remains closed, so no one else can pick up the tab to update and improve it. I've also had issues in trying to use it on Windows 10... I eventually settled on something else'. <https://jmbriiffa.wordpress.com/2020/08/22/free-software-for-a-harris-matrix/>

In fact, practitioners undertaking post-excavation stratigraphic analysis were not necessarily expected to use stratigraphic software at all, and the choice of whether to use dedicated software tended to be left to the individual to make and, to a certain degree, was dependent upon the archaeological experience and especially their IT experience and skills. A related issue that compounded this variability in methods chosen was the availability of any of the software programs. In many cases, spreadsheet software (Excel) was used for 'working' matrix diagrams because it was the only software commonly available within the organisation's software suite.

This issue of software accessibility, even within the same organisation, was noted a number of times where stratigraphic analysis was carried out in spreadsheets as few commercial operators seem to be using any dedicated matrix software, although several highly detailed 'hybrid' solutions were noted, e.g. using Excel for the sub-grouping, grouping, correlation and then using matrix building software like Stratify to produce interim matrix diagrams for analysis, and then using drawing packages (e.g. [Diagrams](#)) for the final 'publication ready' Harris matrix diagrams once they were



fully analysed. A particular example cited by some consultees was use of the [Bonn matrix package](#) to check 'logical integrity' of the initial stratigraphic relations (Figure 11).

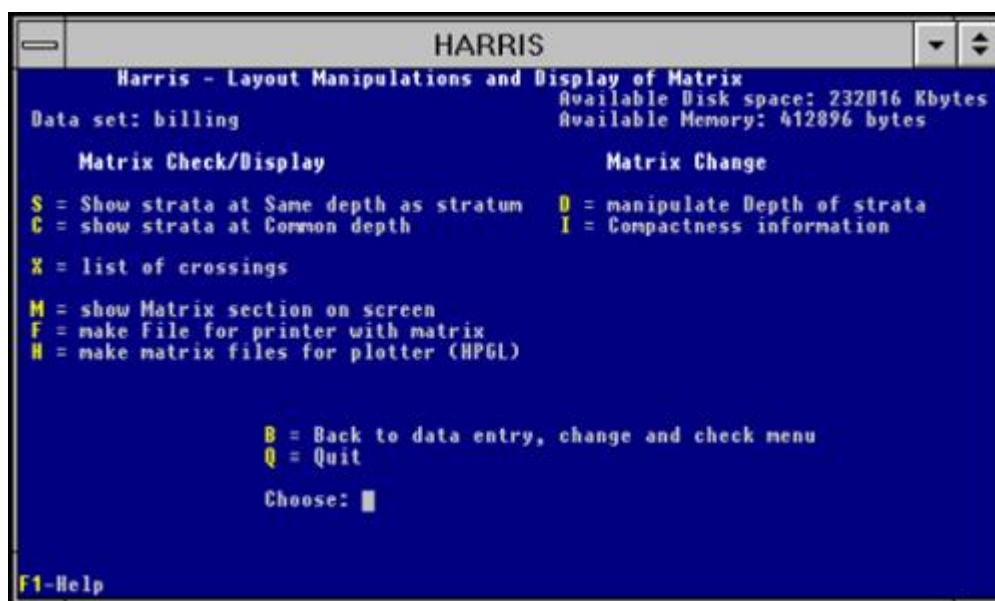


Figure 11: Screen shot of Bonn Matrix software (version 5.43)

This stage was followed by export of a version to ArchEd (or Stratify) for visualisation of the diagram, then Excel was used for sub-group and group relationship construction and analysis. Each of these activities would have separate digital outputs but no indication of how they might be deposited or accessed via an archive.

Inevitably financial cost was an issue raised in a number of cases. Any cost associated with software meant it was less likely to be used. This was raised most often by commercial archaeologists in respect to Harris Matrix Composer.

6.2 GIS and spatial records

For archaeologists, the digitisation of the spatial record has a number of obvious advantages over standard two-dimensional paper maps and plans, the most obvious being related to data manipulation, since data can be edited, duplicated and printed cheaply and efficiently. Producing a 'map-series' to display diachronic spatial change (phases) or distribution of material culture is relatively straightforward using digital methods.

The most common tools available to the archaeologist for this purpose fall broadly into two types, Geographic Information Systems (GIS) and Computer-Aided Design (CAD). GIS were already being developed by the late 1980s; however, they have only been sufficiently affordable or accessible enough to be regarded as a commonplace disciplinary tool over the last twenty years. Prior to this more widespread use of GIS, the digitising process usually involved the manipulation of raw digitised spatial data inside a CAD software package. In essence, the layer functionality of most vector-based CAD software allows for the straightforward



overlaying of archaeological features, structures or even stratigraphic units, which makes it a particularly elegant solution for manipulating and visualising single context excavation data (Wright [2011](#), 134).

This type of data can be used to create sophisticated 3D vector models of spatial data and can also very effectively plot distribution patterns within those models. Historically, however, within archaeology at least, the uses of CAD beyond the level of spatial modelling have tended to be fairly limited because this kind of software was not initially intended to record further attributes about the vectors it stored. From a stratigraphic perspective, if one is simply using CAD to digitise plans, the software effectively acts as a more efficient way of overlaying plans while the actual analysis is still done by the archaeologist, or perhaps separately in a third-party software. This might be seen as an advantage since the archaeologist is not too detached from the interpretative process. Indeed, there have been clear advantages in the use of CAD packages in terms of quickly drawing together atomised contexts as multi-context plans (Alvey [1993](#)), and breaking down 'the traditional barriers between excavation and post-ex' (Wright [2011](#), 134; see also Lock [2003](#), 105-6).

Nevertheless, GIS has a distinct advantage over CAD because it offers the data structure required to make more meaningful spatial analyses. As a fully integrated spatial database with a spatial graphical front-end, its users tend to have a feature-based perspective on their data, and certainly GIS allows for unparalleled querying and for the semi-automated manipulation and filtering of spatial data. Recently off-the-shelf GIS packages have begun to address longstanding critiques relating to their capacity to represent space beyond 2D (or 2.5D; see Conolly and Lake [2006](#), 38-39, or Harris and Lock [1996](#), 309). The 3D capacity of modern GIS is improving rapidly and becoming more routinely deployed within archaeology for the recording and analysis of stratigraphic deposits (Dell'Unto and Landeschi [2022](#)). However, issues remain regarding the potential for any relational database management system to handle temporal data, which presents challenges for the sophisticated modelling of spatiotemporality within GIS (see discussion in Taylor [2020](#)).

Ultimately, although these technologies have had an impact on the way we manage and deal with the spatial record, their inability to manage temporal data or facilitate stratigraphic analysis means that despite the increasing adoption of a digital approach, the underlying practice of creating a Harris Matrix has not really changed much since 2001 when Roskams argued for more consistency in post-excavation analysis:

'Turning next to data manipulation after excavation, there is a great need to sort out the concepts used in stratigraphic analysis ... to match the systematization which has been developed in the production of the site record.' Roskams [2001](#), 278-79).

6.3 Can we move towards a FAIR and 'frictionless' data package for archaeological stratigraphic data?



The work of the Matrix project has highlighted the need to take a new approach to heritage data characterisation and data packaging in archives and how such an approach could help in making stratigraphic, and other associated, data more reusable and interoperable across different site records. The project has attempted to identify a coherent and 'frictionless' (<https://specs.frictionlessdata.io/>) data package for stratigraphic and chronological data that would enable finding and reuse ([Recommendation 5.1](#)). One possible avenue to explore could be a 'FAIR Cookbook for Heritage Data' (or at least for Archaeological Stratigraphic Data) along similar lines to the online [FAIR Cookbook](#) for the Life Sciences ([Recommendation 5.2](#)).

As a starter for such a data reuse package (or 'recipe' in a FAIR Heritage Data cookbook), we would propose that the 'stratigraphic package' would need to contain a minimum of:

- Site record of stratigraphic relationships ([example](#))
- Grouping and sub-grouping if undertaken ([example](#))
- Phasing carried out ([example](#))
- A matrix diagram, ideally in a digital format but at least as a PDF to refer to - XSM10 example would be derived from the [Stratify file.LST](#)
- *In addition, if used*
Periods (if incorporated) ([example](#))

The MOLA Index of Archaeological Association file ([IAA.csv](#)) covers a fair amount of the data required for this practice (grouping and phasing) but for stratigraphic analysis and reuse, people would also need the 'raw' stratigraphic relationships (i.e. the excavation stratigraphy) and a diagram, as well as more consistent recording of the limits of excavation and the interfaces with natural stratigraphy. In the case of the XSM10 archive, some of this information only came from resurrecting the Stratify (.LST) files for the Context level data.

6.4 Data fields most commonly used in stratigraphic analysis - minimum requirements)

As part of modelling the data required for stratigraphic analysis, a list of minimum required data fields was derived to inform the development of the prototype matrix analysis tool. These are set out in Table 4 and are intended to highlight which data elements from excavation records are needed for a stratigraphic analysis tool.

The approach to this may vary slightly depending on whether the archaeologist is entering data as a new record of a single stratigraphic unit, or whether they want to import a pre-existing set of stratigraphic data records from a digital file format (e.g. CSV). Either way, the initial fields used (Step 1 illustrated in [Figure 8](#)) are expected to be derived from fieldwork records of basic stratigraphic relationships between individual context records. These data can be directly imported from a site database or alternatively can be entered manually from data sources such as context sheets. To save time during the Matrix project, the majority of data was at least partly imported from existing datasets.



The prototype tool imports the initial five minimum required fields as shown in Table 4.

Table 4: Main (minimum) fields recommended for import to create a stratigraphic relationship in Phaser

Field	Note	Example data
Context number	The primary context. Derived from fieldwork records	123
Stratigraphic relationship	Derived from fieldwork records	Above/Below Equal
Related context number	Any related stratigraphic unit. Derived from fieldwork records	321
Context type	Derived from fieldwork records Some might match to approved LOD vocabulary such as http://purl.org/heritagedata/schemes/eh_tmt2	<u>Layer</u> , deposit, cut, fill
Sitecode	To keep track of datasets from different sites	XSM10

Once the stratigraphic record data has been entered in Step 1 (see [Figure 7](#)), then further data can be entered and subsequent data generated as part of the analysis process. This would include data derived from analysis of grouping and sub-grouping and phasing activities, along with dating evidence. For initial prototype testing purposes, the test datasets, sourced from archive records, already had these 'analysis' data available from completed analysis that could be entered from the archive records.

7. Phaser Prototype Software Development

The Phaser prototype enables creation, editing, analysis, import and export of archaeological stratigraphy data. The development work took place in two stages. The first stage produced an initial working prototype and the second stage made revisions and additions incorporating feedback from user workshops and usage in the interim period.



The prototype we have produced is an open-source, responsive, single page application written using the Vue reactive framework, with no server-side dependencies. This allows the [application](#) to run on a wide variety of devices using a modern web browser, with no additional installation or configuration required on the part of the user. Hosting of the application and its [associated documentation](#) is via GitHub, facilitating the ongoing availability of a documented working deployment of the prototype to persist beyond the official end of the funded project.

Access the [Phaser tool](#)

7.1 Data Model

Traditional stratigraphic matrices can be implemented as directed graph structures - the archaeological contexts being the vertices (nodes) of the graph, and the stratigraphic relationships (above, below, equals) being typed directional arcs (edges) between nodes, with a cardinality of many-to-many (Figure 12).

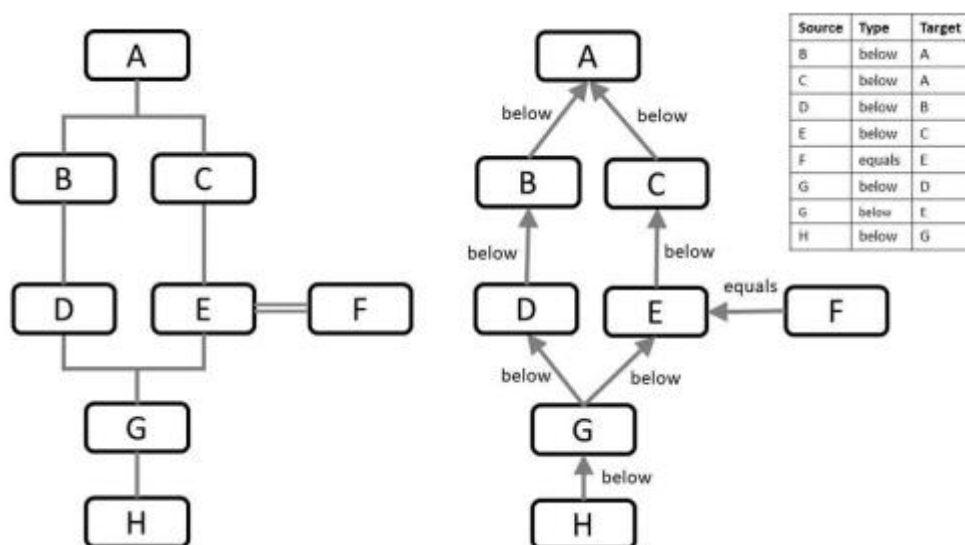


Figure 12: Context stratigraphy as a directed graph structure

Although the Phaser tool performs import/export of delimited tabular data, the internal data structure for the matrix diagram is a directed graph. Context nodes have coordinates to enable their positions in the diagram to be stored, plus specialised properties such as site code, identifier, label, description etc.

An additional requirement was the grouping of elements using a containment hierarchy, as shown in Figure 13. Note that there is some flexibility in the containment hierarchy in that contexts may be contained directly within either a subgroup, a group, or a phase.

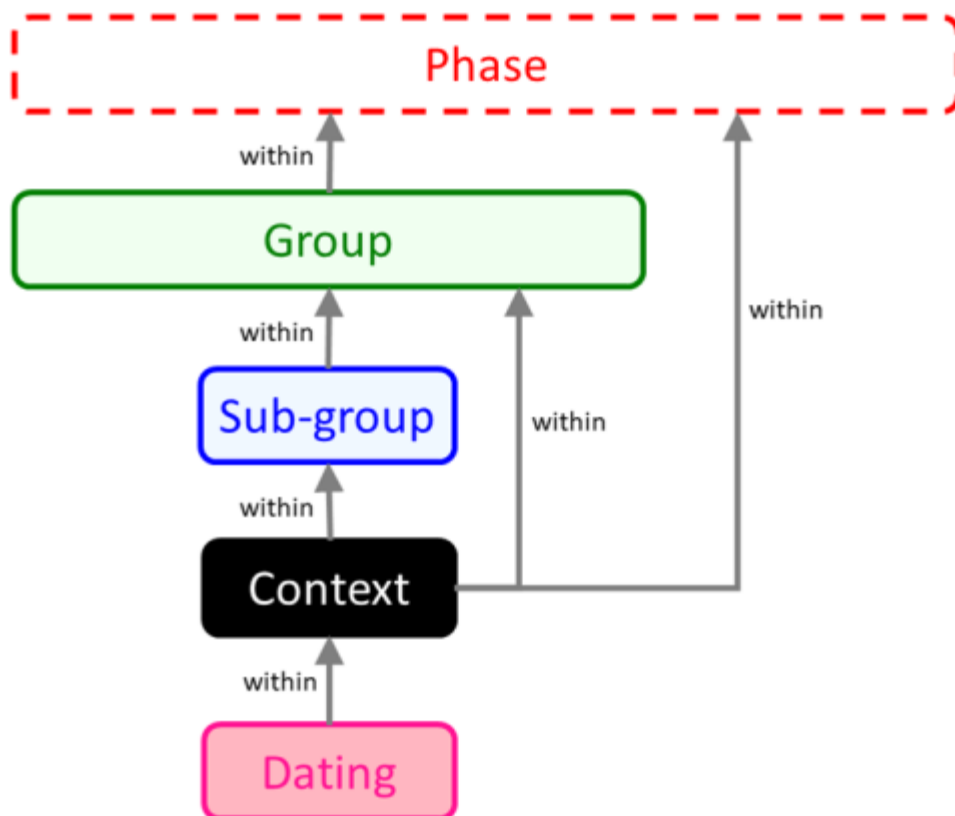


Figure 13: Matrix element containment hierarchy

This containment structure is the mechanism used for grouping and phasing within the prototype application. The cardinality of the containment relationships is one-to-many (e.g. one group can contain many sub-groups; any sub-group can only be within one group). Recursive or cyclical containment is not allowed.

7.2 Design decisions - user interface design decisions

7.2.1 Matrix diagram

The user interface is centred around a stratigraphic matrix diagram with pan and zoom controls, a menu, and an interactive sidebar to display further details for the currently selected elements. Early experimentation to identify a suitable JavaScript diagramming library that might support the use of graph layout algorithms looked at libraries including [Mermaid](#), [JointJS](#), [Dagre-d3](#), and [Cytoscape.js](#). Both Mermaid and JointJS have impressive visual diagramming and charting capabilities. The former uses an interesting declarative markdown to express the diagram and the latter has comprehensive configurable features (though with some associated licensing restrictions). Dagre-d3 offers an appropriate level of diagramming and graph-based functionality for our requirements but now appears to be unsupported. Cytoscape.js offers multiple configurable visualisations and extensions based on an underlying graph library - and was the component eventually chosen for use in the initial



implementation of the prototype application's matrix diagram. An additional SVG layer allows creation of content on the diagram beyond what is drawn by the Cytoscape library itself.

7.2.2 Layout algorithms

There is a substantial existing body of research on algorithms for layout of directed graphs and minimisation of crossing edges, including theoretical background and practical implementations. With the limited time available for development, it was decided to use a suitable existing algorithm and tailor its configuration, rather than embarking on a time-consuming and complex bespoke exercise. Some informal testing of hierarchical layout algorithm implementations (e.g., BreadthFirst and DAGRE) was undertaken using a small example stratigraphic matrix. The DAGRE implementation was chosen based on the encouraging results obtained. As these algorithms are intended to run on *acyclic* graphs, it was required to omit reciprocal relationships and equality relationships during recalculation of node positions to avoid the presence of cycles (paths starting and ending at the same node), which could adversely affect the calculated layout.

7.2.3 Interactivity

The matrix diagram and the sidebar data tables allow interactive selection of elements in the diagram to highlight the appropriate element within the appropriate sidebar table, and vice-versa. This proved particularly useful when dealing with larger datasets as it became more difficult to locate the corresponding elements while viewing and editing a larger matrix.

7.3 Scaling and layout issues of matrix data and implementation decisions

A recurring issue encountered during initial testing was in scalability for anticipated data volumes. There was some discussion during the project of what might be regarded as 'normal' data volumes for stratigraphic data, and some effort was devoted to identifying representative example datasets for use in testing. While the application worked well for smaller numbers of contexts (in the hundreds), the performance was seen to degrade when importing and processing larger numbers of contexts (in the thousands). On investigation of the causes of these issues, some changes were made to the initial prototype.

Certain synchronous functions that took time to complete and caused the user interface to become unresponsive were replaced with asynchronous functions. They would still take the same time to complete but would not cause the user interface to freeze waiting for completion. Secondly, a background local cache was being maintained to avoid losing data in the case of the browser window being inadvertently refreshed without saving the dataset. Although this worked, it proved demanding in terms of processor and memory usage when caching larger amounts of data. As it was not strictly required, the cache was removed. Thirdly, the displayed



data tables were rendering all rows, meaning many HTML elements were being created and reactively maintained even though they were not necessarily visible. Pagination controls were introduced on each table to reduce the amount of HTML elements existing at any one time. Finally, the reactive nature of the application means that any changes to data values will cause individual components to refresh their display as and when necessary. However, in the case of the matrix diagram, a recalculation of the layout plus a refresh of the display can be expensive in terms of processor usage, memory resources and time. A decision was made to decouple the diagram, allowing it to be refreshed independently on demand rather than automatically.

These changes, in conjunction with efficiency improvements to the underlying data storage mechanism, reduced resource usage overall allowing the import and processing of larger datasets and improving the general performance. It would still be recommended to logically subdivide very large datasets (where possible) to obtain better performance.

7.4 Validation testing and rule implementation

When importing externally created data, it is important for it to be checked to conform to the minimum requirements of the importing application. Some fundamental validation can take place during the import, but smaller inconsistencies should not necessarily stop data being imported into the application where the issues identified can be fixed.

A series of rule-based validation checks were implemented to assess and improve the consistency of compiled or imported data. All checks run across the entire dataset and may be repeated on demand, allowing the user to make changes and then revalidate. [Appendix B](#) gives a listing of validation checks incorporated in Phaser. The validation rules are classed as either mandatory (MUST) or optional (SHOULD), and results are appropriately colour-coded to visually indicate where there may be potential problems within the dataset. Elements failing the validation are listed.

The validation was implemented as a separate on-demand bulk process within the application, making it easier to determine where there may be a repeating error pattern occurring e.g., where the imported data omits a particular field. The rationale for this is that much of the initial testing and use cases of the prototype centred around reuse of existing legacy datasets. Given more time, the validation rules could also flag individual fields with appropriate error messages during manual data entry and editing.

7.5 Temporal data analysis implementation choices



The prototype application uses a granularity level of years relative to Common Era (CE) to define date ranges. Positive integers represent CE years, and negative integers represent BCE years. Each dating record has a minimum (earliest) and maximum (latest) year property reflecting a range within which the date of creation/manufacture or deposition must have occurred. In addition, a tolerance may be applied to year values, in terms of either a specific number of years or a percentage. These year ranges are then used to derive temporal relationships between elements. The application has an option to ignore specific dating records, retaining the data but excluding them from any subsequent calculations.

7.5.1 Derived temporal relationships

The element containment hierarchy is utilised to successively calculate broader inherited earliest/latest year limits for each element (contexts, sub-groups, groups, phases). Temporal relationships between elements are then derived based on these inherited year ranges and can be compared to the specified stratigraphy and grouping/phasing structure for the site. These derived relationships are colour-coded where displayed, to indicate where the entered dates of individual items either align or disagree with the stratigraphic analysis.

7.5.2 Derived stratigraphic relationships

The application uses a combination of the entered stratigraphic relationships between contexts and the entered hierarchical grouping information to derive additional stratigraphic relationships between contexts and groups, or between groups (including construction of a Group Matrix diagram). The inherited year ranges for these elements can then be used as in the derived temporal relationships, to indicate alignment or discrepancy with the derived stratigraphy.

The cross-checking that is enabled between the stratigraphic relationships and the temporal dating evidence highlights where there is key evidence from the finds data to focus on and therefore, we would argue, it highlights where there is greater evidence for the chronological sequencing.

7.6 Temporal validation and potential user interface enhancement work

Adding in the Allen operators to express the complexity of the temporal relationships at a human interaction level has added a considerable degree of additional information for the archaeological user to work with. This enables the archaeologist to see more explicitly where there may be discrepancies between the stratigraphic relationships recorded for stratigraphic units and the dating evidence recorded for the objects retrieved from those stratigraphic units. We found, as we developed the Phaser application, that this additional information, although highly useful for checking dating correlations, needed some management through the user interface to enable the nature of any discrepancies to be mediated and highlighted most helpfully to the user. In the prototype, we have used various colour bandings to show



different degrees of mis-match or agreement between the temporal relationships and the stratigraphic relationships (Figure 14). It would be feasible to give a user more control in choices of colours and rule expression as part of further software development if considered useful.

[ONLINE ONLY]

Figure 14: Animation showing the derived temporal relationship 'overlapping' between contexts (1392) and (1396) on separate strands of the stratigraphic matrix.

7.7 Temporal validation and potential export functionality for archive

Largely owing to practical limitations of available time for development and user feedback testing on the project, the current version of the Phaser application does not export the Allen relationships in an archivable (or at least FAIR) format. An agreed formal representation of the temporal relationships could act as a more permanent, and more reusable, method for documenting the temporal relationships often only presented as a visual diagram in PDF format ([Recommendation 6.1](#)). However, any such further development of the Phaser software will need to be based on additional user testing and feedback on the novel functionality for temporal reasoning that is provided in the current prototype version.

The temporal operators highlighted in Phaser, used for analysing (and hopefully resolving) discrepancies between the stratigraphic records and the dating of objects within the stratigraphic units, are reflected in the chronological modelling that is undertaken during Bayesian chronological modelling. An export of the derived temporal relationships within Phaser could also enable their inclusion in chronological modelling software. Again more work would be required to consider how best to surface the temporal relationships that could be incorporated into the chronological modelling software because of the potentially large number of relationships in cases of dating evidence. In addition, the software could be enhanced to export a suitable set of the temporal relations to a JSON format so that they could be reusable without needing the Phaser software. It would also be possible to add some functionality to copy or export from the 'derived temporal relationships' table, or add some form of special report/archival output for combination of imported/compiled/derived data in future development.

The Phaser software calculates the whole set of Allen temporal relationships that hold between the different stratigraphic entities, namely contexts, sub-groups and groups. A decision was made to only compare the set of relationships that hold within a single phase. The software allows the user to select the phase they wish to analyse from a pull-down list of each phase identified in the matrix. Because we define phases to broadly reflect the relationship that a phase 'meets in time' with the phase that follows it, then it is evident that contexts within different phases would not have any overlapping chronological relationship, or indeed any other Allen temporal relationships other than before/after. Therefore only the temporal relationships within each single phase are compared. Even so, some phases with a reasonable number of finds will exhibit a considerable number of relationships.



A way to 'highlight' the more significant relationships has been developed using colour coding so that temporal relationships are highlighted between stratigraphically significant entities (Figure 15). In short, contexts are compared most closely with other contexts, sub-groups with other sub-groups, and groups compared most closely with other groups.

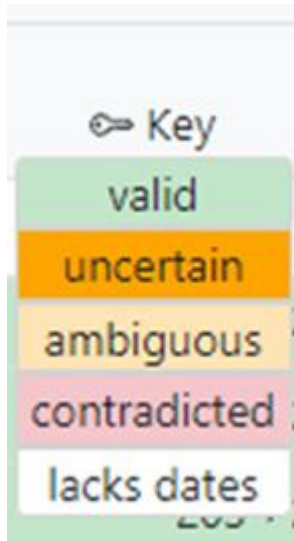


Figure 15: Colour coding to signal measures of confidence and evidence for temporal relationships (see Figure 20)

The temporal relationships are then compared against the recorded stratigraphic relationships and any discrepancies can be highlighted. If the temporal dating evidence is in agreement with the stratigraphic record, we highlight green. If there is a direct contradiction between the temporal dating and the stratigraphy, it is coded red. The current 'rule of thumb' used in the Phaser prototype is to highlight the 'Context to Context' relations with dating evidence. A possible route for analytical recording purposes would be to archive the relationships that highlight green for 'valid' and that hold between two Contexts that both have dating evidence. In other cases, the temporal dating may suggest or reflect some uncertainty in the dating presented e.g. if the date range on finds in one context overlaps with the date ranges of another. The measure of uncertainty is something we will return to. In the case where a temporal relationship is more ambiguous, such as a discrepancy between a context date and a group with which its date range overlaps, then we highlight an ambiguity rather than a direct contradiction.

This is not intended to be a prescriptive error check, but rather a tool for enabling the stratigraphic analyst to re-visit dating evidence and associated grouping and phasing to see if the discrepancies in the phasing need adjusting.

7.8 Reuse of Phaser outputs by Bayesian chronological modellers

One immediate positive outcome from the Phaser development was that it enabled the import of legacy data (such as XSM10 data originally in Stratify .LST format on



ADS) and subsequent export of that data in .CSV format for incorporation into chronological modelling software. This was tested successfully by Moody during her research using her prototype Bayesian chronological Modelling software 'PolyChron', reporting that:

'Inputting the data from Phaser into PolyChron was very simple. All the relative dating evidence required was presented in the datasets. Some datasets required mild formatting, but this is primarily due to PolyChron being a prototype at present and is still fairly rigid in what data formats it can take. However, the time taken to manipulate the datasets into the right format was of the magnitude of minutes rather than the hours that it has taken me to convert data from PDF site reports into the correct format' (Moody in [prep.](#)).

7.9 Further possibilities for R&D

Future possibilities could follow multiple strands depending on the perceived importance of the different aspects highlighted during the development of the prototype application. Some examples that could warrant further discussion include:

- Interactive diagramming and visual features – the application produces the equivalent of traditional paper-based 2D schematic stratigraphic diagrams, but the digital environment offers potential interactive enhancements (e.g. information boxes on hovering the cursor over elements) and enabling alternative displays of the data.
- Improving layout algorithms – owing to time pressures, an off-the-shelf graph layout algorithm was employed. It may be possible to produce a more tailored algorithm better reflecting the requirements of stratigraphic diagramming.
- Improving performance for larger graphs – the initially encountered performance issues were alleviated to some extent but similar problems re-emerge when dealing with very large datasets.
- Data validation and consistency – the basic validation checks introduced could be extended to aid greater consistency.
- Common data structures and archival formats for stratigraphic and grouping/phasing data - a variety of textual formats were encountered in legacy stratigraphic data. The prototype application included facilities to import/export delimited (CSV) files plus JSON file load/save functionality based on the format used by the Cytoscape library. An openly documented common interchange schema capable of capturing and representing all the data required to reproduce the matrix diagram would be a useful contribution and would allow any number of useful applications and tools to archive and interchange stratigraphic matrices.
- Incorporation and use of Linked Open Data reference resources. Pick lists of options for data values would help to improve consistency of datasets by avoiding manual data entry. The consistency would further be improved by listing values originating from established controlled reference vocabularies.
- Reporting – tabular data export facilities exist, but specialised reports for offline analysis of the status of the dataset would be useful. Inherited and derived relationships were used to improve validation and to flag inconsistencies, but there is currently no facility to export or use these outside the application.



8. Conclusions: Challenges and Recommendations

8.1 Challenge: Improving archaeological reporting, publication and data archiving

Data in archaeological digital archives are growing exponentially as more digital technologies are invented and adopted, but the contents of project archives remain to some extent rather piecemeal, possibly partly as a side-effect of the publication of results in a proliferation of formats (Jones *et al.* [2003](#)), and especially when viewed from different international perspectives (Wright and Richards [2018](#)). The piecemeal nature may also be partly explained by the tendency for a systemic hiatus in the overall process by which data are produced during the analysis stage of projects, and the length of time it can often take archaeological information to travel through the business process swim-lanes of commercial and research channels, from primary data recording to publication outcomes. The inconsistency in archive products is reflected in a current approach by most archaeologists to digital archiving that seems to result in a rather 'flotsam and jetsam' collection of products ending up in archaeological digital archives. Or, as one might more tactfully put it, 'archaeology needs not only better policies for data curation, but also the harmonising of the processes of data creation and its deposition for archiving' (Richards *et al.* [2021](#)).

What is definitely not needed is more of the same. We are not suggesting the answer is to just increase the amount of stratigraphic data in the archives, although improving the consistency of what is already there is important. What is needed is better research data management (RDM) (Higman *et al.* [2019](#)) of the stratigraphic data that people are already trying to archive. Rather than just depositing more archived data, the need is to better characterise the indexing of subject matter (meta)data for what has been loosely termed 'heritage data' through good RDM. It follows that to enable a better understanding of how different types of heritage data could be reused, it is necessary to better understand the nature and character of what is meant by the much vaunted phrase 'heritage data' (Albuerne *et al.* [2018](#)). Such 'heritage data characterisation' should enable researchers to better qualify, and ideally quantify, what research questions have, or have not, been addressed. In doing so, it should be possible to better signpost the utility of the data products that are currently being archived so that there is consistency and consensus about what is actually placed in the digital archive for particular reuse cases. This may be as much about better sign-posting, as it is about FAIR metadata (and the current buzz in archive circles around paradata). It is an argument against the sometimes rather simplistic approach taken among archaeologists to archive everything 'just in case' someone needs it someday, that can just lead to a proliferation of unmanageable (unFAIR) data. It may be worth repeating here that an archive item that is never used is literally useless. The results from the Matrix are not that we must archive all stratigraphic matrices, but rather that we need greater professional, academic, and probably international, agreement on when a matrix diagram (among many other



things), is not necessary to still provide a perfectly adequate and FAIR archive of a project.

There may be those who would argue there is a risk in asking for anything more from archaeologists about what they should archive. It is true that during the Matrix project, a recurring argument we heard was 'digital archiving costs too much'. Digital archiving costs generally remain in the region of 2-3% of the overall cost of an archaeological project (Richards *et al.* [1999](#)). Given the scale of resources already required for post-excavation analysis work, we are arguing that people need to maximise the cost-benefits that derive from that analysis work. Rather than simply continuing to archive the digital by-products of a project, we need better measures of the research value (significance) of the resulting data. Here, one is reminded of Moody's identification of the number of deposited but 'empty' (i.e. paid for depositing nothing) files discovered in the ADS archive (Moody *et al.* [2021](#)). There are obvious cost-benefits (in risk management and commercial terms) for making it clearer what the expected by-products for digital archiving should be according to agreed procedures and taking into account, according to agreed professional good practice, the nature of the archaeology discovered. We argue that for 'significant' stratigraphic data, it should be possible to characterise such information by quantification and evidence for the complexity (significance) of the stratigraphy encountered. This article is not arguing for more, just better. To paraphrase a popular aphorism, the digital archive of an archaeological project needs to be 'as complete as necessary, as useful as possible'.

The requirement to record and deposit more useful data to inform better synthetic research was highlighted in the Publication User Needs (Jones *et al.* [2003](#)) project, especially among respondents to the survey who felt there was an inadequate relationship between fieldwork publications and research/publications concerned with synthesis.

'Amongst those who consider the relationship inadequate, 79% feel that measures should be devised to encourage more research and publication that combines results from a number of fieldwork projects to produce broader syntheses' (Jones *et al.* [2003](#), see [fig.19](#)).

This issue is also more recently re-visited and summarised in a recent article assessing the future for Cultural Heritage Management in the USA.

'The two communities — CRM and academia — need each other to move beyond project-based studies to large-scale comparative research. We can analyze long-term socioenvironmental processes posed by such issues as warfare, disease, famine, biodiversity, sustainability, wealth inequality, climate variability, and natural disasters only if we collaborate with each other and with other stakeholders, including scientists in allied fields and members of local and descendant communities. Synthetic research in which we use datasets obtained at the behest of the public to address issues of interest to the public in ways that such results can be impactful (see Kohler and Rockman [2020](#)) is not only in archaeology's best interest but also a moral imperative' (Altschul and Klein [2022](#)).

To enable the best forms of synthetic research, the challenge to the archaeological community is to better define what digital data is needed for reuse and make sure that data are deposited in the FAIR-est ways possible. The challenge to the archives



8.3 Recommendations: What might be addressed by further work?

Addressing the lack of post-excavation analysis procedures documentation

Recommendation 1.1 Further work should be undertaken with stakeholders across the sector, and particularly the major contracting archaeology organisations, to develop shared good practice documentation in the form of an online handbook. This online resource could be piloted initially for stratigraphic analysis and related practices but could potentially be developed for a wider set of related post-ex analysis practices, including submission and format of specialist analysis data. This in turn would help improve the practices for sharing, interoperability and reuse (FAIRness) of data deposited in resulting archaeological archives. A bid to AHRC ([AH/X006735/1](#)) for follow-on funding is seeking to investigate the work needed in this area.

Recommendation 1.2 Work should be undertaken collaboratively across the sector to develop more consistent approaches for creating the (digital) outputs from stratigraphic analysis, using existing stratigraphic principles and recognised standards, and especially to identify common stratigraphic data that are required to form part of a consistent and completed digital archive deposition.

Recommendation 1.3 Explore the feasibility, cost-benefits, and business viability for a consortium of commercial archaeological practitioners to develop and sustain online tools and resources for promoting best practice in post-excavation and analysis methods. Such a Consortium should support the uptake of the proposed online handbook in promoting best practice and FAIR and open principles and research the requirements for a related Community of Practice across the UK heritage sector and internationally to maintain, sustain and grow these online resources. Work on the Matrix prototype tool can be used as a test bed for gathering user requirements for the feasibility of sharing online tools and synchronous data exchange with a multiple range of registered organisations. Phaser-style online tools that use Linked Open Data (LOD) terminologies to help map and cross-reference data in the UK and globally (using multi-lingual vocabularies).

DMPs - Requirements for inclusion of stratigraphic data in Data Management Plans (DMPs)

The adoption and promotion of data management planning is still in its relative infancy among archaeological practitioners. DMPs are currently only produced for archaeological research projects that are funded by the UKRI research bodies. The Chartered Institute for Archaeologists has recommended that 'archaeological projects should include a DMP as part of the archaeological project's WSI or project design, and then maintained throughout project delivery'. It is further recommended that any DMP produced for an archaeological project should be deposited as part of the digital archive. The most recent upgrade to the [OASIS system](#) for reporting



archaeological investigations now includes the functionality for the DMP to be included as part of the OASIS submission.

For projects working under EU funding requirements, a collection of Data Management Tools has been created by the ARIADNE+ project. As the ARIADNE+ project team explains in their introduction on the website: 'The rationale for this is the idea that scientific research should be transparent and replicable, and that its results, including research data, should be shared whenever possible'. The DMP guidance is available from the [ARIADNE website](#).

A key point here is that archaeological projects usually follow a path through a number of stages of data recording, analysis, publication and archive, and so data are updated and created at different points in the process. It is therefore important to include review points in the data management planning so that the DMP can be updated to reflect the stage of data work undertaken and that any final archive of a project should include the latest updated copy of the DMP so it is available for anyone wanting to reuse the archive.

Recommendation 2.1 All investigative projects undertaken or commissioned (i.e. grant aided) by Historic England should be required to prepare a Data Management Plan (DMP) as part of project management documentation.

Recommendation 2.2. DMPs should be updated at the relevant MoRPHE project management review points. An initial DMP will set out what is expected for data at the initiation of a project based upon the anticipated research aims of the project. However, as often demonstrated, an archaeological investigation may encounter unexpected discoveries, and research aims and objectives will need updating accordingly. If new discoveries are recorded on site, or made during the subsequent stages of a project, then the project's DMP will need updating to incorporate adequate digital archiving to reflect the scale and significance of those new data.

Recommendation 2.3. Contracting units should be expected to state the accredited digital repository where the stratigraphic archive data will be deposited within their WSIs and include this in the DMP. This should be incorporated into any best practice guidance that is produced by Historic England (and ALGAO, ClfA, ADS, SMA etc..) for investigative and R&D projects that produce material requiring digital archiving.

Recommendation 2.4. Contracting units should be expected to provide a DMP from the outset of each project. This should be required within a WSI prior to approval (ALGAO [2019](#), 23).

Phase Indexing - Minimum record requirements of phasing process for digital archive

Recommendation 3.1. Develop a federated online system, using the online handbook, for promoting best practice and minimum requirements for phasing and stratigraphic analysis procedures in the UK. Promote FAIR and sustainable best practice within the commercial archaeological sector for the wider public benefit across the UK and internationally through an online registration system for



stratigraphic analysis, in the same way that OASIS has helped raise the standards and access to reports of archaeological investigations.

International Convention on Stratigraphic Methods and Data

Recommendation 4.1. Re-visit, re-affirm and refresh as necessary the existing Harris Principles of Stratigraphy ([Harris 1989](#)) as part of promoting an international convention on archaeological stratigraphic and chronological methods and data.

Recommendation 4.2. A new Universal Law of Spatio-Temporal Succession is proposed for adoption globally (and beyond) by those undertaking archaeological investigation, this new law to be affirmed within the said international convention.

'A unit of archaeological stratification takes its spatio-temporal position in the archaeological stratigraphic sequence from its spatio-temporal juxtaposition between the end of the prior archaeological stratigraphic unit (which lies spatio-temporally prior to it) and beginning of the posterior archaeological stratigraphic unit (which lies spatio-temporally posterior to it), regardless of any other superpositional relationships in the sequence, and presuming gravity has remained constant between the stratigraphic unit's deposition and excavation.'

Recommendation 4.3. A specific working group on stratigraphic standards should be created as part of the work to investigate the best practice documentation for post-excavation analysis. Funds for the initial setting up of this group in the UK will be sought from AHRC and further funding investigated by that group to build a pan-global international convention. An online forum for the former 'Interpreting Stratigraphy' mini-conference is one possible route for taking such initiatives forward, including a related community of practice.

Recommendation 4.4 The said international convention on stratigraphic standards should provide more accurate and interoperable records of the 'Jinji' boundary between human-made strata and naturally deposited strata for (re-)use in anthropocene and ecological research within and beyond archaeology. See [Appendix A](#)

Investigate Heritage Data Classification as a strategy for data management, archiving and reuse of heritage data

Recommendation 5.1 Improve data packaging and related data sign-posting in digital archives to better reflect heritage data characterisation of most common examples of reuse (re-mixing, re-cycling of data) based upon analysis of common use case scenarios and 'action mapping' of typical user journeys through digital archives. Stratigraphic and chronological reuse examples could be used as case studies.

Recommendation 5.2. A project should be instigated to investigate the usefulness of a 'FAIR Cookbook for Heritage Data' for providing examples of good research



data management practices and offer guidelines, information, and pointers to help researchers with problems throughout the data lifecycle (see [Section 6.3](#)).

Enhance the recording of implicit and explicit spatio-temporal and temporal relationships in the digital stratigraphic records of archaeological investigations

Recommendation 6.1. Further work should be undertaken to exploit the advances made possible by the successful development in the Matrix project of an enhanced temporal representation methodology. The Phaser software demonstrates the practical extension of the Harris Matrix method by using the implicit Allen temporal relationships and phasing relationships (Allen [1983](#)) within existing stratigraphic records to strengthen and verify spatiotemporal reasoning within the stratigraphic analysis. This opens further possibilities for improved stratigraphic analysis and related chronological reasoning to increase the use of the additional Allen temporal operators (Dye *et al.* [2023](#)), already implicitly held but not explicitly represented or documented, in the stratigraphic records currently created.

Acknowledgements

This work was supported by the Arts and Humanities Research Council via funding under their grants (ref AH/T002093/1), without which this research could not have taken place.

The authors would like to thank Prof. Doug Tudhope, Prof. Caitlin Buck, Dr Tom Dye, Dr Holly Wright, Dr Edward Harris, Steve Roskams, Tim Williams, Barney Sloane, and Dr Jen Heathcote, for their advice and critical insight, which has helped to inform this project from the outset. Thanks are also due to Bryony Moody, whose complementary research and its own need for robust stratigraphic data has helped us to think through many of the issues we encountered over the course of the project.

Thanks, as ever, to the staff at ADS who have helped to track down various data items we could not find or confirm what were the 'known unknowns' in the ADS archive. Thanks also to the anonymous referee whose comments on the draft text have been most helpful in wrestling this quite extensive article into a rather more digestible form.

We would sincerely like to thank our workshop participants and various consultees, whose engagement and feedback was essential both in understanding disciplinary needs and evaluating prototypes of the Phaser software, including: Dr Kenneth Aitchison, Dr Alex Smith, Dr Claire Christie, Dr Dominic Perring, Guy Hopkinson, Ian Milstead, Vicki Ridgeway, Becky Haslam, Dr Gill Hey, Ken Welsh, Gail Wakeham, Dr Alistair Barclay, Dr Dave Gilbert, John Layt, Guy Hunt, Dr Manda Forster, Doug Killock, Dr Matt Edgeworth, and Prof. Dominic Powlesland.



Particular thanks to Dr Alex Smith at Headland Archaeology for his help with a number of our enquiries and especially for generously enabling our reuse of the MOLA-Headland guidance for illustration purposes.

Special thanks go to the team at MoLA: Louise Fowler, Dr Sara Perry, David Bowsher and all the XSM10 team, especially Al Telfer, Serena Ranieri and Rob Hartle, for so diligently archiving the outputs of their work in ways that enabled FAIR and Open play with their data.

Bibliography

Aitchison, K., German, P. and Rocks Macqueen, D. 2021 'State of the archaeological market 2020', FAME report. <https://famearchaeology.co.uk/wp-content/uploads/2021/10/State-of-the-Archaeological-Market-2020.pdf>

Albuerne, A., Grau-Bove, J. and Strlic, M. 2018 'The role of heritage data science in digital heritage' in M. Ioannides, E. Fink, R. Brumana, P. Patias, A. Doulamis, J. Martins and M. Wallace (eds) *Digital Heritage. Progress in Cultural Heritage: Documentation, Preservation, and Protection*, Lecture Notes in Computer Science, Cham: Springer International Publishing, 11196. 616-22. https://doi.org/10.1007/978-3-030-01762-0_54

ALGAO 2019 *Planning for archives: opportunities and omissions*, <https://historicensland.org.uk/images-books/publications/planning-for-archives/planning-for-archives/>

Allen, J.F. 1983 'Maintaining knowledge about temporal intervals', *Communications of the ACM* **26**(11), 832-43. <https://doi.org/10.1145/182.358434>

Altschul, J.H. and Klein, T.H. 2022 'Forecast for the US CRM Industry and Job Market, 2022-2031', *Advances in Archaeological Practice* **10**(4). <https://doi.org/10.1017/aap.2022.18>

Alvey, B. 1993 'Interpreting archaeology with Hindsight: the use of three dimensions in graphic recording and site analysis' in E.C. Harris, M.R. Brown and G.J. Brown *Practices of Archaeological Stratigraphy*, London: Academic Press. 218-228. <https://www.sciencedirect.com/book/9780123264459/practices-of-archaeological-stratigraphy>

Archaeology Data Service and Digital Antiquity (eds) 2009 *Guides to Good Practice*, Archaeology Data Service, University of York, UK. <https://archaeologydataservice.ac.uk/help-guidance/guides-to-good-practice/>

Ove Arup & Partners, York University, English Heritage and York City Council 1991 *York Development and Archaeology Study*, Ove Arup & Partners, Manchester. <https://www.york.gov.uk/downloads/file/1771/sd114b-english->



[heritage-and-york-city-council-york-development-and-archaeology-study-may-1991-part2](#)

Balm, R. 2015 *Archaeology's Visual Culture. Digging and Desire*, London: Routledge. <https://doi.org/10.4324/9781315673868>

Barker, P. 1977 *Techniques of Archaeological Excavation*, Routledge: London.

Bayliss, A. 2009 'Rolling out the revolution: using radiocarbon dating in archaeology', *Radiocarbon* **51**(1), 123-47. <https://doi.org/10.1017/S0033822200033750>

Berggren, Å. and Hodder, I. 2003 'Social practice, method, and some problems of field archaeology', *American Antiquity* **68**(3), 421-34. <https://doi.org/10.2307/3557102>

Biddle, M., Hudson, D.M. and Heighway, C.M. 1973 *The Future of London's Past; A Survey of the Archaeological Implications of Planning and Development in the Nation's Capital*, Rescue (Trust) Publications, No. 4, Worcester.

Binding, C. 2010 'Implementing Archaeological Time Periods Using CIDOC CRM and SKOS' in L. Aroyo, G. Antoniou, E. Hyvönen, A. Teije, H. Stuckenschmidt, L. Cabral and T. Tudorache (eds) *The Semantic Web: Research and Applications, ESWC 2010*, Lecture Notes in Computer Science **6088**, Berlin, Heidelberg: Springer. 273–287. https://link.springer.com/chapter/10.1007/978-3-642-13486-9_19

Bradley, R. 2006 'Bridging the two cultures - commercial archaeology and the study of prehistoric Britain', *The Antiquaries Journal* **86**, 1-13. <https://doi.org/10.1017/S0003581500000032>

Börjesson, L., Sköld, O. and Huvila, I. 2020 'Paradata in documentation standards and recommendations for digital archaeological visualisations', *Digital Culture & Society* **6**(2), 191-220. <https://doi.org/10.14361/dcs-2020-0210>

Bronk Ramsey, C. 2008 'Radiocarbon dating: revolutions in understanding', *Archaeometry* **50**(2), 249-75. <http://dx.doi.org/10.1111/j.1475-4754.2008.00394.x>

Buck, C.E., Cavanagh, W.G. and Litton, C.D. 1996 *Bayesian Approach to Interpreting Archaeological Data*, Chichester: Wiley and Sons Ltd.

Carey, C., Howard, A.J., Knight, D., Corcoran, J. and Heathcote, J. 2018 *Deposit Modelling and Archaeology*, University of Brighton. <https://www.brighton.ac.uk/pdf/research/set-groups/deposit-modelling-and-archaeology-volume.pdf>

Carver, M.O.H. 1979 'Three Saxo-Norman tenements in Durham City', *Medieval Archaeology* **23**, 1-80. <https://doi.org/10.1080/00766097.1979.11735413>

Carver, M.O.H. 1987 *Underneath English Towns: Interpreting Urban Archaeology*, London: B.T. Batsford.



Carver, M.O.H. 1990 'Digging for Data: Archaeological Approaches to Data Definition, Acquisition and Analysis' in R. Francovich and D. Manacorda (eds) *Lo Scavo Archeologico: Edizioni all'Insegna del Giglio*, Firenze: All'Insegna del Giglio SAS. 42-120.

Carver, M.O.H. 2009 *Archaeological Investigation*, London: Routledge.

Chadwick, A.M. 2001 'What have the post-processualists ever done for us? Towards an integration of theory and practice; and radical field archaeologies'. Paper given at Interpreting Stratigraphy Group Meeting, University of York, York. [PDF](#)

Chartered Institute for Archaeologists 2022 *Dig Digital* [website], <https://www.archaeologists.net/digidigital> [Last accessed: 11 November 2022].

Clarke, A., Fulford, M. and Rains, M. 2007 *The Silchester Project: Roman Town Insula IX The Development of an Urban Property c. AD 40-50-c. AD 250* [dataset], York: Archaeology Data Service [distributor] <https://doi.org/10.5284/1000259>

Coe, A. 2022 *Deciphering Earth's History: The Practice of Stratigraphy*, The Geological Society of London. <https://www.geolsoc.org.uk/GIP001> [Last accessed: 23 December 2022].

Cohen, K.M., Harper, D.A.T. and Gibbard, P.L. 2023 *ICS International Chronostratigraphic Chart 2023/01*, International Commission on Stratigraphy, IUGS. www.stratigraphy.org (Last accessed: 2 March 2023).

Conolly, J. and Lake, M. 2006 *Geographical Information Systems in Archaeology*, Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511807459>

Costa, S., Beck, A., Bevan, A. and Ogden, J. 2013 'Defining and Advocating Open Data in Archaeology' in G. Earl, T. Sly, D. Wheatley, I. Romanowska, K. Papadopoulos, P. Murrieta-Flores and A. Chrysanthi (eds) *Archaeology in the Digital Era: Papers from the 40th Annual Conference of Computer Applications and Quantitative Methods in Archaeology (CAA), Southampton, 26-29 March 2012*, Amsterdam University Press. 449-56. <https://doi.org/10.2307/j.ctt6wp7kg.50>

Crossrail Ltd. 2019 *Crossrail: Archaeological Investigations Conducted in Advance of Construction of the Elizabeth Line* [dataset], York: Archaeology Data Service [distributor] <https://doi.org/10.5284/1055125>

Dell'Unto, N. and Landeschi, G. 2022 *Archaeological 3D GIS*, Taylor & Francis. <https://doi.org/10.4324/9781003034131>

Davies, D. 2017 *The Development Of Archaeological Post-Excavation Within British Professional Archaeology*, BAJR Guide 46 http://www.bajr.org/BAJRGuides/46_PostEx_Theory_Application/46%20Archaeological%20PostExcavation.pdf



Doneus, M., Neubauer, W., Filzwieser, R. and Sevara, C. 2022 *Stratigraphy from Topography II. The Practical Application of the Harris Matrix for the GIS-based Spatio-temporal Archaeological Interpretation of Topographical Data*, *Archaeologia Austriaca* **106**, 223-252. <https://doi.org/10.1553/archaeologia106s223>

Dye, T.S., Buck, C.E., DiNapoli, R.J. and Phillippe, A. 2023 'Bayesian chronology construction and substance time', *Journal of Archaeological Science* **153**, 105765. <https://doi.org/10.1016/j.jas.2023.105765>

Dye, T.S. and Buck, C.E. 2015 'Archaeological sequence diagrams and Bayesian chronological models', *Journal of Archaeological Science* **63**, 84-93. <https://doi.org/10.1016/j.jas.2015.08.008>

Edgeworth, M. 2018 'Humanly modified ground' in D.A. DellaSala and M.I. Goldstein (eds) *The Encyclopedia of the Anthropocene*, Oxford: Elsevier. 157-161. <https://doi.org/10.1016/B978-0-12-809665-9.09999-7>

Edgeworth, M., deB Richter, D., Waters, C., Haff, P., Neal, C. and Price, S.J. 2015 'Diachronous beginnings of the Anthropocene: the lower bounding surface of anthropogenic deposits', *The Anthropocene Review* **2**(1), 33-58. <https://doi.org/10.1177/2053019614565394>

English Heritage 1991 *Management of Archaeological Projects 2* (MAP2), London.

Farid, S. 2014 'Timelines: phasing Neolithic Çatalhöyük' in I. Hodder (ed) *Çatalhöyük Excavations: The 2000-2008 Seasons*, Los Angeles: British Institute at Ankara and The Cotsen Institute of Archaeology, UCLA. 91-129.

Griffiths, S. 2017 'We're all cultural historians now: revolutions in understanding archaeological theory and scientific dating', *Radiocarbon* **59**(5), 1347-57. <https://doi.org/10.1017/RDC.2017.20>

Hammer, F. 2002 *Post-Ex. Manual* [Online], Interpreting Stratigraphy Group. <http://www.york.ac.uk/archaeology/strat/> [Last Accessed: 1 March 2023. Website currently not supported].

Harris, E. 1979 'The laws of archaeological stratigraphy', *World Archaeology* **11**(1), Early Chemical Technology, 111-17. <https://doi.org/10.1080/00438243.1979.9979753>

Harris, E. 1989 *Principles of Archaeological Stratigraphy*, 2nd Edition, London and New York: Academic Press. <https://doi.org/10.1016/C2009-0-21688-6>

Harris, T.M. and Lock, G.R. 1996 'Multi-dimensional GIS: exploratory approaches to spatial and temporal relationships within archaeological stratigraphy' in H. Kamermans and K. Fennema (eds) *Interfacing the Past, Computer Applications and Quantitative Methods in Archaeology CAA95*, Leiden: Leiden University Press. 307-316.



Harris, E.C., Brown, M.R. and Brown, G.J. (eds) 1993 *Practices of Archaeological Stratigraphy*, London: Academic Press. <https://www.sciencedirect.com/book/9780123264459/practices-of-archaeological-stratigraphy>

Hiebel, G., Theodoridou, M., Doerr, M. and Eide, Ø. 2015 *An Extension of CIDOC-CRM to link the CIDOC CRM to GeoSPARQL through a Spatiotemporal Refinement*, Version 1.2. FORTH <https://cidoc-crm.org/crmgeo/ModelVersion/version-1.2>

Higman, R., Jones, S. and Bangert, D. 2019 'Three camps, one destination: the intersections of research data management, FAIR and Open', *UKSG Insights* **32**(1), 18. <https://doi.org/10.1629/uksg.468>

Historic England 2015 *Management of Research Projects in the Historic Environment: The MoRPHE Project Managers Guide* (MoRPHE). <https://historicengland.org.uk/images-books/publications/morphe-project-managers-guide/>

Historic England 2017 *Research Agenda*, Swindon: Historic England. <https://historicengland.org.uk/images-books/publications/he-research-agenda/research-agenda/>

Hodder, I. 1997 'Always momentary, fluid and flexible towards a reflexive excavation methodology', *Antiquity* **71**, 691-700. <https://doi.org/10.1017/S0003598X00085410>

Hodder, I. 1998 'Whose rationality? A response to Fekri Hassan', *Antiquity* **72**(275), 213-17. <https://doi.org/10.1017/S0003598X00086439>

Hodder, I. 1999 *The Archaeological Process: An Introduction*, Oxford: Blackwell.

Hodder, I. (ed) 2000 *Towards Reflexive Method in Archaeology: The Example at Çatalhöyük*, Cambridge/London: McDonald Institute for Archaeological Research, British Institute of Archaeology at Ankara, Çatalhöyük Research Trust.

Hodder, I. 2005 'Changing entanglements and temporalities', in I. Hodder (ed) *Changing Materialities at Çatalhöyük: Reports from the 1995-99 Seasons*, Cambridge/London: McDonald Institute for Archaeological Research, British Institute of Archaeology at Ankara, Çatalhöyük Research Trust. 1-22.

Huggett, J. 2018 'Reuse remix recycle: repurposing archaeological digital data', *Advances in Archaeological Practice* **6**(2), 93-104. <https://doi.org/10.1017/aap.2018.1>

Hutton, W. 1788 'Theory of the Earth; or an investigation of the laws observable in the composition, dissolution, and restoration of land upon the Globe', *Transactions of the Royal Society of Edinburgh* **1**(2), 209-304. <https://doi.org/10.1017/S0080456800029227>



- Huvila, I. 2022 'Improving the usefulness of research data with better paradata', *Open Information Science* **6**(1), 28-48. <https://doi.org/10.1515/opis-2022-0129>
- Huvila, I., Sköld, O., & Börjesson, L. 2021 'Documenting information making in archaeological field reports', *Journal of Documentation* **77**(5), 1–24. <https://doi.org/10.1108/jd-11-2020-0188>
- Jones, S., MacSween, A., Jeffrey, S., Morris, R. and Heyworth, M. 2003 'From the ground up. The publication of archaeological projects: a user needs survey. A summary', *Internet Archaeology* **14**. <https://doi.org/10.11141/ia.14.4>
- Joukowsky, M.S. 1980 *A Complete Manual of Field Archaeology: Tools and Techniques of Field Work for Archaeologists*, Englewood Cliffs, N.J.: Prentice-Hall, Spectrum.
- Kohler, T.A. and Rockman, M. 2020 'The IPCC: a primer for archaeologists', *American Antiquity* **85**(4), 62-51. <https://doi.org/10.1017/aaq.2020.68>
- Lock, G.R. 2003 *Using Computers in Archaeology: Towards Virtual Pasts*, Oxford: Routledge.
- Lucas, G. 2001 *Critical Approaches to Fieldwork: Contemporary and Historical Archaeological Practice*, London: Routledge.
- Lyell, C. 1830-1833 *Principles of Geology: Being an Attempt to Explain the Former Changes of the Earth's Surface, by Reference to Causes Now in Operation*, London: John Murray. <https://doi.org/10.5962/bhl.title.50860>
- Masur, A., May, K., Hiebel, G. and Aspöck, E. 2013 *Comparing and mapping archaeological excavation data from different recording systems for integration using ontologies*, [PDF](#) [Last accessed: 2 March 2022].
- May, K. 2020 'The Matrix: connecting time and space in archaeological stratigraphic records and archives', *Internet Archaeology* **55**. <https://doi.org/10.11141/ia.55.8>
- Moody, B.C. 2019 *Heritage Data Analytics Year 1: The Quality and Utility of Resources in Digital Heritage Repositories*, England: Historic England. Unpublished.
- Moody, B.C. in prep *Semi-Automation of Bayesian Chronology Construction Using a Graph-Theoretic Approach*, PhD thesis, University of Sheffield.
- Moody, B., Dye, T., May, K., Wright, H. and Buck, C. 2021 'Digital chronological data reuse in archaeology: three case studies with varying purposes and perspectives.' *Journal of Archaeological Science: Reports* **40**, 103188. <https://doi.org/10.1016/j.jasrep.2021.103188>
- Naylor, J.C. and Smith, A.F.M. 1988 'An archaeological inference problem', *Journal of the American Statistical Association* **83**(403), 588. <https://doi.org/10.2307/2289281>



Paice, P. 1991 'Extensions to the Harris Matrix System to illustrate stratigraphic discussion of an archaeological site', *Journal of Field Archaeology* **18**(1), 17. <https://doi.org/10.2307/530148>

Papadakis, M., Doerr, M. and Plexousakis, D. 2014 'Fuzzy times on space-time volumes', *eChallenges e-2014 Conference Proceedings*, Belfast, UK, 1-11. [PDF](#)

Pavel, C. 2010 *Describing and Interpreting the Past - European and American Approaches to the Written Record of the Excavation*, Bucharest: University of Bucharest Press.

Pearson, N. and Williams, T. 1993 'Single-context planning: its role in on-site recording procedures and in post-excavation analysis at York', in E.C. Harris, M.R. Brown and G.J Brown (eds) *Practices of Archaeological Stratigraphy*, London: Academic Press. 89-103. <https://doi.org/10.1016/B978-0-12-326445-9.50013-7>

Perry, S. and Taylor, J.S. 2018 'Theorising the digital: a call to action for the archaeological community' in M. Matsumoto and E. Uleburg (eds) *Oceans of Data: Proceedings of the 44th Conference on Computer Applications and Quantitative Methods in Archaeology*, Oxford: Archaeopress. 11-22. <https://www.torrossa.com/it/catalog/preview/5245174>

Pitt-Rivers, A.L.F. 1886-90 *Excavations at Cranborne Chase*, 4 volumes privately printed.

Popper, Karl R. 2002 *The Logic of Scientific Discovery*, Psychology Press.

Richards, J.D., Robinson, D., Wise, A. and Condron, F. 1999 *Strategies for Digital Data: findings and recommendations from digital data in archaeology -- a survey of user needs*, Archaeology Data Service, York.

Richards, J.D., Jakobsson, U., Novák, D., Štular, B. and Wright, H. 2021 'Digital archiving in archaeology: The State of the Art. Introduction', *Internet Archaeology* **58**. <https://doi.org/10.11141/ia.58.23>

Rocks-Macqueen, D. and Lewis, B. 2019 *Archaeology in Development Management. Its contribution in England, Scotland & Wales*, Landward Research Ltd. <https://www.algao.org.uk/archaeology-development-management>

Roskams, S. 2001 *Excavation*, Cambridge: Cambridge University Press.

Roskams, S. 2013 "'Late Antique Field Archaeology": a Legitimate Aim?' in I. Lavan and M. Mulryan (eds) *Field Methods and Post-Excavation Techniques in Late Antique Archaeology*, Leiden: Koninklijke Brill NV. 17–50.

Roskams, S. 2020 'The post-excavation analysis and archiving of outputs from complex, multi-period landscape investigations: the example of Heslington East, York', *Internet Archaeology* **55**. <https://doi.org/10.11141/ia.55.7>



Shanks, M. and Tilley, C. 1992 *Reconstructing Archaeology: Theory and Practice*, New Studies in Archaeology, London: Routledge.

Shanks, M. and McGuire, R. 1996 'The craft of archaeology', *American Antiquity* **61**, 75-88. <https://doi.org/10.1017/S0002731600050046>

Smith, W. 1815 *A Delineation of the Strata of England and Wales with part of Scotland: Exhibiting the Collieries and Mines, the Marshes and Fen Lands Originally Overflowed by the Sea, and the Varieties of Soil According to Variations in the Substrata, Illustrated by the Most Descriptive Names*, London. John Cary.

Spence, C. 1990 *Archaeological Site Manual*, London: Museum of London.

Spence, C. 1993 'Recording the archaeology of London: the development and implementation of the DUA recording system' in E.C. Harris, M.R. Brown and G.J. Brown (eds) *Practices of Archaeological Stratigraphy*, London: Academic Press. 23-46. <https://doi.org/10.1016/B978-0-12-326445-9.50008-3>

Steane, K. 1993 'Land use diagrams - a hierarchy of site interpretation' in J. Barber (ed) *Interpreting Stratigraphy*, Edinburgh: AOC (Scotland) Ltd. 11-14.

Steensen (Steno), N. 1669 *De solido intra solidum naturaliter contento dissertationis prodromus*. <https://archive.org/details/cu31924012131458>

Taylor, J.S. 2016 *Making Time For Space At Çatalhöyük: GIS as a tool for exploring intra-site spatiotemporality within complex stratigraphic sequences*, PhD thesis, University of York. <https://etheses.whiterose.ac.uk/13500/>

Taylor, J. 2020 'Space and time' in M. Gillings, P. Hacıgüzeller and G. Lock (eds) *Archaeological Spatial Analysis: A Methodological Guide*, Wiley-Blackwell Publishing. 408-429. <https://doi.org/10.4324/9781351243858-21>

Thorpe, R. 2012 'Often fun, usually messy: fieldwork, recording and higher orders of things' in H. Cobb, O.J. T. Harris, C. Jones and P. Richardson (eds) *Reconsidering Archaeological Fieldwork: exploring on-site relationships between theory and practice*, Boston, MA: Springer US. 31-52. <https://doi.org/10.1007/978-1-4614-2338-6>

Trigger, B.G. 2006 *A History of Archaeological Thought*, Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511813016>

Tsang, C. 2021 'Red sky at night: digital archiving in England 2020', *Internet Archaeology* **58**. <https://doi.org/10.11141/ia.58.6>

Tudhope, D., May, K., Binding, C. and Vlachidis, A. 2011 'Connecting archaeological data and grey literature via semantic cross search', *Internet Archaeology* **30**. <https://doi.org/10.11141/ia.30.5>

Vandavelde, S., Brochier, J.É., Desachy, B., Petit, C. and Slimak, L. 2018 'Sooted concretions: a new micro-chronological tool for high temporal resolution



archaeology', *Quaternary International* **474**(Part B), 103-18. <https://doi.org/10.1016/j.quaint.2017.10.031>

Watson, S. 2019 'Whither archaeologists? Continuing challenges to field practice', *Antiquity*, **93**(372), 1643-1652. <https://doi.org/10.15184/aqy.2019.141>

Westman, A. and Shepherd, L. 1992 'From city site to research archive...', *London Archaeologist*, London Archaeologist Association [dataset], York: Archaeology Data Service [distributor], 435-43. <https://doi.org/10.5284/1070821>

Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N. *et al.* 2016 'The FAIR guiding principles for scientific data management and stewardship', *Scientific Data* **3**(1), 160018. <https://doi.org/10.1038/sdata.2016.18>

Wheeler, R.E.M. 1954 *Archaeology from the Earth*, Oxford: Clarendon Press.

Whitehouse, N, Schulting, R., McClatchie, M., Barratt, P., McLaughlin, T., Bogaard, A., Colledge, S., Marchant, R., Gaffrey, J. and Bunting, M. 2014 'Neolithic agriculture on the European western frontier: the boom and bust of early farming in Ireland', *Journal of Archaeological Science* **51**, 181-205. <https://doi.org/10.1016/j.jas.2013.08.009>

Whittle, A., Healy, F. and Bayliss, A. 2011 *Gathering Time. Dating the Early Neolithic Enclosures of Southern Britain and Ireland*, Oxford: Oxbow. <https://doi.org/10.2307/j.ctvh1dwp2>

Wright, H.E. 2011 *Seeing Triple: Archaeology, Field Drawing and the Semantic Web*, PhD Thesis University of York. <https://etheses.whiterose.ac.uk/2194/>

Wright, H. and Richards, J.D. 2018 'Reflections on collaborative archaeology and large-scale online research infrastructures', *Journal of Field Archaeology* **43**, supp. 1, S60-S67 <https://doi.org/10.1080/00934690.2018.1511960>