TRAJ Theoretical Roman Archaeology Journal

Lawrence, Andrew Kenneth. 2022. Harder – Better – Faster – Stronger? Roman Archaeology and the Challenge of 'Big Data'. *Theoretical Roman Archaeology Journal*, 5(1): pp. 1–29. DOI: https://doi.org/10.16995/traj.8881

Open Library of Humanities

# Harder – Better – Faster – Stronger? Roman Archaeology and the Challenge of 'Big Data'

**Andrew Kenneth Lawrence,** Institute for Archaeology, Archaeology of the Roman Provinces, University of Bern, CH, andrew.lawrence@iaw.unibe.ch

This paper deals with the integration of large datasets in Roman archaeological research. Starting with an overview of the definition of the term and its use in the humanities and in Roman archaeology, the paper then traces the use of big data in four sub-fields of Roman archaeology: pottery studies/Samian ware, zooarchaeology, numismatics, and landscape archaeology. Finally, the sub-fields are compared, and recommendations are offered for future data-driven research. Three conclusions can be drawn. 1) The managing and understanding of ever-increasing datasets have been a dominant theme of the last 40–50 years in Roman archaeology but possibly has not been explicitly discussed. 2) Many databases are concerned with only one specific unit or artefact genre, thus reinforcing a certain 'siloisation'. 3) Data-driven Roman archaeology has the potential to challenge existing narratives and even act as a corrective for traditional frameworks or narratives.

## Introduction

The last decades have seen an explosion in data-driven research, not only in archaeology more generally but also in Roman archaeology specifically. Many large and smaller projects are not only working with large amounts of data but are also constantly producing and generating new methodologies, scripts, and even stand-alone computer programs for collecting, curating, and analysing datasets. Furthermore, several conference sessions are now also focusing on data-driven projects and methods, such as computational modelling in the frontier regions of the Roman Empire (Verhagen 2022) or rich datasets from large excavation schemes (Mazzilli and Aldred 2022). Thanks to collaborative software solutions and a greater awareness of the available statistical toolkits, processing and visualising data are now also easier, with many papers benefiting from complex yet easy-to-read graphs and tables (Garland 2021; Kamash 2021; for a plea for better graphics in theoretical archaeology journals see Baxter and Cool 2016: ix–x). The possibilities of data-driven research seem almost unlimited.

Bearing that in mind, it may be time to zoom out and take stock of the development of data-driven Roman archaeology as a whole. Where are these petabytes leading us? Are our pictures of the Roman past becoming more complete through ever-increasing datasets? Do the results from big data projects really lead to greater leaps in our knowledge within the field of Roman archaeology? Is Roman archaeology experiencing a 'big data revolution?' If so, then a case should be made for pushing data-driven research even harder within our discipline, facilitating data collection, analyses, and deposition. If not, we should ask ourselves whether and how large-scale projects might be recalibrated—the quantitative methodology toned down in favour of something more interpretative, with fewer graphs and tables. The aim of this paper is thus to give an overview of developments within the field of Roman archaeology in the last decades about the epistemology of data-driven archaeology on one hand and the practices of data collection and handling within projects or papers on the other.

The first, and most important, caveat of this paper is that I am neither a data scientist nor an expert in any type of data analysis using the relevant programs, such as SPSS, R, Python, or even Excel. My hitherto 'expertise' in this specific field is self-taught; my academic education, competencies, and focus have been on the qualitative side of Roman archaeology—i.e. the description, identification, interpretation, and contextualisation of finds, assemblages, findings, structures, or sites. Within the scope of my own post-doc project 'Reassessing Roman Impact – Geographies of Power and Domination in the Roman West and Beyond', I have become engaged and interested in the data-fication of archaeological sites and their cross-comparability over time and space.

A second point concerns this paper's scope. I will focus on Rome's north-western provinces, as this is the area I am currently working on. In addition, I am not an expert on any of the four case studies I have chosen; I am, however, very interested and invested in how the results from these four sub-disciplines can feed into our broader narratives of the Roman past.

## (Big) Data, the Humanities, and Roman Archaeology

Systematic usage of the term *big data* can be traced back to the late 1990s when it was used in the graphics departments of film studios creating special effects in blockbuster movies. Within data science specifically and also in broader terms, big data has come to be characterised by the following traits (colloquially known as the 'three Vs'; Kitchin and McArdle 2016):

- Volume – i.e. the size of the data involved, usually measured in bytes.
- Variety – i.e. the range of formats to which data is stored (xls, csv, pdf, html, jpeg, tif, audio files, etc.)
- Velocity – i.e., the speed at which data is generated and transferred and the speed at which storage becomes obsolete (defunct websites, unreadable media)

In Archaeology and in Roman Archaeology the "Big" in Big Data should not merely be reduced to a question of scale. On one hand, there is no cut-off point or threshold where Data becomes Big Data (Van Valkenburgh and Dufton 2020: S2), on the other, as will be demonstrated in the case studies below, archaeology has been working with large datasets since its very beginnings, albeit in non-standardized forms.

Rather, it is the development of the interlinking of existing data repositories—for example via Linked Open Data design principles (Schmidt, Thiery, and Trognitz 2022)—which is now facilitating a stark increase in the amount and variety of data which can be cross-queried and analysed (see also Cooper and Green 2016: 294–299 and more generally Boyd and Crawford 2012: 662–665).

A further trait of archaeological Big Data is its accessibility. Nowadays, it is not only easier to upload but also to gain access to thousands of data entries (be it sites or different categories of finds) which can then be compared and interrogated with relative speed and ease.

Recently, archaeological data has been termed *slow data*, shifting the focus away from the volume or size of the data towards the process and speed of its creation. Despite its richness, archaeological data—especially in its structured, normalised, and reliable form—can take many years to generate (Bickler 2021: 186). In 2015, Andrew

Bevan stated that—compared, for example, to the data produced by social media and Big Tech companies—archaeological big data isn't actually that big in terms of bytes (2015: 1473). However, the steep rise in quantity and quality and its increased use in Roman archaeology in recent years merits the adjective 'Big'.

### *Uneasy bedfellows? Roman archaeology and quantifications (early 20th cent – ca. 1960)*

In the first half of the twentieth century, data-based methodologies were unsurprisingly not widely integrated within Roman archaeology. Nonetheless, a tendency to systematically collect and collate information can be observed. In Germany, between 1894 and 1937, the Reichslimeskomission published 15 sections (550 km in total) of the Uppergermanic–Rhaetian Limes (ORL) (Moschek 2011: 30). In France, the first edition of the *Carte Archéologie de la Gaule* was published in 1931 (Couisson 1931). In Britain, from the 1920s onwards, Osbert Guy Stanhope Crawford pioneered quantitative methods in the shape of distribution maps and aerial photographs (Bowden 2001: 38–43). In the same period, Cyril Fox (2010) published his doctoral thesis on the archaeology of the Cambridge region. The Roman period listed eight Roman Towns, Earthworks, etc. and canals as well as 15 rural sites (Fox 2010: 159–236).

After the middle of the twentieth century, a possible paradigm shift can be observed in Roman archaeology in Britain and Europe. In his assessment of post-war Romano-British archaeology, Martin Millett names the growth of archaeology outside of academia as *the* defining characteristic of this period (2015: 27–33). The increase in excavations, both in quantity and in size, will have led to an exponential increase in archaeological data.

A similar assessment is made by Paul-Marie Duval in his overview of Gallo-Roman archaeology in post-war France, naming the creation of the CNRS (Centre national de la recherche scientifique), founded in 1939, and the institutionalisation of archaeology along with aerial photography and subaquatic archaeology as two new fields which contributed substantially to the understanding of Roman Gaul (Duval 1989). Aerial archaeology contributed substantially to a growth in the dataset for Iron Age and Roman rural settlements in northern France. A growing interest in pottery studies—especially terra sigillata—not only from an art-historical perspective but also as a mass-produced and widely traded commodity, can also be registered. Both the archaeology of rural settlements and pottery studies were topics which would become synonymous with Roman archaeology in western Europe.

In a chapter titled 'Short History of the Beginnings of Mathematics in Archaeology', the French prehistorian François Djindjian notes especially the years between 1950 and 1965 as the period when quantification truly began to permeate archaeological methods, listing a dozen papers which contributed to this movement (2015: 66). From

a theoretical point of view, the emergence of the Annales school of thought in France in the second quarter of the twentieth century, with its focus on interdisciplinarity, long-term change, and overarching social phenomena, can be seen as further anticipation of data-driven research—as many of the questions posed by the l'école des Annales could be operationalised by systematic data collection and analyses (Bintliff 1991: 1–32; Schnapp 1981).

### *New Roman archaeology? (ca. 1960 to the turn of the millennium)*

In archaeology in general, the 1960s saw the publication of landmark essays and volumes which would leave permanent marks on the discipline. The rallying call of New Archaeology (Binford 1962; Clarke 1973) and the subsequent responses have been discussed in depth elsewhere (e.g. Malone and Stoddart 1998).

In France, the rebranding of the Centre d'Analyse Documentaire pour l'Archéologie in 1958 by Jean-Claude Gardin saw computational archaeology make a quantum leap in terms of methodology and theoretical underpinning (Moscati 2016). In the subsequent years, the centre would focus on objects, iconography, and texts and would be at the forefront of the discussion on the integration of computing methods into more classical archaeological topics.

In Germany, the discussion revolved less around the quantifying of archaeological source material and gravitated more towards the interpretation. Hans Jürgen Eggers, the chair of prehistory in Hamburg, published his *Introduction to Prehistory* in 1959, which rallied against the narrow interpretation of material culture as a marker for ethnicity (Eggers 1959: 255–297; von Carnap-Bornheim 2001: 179–197) and made the case for stronger methodological reflection (Theune 2006: 328).

The impetus for New Archaeology was slow to arrive in Roman Archaeology. In an article titled 'A Classical Archaeologist's Response to the "New Archaeology"', Stephen Dyson (1981) reflects on why 'New Archaeology' penetrated so superficially into classical archaeology. One reason he offers is the primacy or even dependency of the subject on literary texts and canonised works: 'Living with Virgil, Homer, or the Parthenon is rather like having a dominant senior partner. One has trouble speaking one's piece' (Dyson 1981: 242). In his review of the role of classical archaeology in the journal *Antiquity*, Nicola Terrenato strikes a similar chord: the mastering of ancient languages and a knowledge of key works is the—sometimes arduous—prerequisite before one's own original research questions and methods can be developed (Terrenato 2002: 1109). The same can be said for Roman archaeology north of the Alps: the discussion of the accounts of events by ancient authors such as Caesar and Tacitus continued to dominate the discourse, even if the archaeological body of evidence for the periods and regions in question has continued to grow in quantity and quality.

How did Roman archaeology deal with these ever-increasing large datasets? A look at the early conferences organised within the scope of the CAA (Computer Applications in Archaeology) shows that with very few exceptions, Roman themes were not really treated until well into the 1980s (Laflin 1982). The creation of the Italian journal *Archeologia e Calcolatori* in 1990 (Cristofani and Francovich 1990)[1] was a further step in this direction.

In many European countries, the anchoring of archaeology within a legal framework and the subsequent growth of excavations in size and frequency meant that many scholars were taught archaeology not only in the lecture hall but also in the field (for England, see Fulford and Holbrook 2011; for Germany, see Heising Unpubl.). The hands-on aspect of excavation archaeology shifted the interest towards managing the sheer quantities of varying archaeological evidence (Laflin 2013: 7–30), and by 1993, Richard Reece was not so much criticising the lack of data-driven research in Roman archaeology, but more specifically the lack of theory with which the data was to be interpreted (Reece 1993).

The creation of the Archaeology Data Service (ADS) at the University of York in 1996 aimed to 'collect, describe, catalogue, preserve, and provide user support for digital resources that are created as a product of archaeological research' (ADS 2017). Almost 25 years later, it has become an indispensable research and archiving facility for archaeology in the United Kingdom (Richards 2017).

Finally, in his stimulating essay from 2004, Greg Woolf wrote of an 'information explosion' in the Roman archaeology of the past few decades (2004: 417–418). Almost 20 years on, the shockwaves of this explosion show no sign of abating, and the data is becoming richer, more varied, and—yes—bigger.

### *Digital Roman archaeologies: from the turn of the millennium to the present and beyond*

In the last 20 to 25 years, archaeology and cultural heritage have embraced the digital turn.

Personal computers have become faster and more powerful but also smaller and more mobile, opening up new opportunities to gather and disseminate gigabytes of data in the field, in museums, archives, libraries, and (home) offices. Likewise, technological advances have created a surge in possibilities for connectivity, leading to speedier rates of data transfer.

As a consequence, new pathways have opened up for digital archaeology to explore and expand on: 3D scanning and reconstruction, neural networks and artificial intelligence, and remote sensing are only some of the fields which make up the world of digital archaeology of today (**Figure 1**).

The creation of research frameworks and infrastructure enabling the linking and exchange of (meta)data is another hallmark of the last decades: the ARIADNE project, part of the EU-funded DARIAH network ran from 2013 to 2017 (and to 2022 under the name of ARIADNEPlus) and ‹aims to bring together and integrate the existing archaeological research data infrastructures so that researchers can use the various distributed datasets› (DARIAH-EU 2012).

The launching of new data infrastructure tools and portals shows no sign of slowing, with databases for Roman-period sites and objects moving online at a speedy pace. The list of open databases on the Project Mercury homepage lists 42 datasets offering information for Roman-period material (Project Mercury 2018). The website of the Oxford Roman Economy Project (OxREP) links to seven datasets which were created between 2005 and 2010. Likewise, image databases and digital journal repositories are also expanding at a phenomenal rate, yielding access to information in hitherto unwitnessed quantities (OxREP 2012).

**Figure 1:** Artifically generated image created with “midjourney” and using this paper’s title as the prompt (Source: Author).

The establishment of university courses, research labs, and even departments for digital or computational archaeology further underlines how integral the digital has become to the discipline. From a 'Romanist' perspective, the courses of digital archaeology offered by universities such as Cologne[2] and York[3] are especially interesting, as these archaeological institutes/departments also offer courses on classical and Roman-provincial archaeology, which could lead to innovative synergies.

The literature on 'digital archaeology' in its broadest sense has thus multiplied exponentially in the last few years. Multi-authored volumes, conference proceedings, e-publications, and journals all contribute to the publication landscape of this dynamic subject.

In 2017, the journal *Archéologies numérqiues*[4] was established in France. A year later, the *Journal of Computer Applications in Archaeology* was launched.

Despite the growth of publications of digital archaeology, papers covering case studies from the Roman Empire or methodological innovations working with Roman-period datasets are still in the minority. This marginalisation seems almost mutual. A brief look at the tables of contents of the journals for Roman archaeology in Europe and the United States of America (*Germania*, *Berichte der Römisch-Germanischen Komission*, *Mitteilungen des Deutschen Archäoloigschen Instituts*, *Gallia, Journal of Roman Archaeology*, *American Journal of Archaeology*) over the past decade reveals that data-driven research has not truly taken off in these periodicals and their respective research traditions.

Finally, similar to the transformations taking place in digital humanities (Klein and Gold 2019), a salient call has been made for 'a form of introspective or more self-aware Digital Archaeology, one which consciously seeks to understand the underlying processes and behaviours that sit behind the tools, technologies, and methodologies applied' (Huggett 2015: 89) or to be more critical of the processes which lead to the structuration of archaeological source material (Hacıgüzeller, Taylor, and Perry 2021: 1711–1714).

## (Big) Data and Roman Archaeology Today: Shining a Light on Some Themes

### *Pottery studies and Samian Ware*

Data—i.e. the systematic quantification and measuring of fragments and vessels—has always played a significant role in Roman pottery studies,[5] even if it was not always designated by this specific term. Due to its standardised repertoire, chronological sensitivity, and popularity within the Roman world and beyond, the study of Samian ware is closely intertwined with quantitative methodologies. Although early forays in Roman pottery studies in the eighteenth and nineteenth centuries were concerned with the motifs on the decorated bowls and the names on the stamps in the late nineteenth

century (Pinon 2015), the potential of the quantitative study of terra sigillata began to be realised, and data-driven methods were quickly adopted.

The introduction of Hans Dragendorff's seminal essay on terra sigillata from 1895 in the *Bonner Jahrbücher* states that 'every Roman shard found outside Italy is very directly a document of cultural and commercial history' (Dragendorff 1895: 2). Without specifically naming it, Dragendorff was demanding and possibly anticipating some kind of database.

In 1904 'Les vases céramiques ornés de la Gaule romaine (Narbonnaise, Aquitaine et Lyonnaise)',[6] Joseph Déchelette (1904) repeatedly uses tabular overviews to illustrate quantities of various motifs and of production centres. At the 1969 Fautores Conference in Speyer (D), a much-discussed paper was submitted with the title 'Computer-aided identification of fragmentary potters' stamps' (Zabehlicky-Scheffenegger 2013).

In 1971, Elisabeth Ettlinger's review of Christian Goudineau's analysis of the pottery from Bolsena (I) underlined the link between the quantity of black gloss ware ('Campana') and terra sigillata on one hand and the chronological development of Late Republican/Early Augustan sites in the western Roman Empire on the other. Ettlinger's (1971) tables show how, in the space of around 15 years, in one of the excavated zones, the ratio of Campana to terra sigillata changes from 20:1 to 0.3:1.

From a methodological point of view, Clive R. Orton's work is of great importance. His papers and publications address not only how to carry out statistical analyses in detail but also why and to what end (e.g. Orton 1980: 15–24). Inspired by Orton's work, Armand Desbat published a succinct yet important paper in the 1990 proceedings of the Société Française d'Étude de la Céramique Antique en Gaule (Desbat 1990). Here, again, the focus is on the method—i.e. how pottery can be counted and what this means for statistical analyses. A conference on Gallo-Roman pottery held at Bibracte and published in 1998 went one step further and collected 18 papers by different authors on pottery quantification. The final paper is a 'protocol' of the results gleaned from the discussion and shows in no uncertain terms how Roman-period pottery can be not only quantified but also published (Arcelin and Tuffreau-Libre 1998).

It is not only pottery forms or types which have been systematically collected. Even before the First World War, Auguste Oxé started to collect all potters' stamps of Arretine ware; it would take several decades until this undertaking was finally published by Howard Comfort in 1968 (Oxé and Comfort 1968). Philipp Kenrick's revised and enlarged edition of this opus magnum, consisting of 36,000 entries, advanced the state of research on early terra sigillata considerably (Oxé, Comfort and Kenrick 2000). Although it is a pity that the database has not been migrated to a new platform (Kenrick 2014). the dataset has been integrated into the 'Samian Research' portal of the Römsich-Germanisch Zentralmuseum of Mayence (Samian Research 2016).

The acknowledgement of pottery's role in the dating of archaeological structures, in answering economic questions, and in addressing topics concerning food and drink consumption led to a growth in pottery studies. The 2018 edited volume by Penelope Allison, Martin Pitts, and Sarah Colley titled 'Big Data on the Roman Table' sets the scene for further studies in this field by implementing data-driven analyses to explore trade and social networks associated with terra sigillata. Future studies on the social behaviour of terra sigillata and on Roman eating and drinking practices would do well to consult this methodologically innovative and theoretically grounded collection of papers on the subject.

Most recently, the Römisch-Germanisch Zentralmuseum im Mayence has been creating a Linked Open Data infrastructure for terra sigillata stamps incorporating data from reference works and connecting them to dated military sites on the frontier in order to establish a more robust stamp chronology (Thiery, Mees, and Gottwald 2020).

### *Zooarchaeology*

In its early years, the investigation of faunal remains concentrated on ahistorical periods, mostly the Paleo-, Meso-, and Neolithic (Gifford–Gonzalez 2018: 19–50). In Roman archaeology, it was again in the decades after the middle of the twentieth century when bone assemblages from Roman sites began to be collected, measured, analysed, and published. In Germany, the faunal remains from military sites, such as the extramural settlement of the fort of Butzbach and the Augustan supply base of Rödgen were published in 1960 (Habermehl 1960) and 1962 (Habermehl 1962), respectively. In Austria, at roughly the same time, a report on animal bones from the Magdalensberg was published (Ehret 1964). In France, early zooarchaeological reports can be found in Marseille (Jourdan 1976). In Britain, a couple of years later, assemblages from London and Exeter were published (Maltby 2019: 11).

The research questions tied to the data were primarily concerned with the diet of the social groupings studied (Davies 1971). Other aspects, such as bone-working (Crummy 1981) and ritual (King 2005: 329) gained traction quickly. The development of the Ossobook database first at the University of Basel (Schibler 1998) in 1990 using dBASE and from 2007 at the Ludwig-Maximilians University of Munich (Kaltenthaler and Lohrer 2018) with its own open-source framework ('xbook') allows for standardised, international collaboration based on common criteria of animal bones. This programme lays the foundation for biometrical analyses, which further our knowledge of the size variations of species before, during, and after the Roman Empire.

It is exactly this type of standardisation which allows for large-scale studies. For example, Anthony King's paper from 1999—'Diet in the Roman World: A Regional

Inter-Site Comparison of the Mammal Bones'—shows, using an impressive dataset collected from published site reports, how meat consumption within the *imperium romanum* could vary from region to region and raises questions regarding what extent Roman imperialism may have impacted foodways and agricultural practices (King 1999).

A more recent database project, consisting of 200 rural sites mainly from northern Gaul, shows how cattle gained importance between the second and the first centuries BC (Méniel et al. 2009). The past decades have seen zooarchaeology going beyond working with single databases, with the compilation and comparison of data from different sources and databases now being pursued systematically. For example, the ZooMWest project is tracing trajectories in animal husbandry in different regions in the western Mediterranean (ZooMWest 2017). With the project's wide geographical and chronological scope, its integration and publishing of large datasets and statistical analyses, long-term (i.e. from the Bronze Age to Late Antiquity) trends in the development of livestock and their interactions with ancient societies can be reconstructed (see for example Nieto-Espinet et al. 2021).

In theory, the integration of multiple datasets seems like a fantastic opportunity to enlarge the sample size of a study and to broaden its scope either regionally and/or chronologically; however, in recent years several methodological issues have come to light regarding the comparability of source material coming from different sources or analysts. Training, experience, access to resources for comparison, and financial and time constraints are all factors which can influence the identification of specimens within bone assemblages, and which can generate biases within the published data. While such distortions cannot be completely nullified, it should be at least possible to pinpoint possible sources of errors: Explicit explanations of the methodologies chosen, of the criteria implemented for the identification of the bones as well as the publishing of images of uncertain identifications, have been put forward as viable recommendations to render datasets more transparent, reusable, and sustainable (Lau and Kansa 2018: 37–38; see also Atici et al. 2013 for a case study and recommendations on working with legacy data). A further step for the interconnecting of databases is the creation and implementation of common ontologies as exemplified by the tDar initiative of the Digital Antiquity organization (tDar 2011).

Quantitative zooarchaeology is a good example as regards to the potential and challenges of the interoperability of multiple databases and the required standardization of the processes involved in recording and publishing. Inevitably, in recent years, critical voices have rightly been raised. In his introductory chapter for the edited volume *Zooarchaeology*, Umberto Albarella acknowledges that while data comparability

is paramount, a certain amount of leeway should be tolerated when it comes to the analysis itself. 'The study of a zooarchaeological assemblage represents an intellectual undertaking rather than a mechanical collection of data' (Albarella 2017: 9). One immediate reaction would be to state that there should not be a binary choice between standardised recording and creative interpretation. A minimal amount of normalized data can be expected in the shape of a database; other, more specialised measurements and their registration can be optional.

### *Numismatics*

Due to their standardised production, Roman and, more recently, Iron Age coinages have lent themselves to quantification and statistics, paving the way for their use in big-data projects. An early example of coin quantification for chronological discussion and socio-political significance can be seen in the Swedish numismatist Sture Bolin's visionary thesis from 1926 on coin finds in Germania libera, in which he painstakingly collected several thousand coins. (Bolin 1926; Bolin 1929). Konrad Kraft's essay from 1956, 'zum Enddatum des Lagers von Haltern', continued in a similar vein (Kraft 1955/56): his method, determining the ratio of issues of types Lyon Altar I to Lyon Altar II in Augustan sites, is still used today (Wigg-Wolf 2018: Fig. 2). 'Coins and the Archaeologist' is an early showcase for implementing statistical analyses and data visualisation on coins (Reece and Casey 1974). Pie charts, bar graphs, distribution maps, and cluster analyses all feature in the papers.[7]

From the 1970s onwards, the concept of cliometrics began to permeate Roman studies. Cliometrics can be described as 'the application of economic theory and quantitative methods to the study of economic history' (Diebolt and Haupert 2021: 1) and was developed in the 1950s by US economic historians (Williamson and Whaples 2003). For the Roman world, the impetus came from ancient history, with scholars such as Keith Hopkins (e.g. 1980), Elio Lo Cascio (e.g. Lo Cascio and Malanmina 2014), and Willem Jongman (e.g. 2014) using data-driven methods to answer questions on economic growth, trade networks, and demography. A debate can be observed on how ancient cliometrics can be integrated into the wider discipline of ancient history (Diebolt 2011). Despite many reservations (e.g. Greene 2006), archaeology could help to bridge this gap, as it is a discipline which has a long tradition of working with fragmented data and models.

Moving away from large-scale statistical models and focusing more on archaeological approaches, a methodology which has only been received sporadically outside its original publication is the analysis of the empirical probability of coin sequences in the Upper Germanic-Rhaetian frontier zone, as devised by Klaus Kortüm

(1998). In this approach, Kortüm compares the coins from military sites situated in the frontier zone with a standard cumulative coin curve consisting of various sites from the left bank of the Rhine. The calculated deviations—especially the first peaks of the graphs—point to intensity in site occupation. While the method has been criticised in some quarters based on a lack of representativity of the samples used, the study is still impressive in its innovative use of large coin datasets.

It is not only coins from sites which have been subject to big data analyses. A recent conference proceeding from 2019 on coin hoards shows the complexity and challenges which come with the processing and analysis of thousands of objects belonging to the same archaeological event. In the conference's final contribution, Jean-Marc Doyen asks, 'If we really have to study mega-deposits, how far can we realistically go?' (Doyen 2019: 321), naming financial and temporal constraints as possible criteria which can influence the study of these huge datasets. Doyen makes the case that the analysis of mega-hoards will yield new information, as opposed to more of the same, only if the right questions are asked: for example, on the chemical composition of the coins, information on the dies, the possible owner(s) of such hoards, or the reasons for these kinds of massive accumulations. It is interesting to note that the 'discovery' of new emperors as issuers of hitherto unknown coin types does not occur from the analysis of mega-hoards but rather from small depots or even from coin finds from sites (Doyen 2019: 316).

It is, however, in the shape of databases and linked open data that the combination of numismatics and large datasets really comes into its own (DARIAH-EU 2018). Early coin databases started in the late 1980s and have been expanding in scale and complexity since the turn of the millennium. The main current trend is the linking of different collections (e.g. in museums) and existing databases to build a networked data collection which can be queried.

The collaborative project of Nomisma.org (NOMISMA 2022) was initiated by the American Numismatic Society in 2010 to facilitate the online presentation of numismatic concepts and employ the methods of the Semantic Web and Linked Open Data. It is thus now possible to jump from a single coin in the collection of Berlin State Museums, to information on its place of minting, to other coins minted from the same site. Thanks to the OCRE project (Gruber et al. 2014), the collation and visualisation of a series of coins belonging to the same type can now be achieved in minutes. The premise for these interlocking databases is the utilisation of the now widely used Research Description Framework (RDF) and a common ontology for all participating databases (RDF Working Group 2014). A further resource which showcases the power of big numismatics datasets is the coin hoards of the Roman Empire database (Lockyear

and Gruber 2013). At the time of writing, 13,170 hoards are in the database. A short query of hoards discovered in the last 20 years (100 BC – 400 AD) shows that only 8.1% of the registered hoards date to decades between 80 and 140 AD. It is these types of preliminary results which can lead to deeper investigations on hoards, hoarding practices, and motives in the Roman world (CHREP 2022).

### *Landscape archaeology*

In Roman archaeology, the study of the built environment has traditionally focused on cities, towns, villas, and farmsteads. While these categories are undoubtedly still relevant for understanding any region, since the 1950s, tentatively at first then at an increased intensity, the emergence of landscapes as an additional unit of analysis can be observed (Trigger 1967).

A case can be made for linking the rise of the prevalence of landscape archaeology to the increasing use of databases and geographical information systems on one hand, as both have facilitated the shift away from the study of a single site towards larger multi-site, regional studies, but also to the more frequent integration of scientific methods such as palynology, paleoecology, geoarchaeology, archaeobotany, and zooarchaeology on the other. According to John Lloyd, one of the strengths of landscape or settlement archaeology is its ability to generate data with which to test accepted ways of looking at past societies and to suggest alternative models (Lloyd 1991: 233). Ian Hodder and Mark Hassal's (1971) famous case study of Romano-British walled towns shows exactly how this could be achieved.

In German-speaking archaeology, the debate on the differentiation between 'Sieldungsarchäologie' and 'Landschaftasarchäologie' is more explicit. Nowadays, according to Sebastian Brather, 'Siedlungsarchäologie', in its widest definition, can be divided into three categories: Environment, Settlements, and Settlement. *Environment* takes ecological factors into account; *Settlements* focus on specific sites, and *Settlement* (singular) on habitation patterns (Brather 2011: 134, Fig. 5).

In France, landscape archaeology (*archéologie du peuplement*/*archéologie du paysage*) really took off with the development of aerial photography. In his many publications on the landscapes of Northern France, Roger Agache (1978), the pioneer of French aerial archaeology, was able to show how dense the hitherto unknown Iron Age and Roman period rural landscape for this area was.

The approaches and methods of francophone *archéologie du paysage* seem to have been better integrated into mainstream archaeological research (Robert 2017). Philip Leveau's work on *Caesarea Mauretania* from 1984 seamlessly integrates urban

analysis, landscape archaeological investigation, and epigraphy into one coherent study. The Archaeomedes Project, financed by the European Commission and running from 1992–1994 with the aim of understanding soil degradation and desertification in Mediterranean landscapes, focused, in one sub-project, on the settlement in the Rhone valley, thus impressively demonstrating how settlement networks can change and persist for several centuries. By integrating geological and environmental data, it also made an explicit connection to contemporary anthropogenic activity and raised awareness of the gradual deterioration of the cultural heritage within the areas studied (van der Leeuw 1998).

If landscape archaeology is understood in its most positivist way, the understanding of the built environment of a landscape should rest on a dataset of all features pertaining to human activity for a specific period in a specific region. The tool *du jour* of this type of landscape archaeology—Geographical Information Systems—thus relies heavily on the conversion of archaeological source material into tabular data (Verhagen 2018: 21). Here is an interface in which the 'translation' of archaeological source material into machine-readable tables is at its most apparent. For Roman archaeology, this could entail defining and classifying sites and, at best, dating the beginning and abandonment thereof with the aim of reconstructing the development of the built environment.

The creation of such a database is not without issues: imprecise coordinates, indistinct chronologies, and a lack of information on the sites themselves (size, internal layout, phasing, etc.) are stumbling blocks which must be overcome. In recent years, however, several projects have been completed showing not only the potential of large-scale and rich datasets but also their relevance for our understanding of the Roman Empire.

In 2016, the publication of the database of the Roman Rural Settlement of Britain, consisting of around 3,650 sites, shows how datasets for landscape archaeology could be constructed (Smith et al. 2016). The immanent publication of the Roman Hinterland Project Database (RHPdb) and the accompanying publications will show the potential and challenges of merged datasets from three different project regions (Attema et al. 2021; 2022). The combined dataset should yield in-depth and broad-ranging insights into urban-rural relations, demography, and economic integration between Rome and its Hinterland. It will be interesting to see if the collective value of the RHPdb is 'greater than the sum of its parts' (Attema 2021: 15) as the authors claim. While not resulting in an actual database, M. Reddé's (2017) double volume *Gallia Rustica* collected over 15 case studies from Northern Gaul and beyond. The project focused on the development and economic systems of the rural settlement in this area and showed how multidimensional and asynchronous the chronological trajectories of the Roman provincial rural settlement may have been.

Future directions for landscape archaeology are being cyclically debated. Ten years ago, Piraye Hacıgüzeller made the salient point that '[n]either walking across the landscape nor collecting the most precise spatial data will reveal a past reality that exists independently from us, the archaeologists' (Hacıgüzeller 2012: 257). Bigger and better datasets cannot and will not compensate for epistemological gaps, but they do help us design and render more complex models, which in turn can open up new research perspectives.

## Conclusion – 'The Most Exciting of Times?'

As we have seen, Roman archaeology has a long tradition of working with large datasets, but it is only recently that they have been more widely disseminated and more easily accessible. This cannot be put solely down to the development of the internet and file-sharing capabilities—in the last 10 years, a definite shift towards systematically implementing and archiving large datasets for research projects has taken place (Kristiansen 2014).

One of the main reservations regarding going 'all in' on big data may have been an uneasiness towards the difficulty of the underlying mathematical formulae in the analysis or out of fear of a loss of control over the interpretation of the source material. In a longitudinal study on data sharing practices in archaeobotanical research, Lisa Lodwick observes that from 239 journal articles, 56% share their primary data and proposes better training and facilitated access to data-sharing platforms as possible measures to improve the situation (Lodwick 2019).

In my opinion, scholarly introvertism and unwillingness to share work-in-progress may also be contributing factors for Roman archaeology in general. This is even more astounding as, due to the standardisation and sheer quantity of Roman material culture and architecture, especially from the Augustan period onwards, homogenous datasets are perhaps easier to come by than in the preceding periods (Pitts 2019: 4–7).

As with many digital projects, a certain cyclical pattern can be observed. An initial enthusiasm, then (often after the financing has stopped) a slower pace when it comes to the maintenance and upkeep of the database. Repositories of archaeological data should not become obsolete when the project funding runs out but should at least be carried forward for as long as possible, thus guaranteeing a more sustainable shelf life for the project data. New project applications with large datasets should and now in some cases must provide proposals for how the data is intended to be archived and reused (e.g. ERC 2022).

In my opinion, national repositories which could be periodically 'tidied-up' with deadwood being systematically cleared or even deleted are a possibility worth

exploring in the future. At the same time, there is a distinct possibility of structured data being forgotten or lost. A further challenge for archaeology in the next years will be the sustainable archiving and cataloguing of existing datasets so that these can be queried and reused by future generations. Online journals such as Open Archaeology, Internet Archaeology and indeed the Theoretical Roman Archaeology Journal are good platforms showing how an article and the accompanying data can be published.

Can we conclude that Roman archaeology is currently experiencing a big data revolution? The answer is yes: the data which is now being used in projects, papers, and monographs is considerably larger than it was 20–30 years ago. However, it is not only the volume which has exponentially increased; the variety of formats and shapes has also expanded significantly. Data can be found in existing databases, in appendixes, in online catalogues, in .csv lists, or PDF files. The heterogeneity of formats, measuring units, chronological frameworks, typologies, and languages make cross-regional and cross-temporal comparisons difficult, if not impossible. However, in cases where several projects are concentrated in one resource, the results are very encouraging (Bond, Dilley, and Horne 2021).

Perhaps we should reframe the question: is Roman archaeology making the most of the big data available? Possibly not, and this is where comparing the different subfields outlined in this paper can help.

The potential of Samian ware as an 'empire-wide' phenomenon has long been recognised, but large datasets and innovative methods—the implementation of ontologies—can now help operationalise extensive chronological and regional studies. Here, the potential is far from exhausted, and the next decade would do well to build upon the research being carried out.

Looking at the changes that numismatics has undergone, one can conclude that here, the most advances have been made towards transformation into a digital discipline, especially with regard to Linked Open Data. We should recall that early coin databases were started in 1988—34 years ago—and this timeframe shows that the sustainable digitalisation of archaeology is a marathon and not a sprint.

In zooarchaeology, the discussion on the standardisation of the recording of material and on the combining of existing datasets must be heeded. Here it has been shown that common terminologies needn't necessarily reflect the same recording practice. Variations from lab to lab or even from researcher to researcher must be expected and appropriate procedures, such as the documentation of the methods and the criteria used (i.e., information describing individual spreadsheet columns) should be put in place to mitigate against possible distortions in the datasets. Furthermore,

zooarchaeology has been using common ontologies to integrate disparate datasets for over a decade (Spielmann and Kintigh 2011). This field can thus boast a wealth of examples which could and should be consulted when other research frameworks are being designed.

However, in some places, a critical, almost diametrically opposed stance, has been taken, suggesting that the diversity of approaches within the discipline should not be sacrificed on the altar of data comparability. It will be interesting to see if and how future research on zooarchaeology can reconcile these two approaches.

A similar discussion on standardisation can also be observed in landscape archaeology. On one hand, data on archaeological sites is 'messy', heterogenous, and incomplete (Verhagen 2018: 21); on the other hand, efforts are being made to connect regional databases with the aim of offering supra-regional analyses of settlement dynamics. In the next years, it will be fascinating to observe how this data is aggregated, archived, and interpreted, which will set the tone for future projects on the study of Roman landscapes.

A further, more general problem is with the analytical units of databases themselves. There is growing concern regarding the lack of cross-specialist analyses. In 2017, D. Mattingly wrote of 'the danger of academic silos developing' within national research traditions (Mattingly 2017: 152). In the sub-fields presented here, a different kind of siloisation is still quite prevalent—one that restricts movement beyond one sole fixed unit of analysis. Projects which integrate data from different sub-fields are still in the minority. In this sense, the Rural Settlement of Roman Britain is again a good example, as many of the sites list bone assemblages, archaeobotanical remains, and coins (for a more detailed discussion on the data gathering processes and practices within the project's scope and how this fed into a synthesis see Fulford and Holbrook 2018).

This approach surely would be an avenue worth pursuing for other projects, as assemblages (i.e. the total sum of artefacts and ecofacts belonging to a defined archaeological structure) can yield more reliable information on the chronology and the individuals or groups connected to these structures (Mattingly 2014: 233). In my opinion, it would be a missed opportunity if database ontologies only used worked with typologies (some of which are over 100 years old) instead of fashioning new classifications and analytical units.

The data itself must also be retraceable, meaning that the primary source—be it an excavation report or a catalogue—must be cited. In my opinion, a dataset with unreliable or only rough references is hazardous especially when it becomes part of a Linked Open Data system.

It, however, is not only the data itself which must be subjected to rigorous source criticism. In his landmark book on the archaeology of knowledge from 1969, Michel Foucault lists a number of consequences that occur in the wake of the emergence of what he calls 'new history' (i.e. the study of so-called discourse formations within history). In his fourth consequence, Foucault makes the case not only for analysing the data itself but also for the choices which led to the data selection, the sampling strategy, the definition of the level of analysis, the categories and groups implemented, and the statistical methods used (Foucault 1969: 10–11). These statements could be transferred to big data studies, meaning that meta-data and para-data should accompany the publication of databases.[8] These ideas have also been put forward by Jeremy Huggett who, as regards to archaeological data, has made the case for documenting its 'circumstances of creation, recording, processing, and manipulation before it ever even becomes Big Data' (Huggett 2020: S14).

In this context, it is worth recalling the FAIR Principles which were published in 2016. These guidelines aim to improve the **F**indability, **A**ccessibility, **I**nteroperability, and **R**eusability of digital assets thus guaranteeing minimal, rigorous standards and protocols (GO FAIR Initiative 2017). New Big Data projects in Roman Archaeology may benefit from adhering to these guidelines and they would prolong the data's 'shelf-life' as it can be re-manipulated in other projects. From a practical point of view, this would mean the increased implementation of APIs (Application Programming Interface) in databases to facilitate the exchange of information. The best way of making sure that these measures are implemented into archaeological practice would be to anchor them within introductory modules at university, archaeological textbooks, and basic literature.

A further step would be to have databases peer-reviewed. Structure, usability, interoperability, and most importantly, data quality are criteria which in my opinion, must be scrutinised. This could also lead to greater visibility of the databases in the medium to long-term. The latest issue of *Germania* has collected four articles which take an in-depth look at the French-based database for archaeological objects artefacts.mom.fr (Feugère et al. 2021). This type of critique is important if we wish to advance the epistemological underpinnings of data-driven archaeology. Databases should not only be used as repositories for collected information but could also be used more explicitly as heuristic devices for ordering and operationalising research questions. A recent paper by Anwen Cooper and Chris Green in the framework of the 'English Landscape and Identities' project is a methodologically thorough example of how the database of the Portable Antiquities Scheme (consisting of 750,000 finds at the time of the paper's publication) was used not only to create a nation-wide map of potential archaeological

sites based on artefact distribution density but also to design chronologically sensitive graphs of artefact loss/deposition by region, thus establishing a novel approach for understanding archaeological patterning in the long term and over large areas (Cooper and Green 2017).

My final observation concerns the theoretical underpinnings of big data projects. This concern is not new, but it does bear repeating. To quote Matthew H. Johnson, 'Theory vs Data is not a zero-sum game'—an increase in data cannot replace a deficit in theoretical underpinning (Johnson 2006: 131; for an opposite opinion, see Anderson 2008). In the same vein, Sara Perry and James Stuart Taylor also make the important point that critical digital reflexivity exists within the discipline but has tended 'to get lost in wider discussions of the technicalities or presentation of results' (Perry and Taylor 2018: 15)—a statement which can be applied to Roman archaeology. They also ask for greater awareness from both theorists and digital archaeologists of the methodological and theoretical toolkits at their disposal.

How this might be achieved is hinted at by Shawn Graham, Ian Milligan, and Scott Weingart in their manifesto from 2015, in which they write that 'big data for the humanities is not only about justifying a story about the past, but generating new stories, new perspectives, given our new vantage points and tools' (Graham, Milligan, and Weingart 2016: xvi). This could mean using data not only to reinforce or reaffirm old narratives and patterns but also to identify, discuss, and propagate new ones (Storey 1999: 231–233; see also Brughmans et al. 2019: 14).

For Roman archaeology, big data research should not only lead to 'more of the same'. If recorded and structured intelligently, it might yield nuanced, differentiated, or narrative-challenging insights into many hitherto unknown or neglected aspects of the Roman world. That truly would be a revolution.

## Notes

[1] http://www.archcalc.cnr.it/journal/year_list.php (last accessed 25.10.2022).
[2] Cologne: https://archaeologie.phil-fak.uni-koeln.de/institut/fachgebiete/archaeoinformatik (last accessed 25.10.2022).
[3] York: https://www.york.ac.uk/study/postgraduate-taught/courses/msc-digital-archaeology/ (last accessed 25.10.2022).
[4] http://www.openscience.fr/Archeologies-numeriques (last accessed 25.10.2022).
[5] For space reasons, this chapter on Samian Ware. Other pottery genres will not be dealt with. For amphorae, see, for example, Laubenheimer 2018.
[6] In English: 'Ornamented ceramic vessels from Roman Gaul (Narbonne, Aquitaine and Lyonnaise).'
[7] For a further overview on quantitative methods in numismatics, see also Doyen 2011.
[8] https://statipedia.org/wiki/index.php?title=Paradata (last accessed 25.10.2022).

## Acknowledgements

This paper was written in the context of an Ambizione-Grant of the Swiss Science Foundation: 'Reassessing Roman Impact: Geographies of Power and Domination in the Roman West and Beyond' (PZ00P1_193441). I would like to thank the editorial board at the Theoretical Journal of Roman Archaeology for allowing me to publish in the journal and for all their assistance in the processes involved. I would also like to thank the two anonymous reviewers for their constructive criticism and feedback. Finally, I would like to thank Markus Peter (Römerstadt Augusta Raurica/Univeristät Bern) for his ideas on the early forays of numismatics with large datasets and James Dodd (Université Catholique de Louvain) for reading and commenting on an early draft. Any errors are my own.

## Competing Interests

The author has no competing interests to declare.

## References

ADS. 2017. The history of the ADS. Available at: https://archaeologydataservice.ac.uk/about/background.xhtml [Last accessed 14 November 2022].

Agache, R. 1978. *La Somme pré-romaine et romaine: d'après les prospections aériennes à basse altitude*. Amiens: Société des antiquaires de Picardie. DOI: https://doi.org/10.3406/pica.1977.1081

Albarella, U. 2017. Zooarchaeology in the twenty-first century: where we come from, where we are now, and where we are going. In: U. Albarella, M. Rizzetto, H. Russ, K. Vickers, and S. Viner-Daniels (eds) *The Oxford Handbook of Zooarchaeology*. Oxford: Oxford University Press: 3–22. DOI: https://doi.org/10.1093/oxfordhb/9780199686476.013.56

Allison, P., Pitts, M., and Colley, S. 2018. Big data on the Roman table. *Internet Archaeology* 50. DOI: https://doi.org/10.11141/ia.50.1

Anderson, C. 2008. The end of theory: the data deluge makes the scientific method obsolete. *Wired* 16: 108–109.

Arcelin, P. and Tuffreau-Libre, M. 1998. La quantification des céramiques: Conditions et protocol. *Actes de la Table Ronde du Centre Archéologique Européen du Mont Beuvray* (Glux-en-Glenne, 7-9 avril 1998). Glux-en-Glenne: Centre Archéologique Européen du Mont Beuvray.

Atici, L. et al. 2013. Other people's data: a demonstration of the imperative of publishing primary data. *Journal of Archaeological Method Theory* 20: 663–681. DOI: https://doi.org/10.1007/s10816-012-9132-9

Attema, P. et al. 2021. The Roman Hinterland Project: integrating archaeological field surveys around Rome and beyond. *European Journal of Archaeology* 25(2): 1–21. DOI: https://doi.org/10.1017/eaa.2021.51

Attema, P., de Haas, T., Tol, G., and Seubers, J. 2022. Towards an integrated database for the study of long-term settlement dynamics, economic performance and demography in the Pontine region and the Hinterland of Rome. In: *Archaeology and Economy in the Ancient World: The Rural Foundations of The Roman Economy. New Approaches to Rome's Ancient Countryside from the Archaic to the Early Imperial Period* 35–53 (Proceedings of the 19th International Congress of Classical Archaeology), Cologne/Bonn, 22–26 May 2018: 35–53.

Baxter, M. and Cool, H. 2016. *Basic Statistical Graphics for Archaeology with R: Life Beyond Excel*. Nottingham: Barbican Research Associates.

Bevan, A. 2015. The data deluge. *Antiquity* 89(348): 1473–1484. DOI: https://doi.org/10.15184/aqy.2015.102

Bickler, S. 2021. Machine learning arrives in archaeology. *Advances in Archaeological Practice* 9(2): 186–191. DOI: https://doi.org/10.1017/aap.2021.6

Binford, L.R. 1962. Archaeology as anthropology. *American Antiquity* 28(2): 217–225. DOI: https://doi.org/10.2307/278380

Bintliff, J.L. 1991. *The 'Annales' School and Archaeology*. Leicester: Leicester University Press.

Bolin, S. 1926. Fynden av romerska mynt I det fria Germanien. *Studier I Romersk Och älder Germansk Historia*. Lund: C.W. Lindström.

Bolin, S. 1929. Die Funde römischer und byzantinischer Münzen im freien Germanien. *Berichte der römisch-germanischen Kommission* 19: 86–145.

Bond, S.E., Dilley, P., and Horne R. 2021. Linked open data for the ancient Mediterranean: structures, practices, prospects. *Institute for the Study of the Ancient World Papers* 20. Available at http://hdl.handle.net/2333.1/gqnk9kz2 [last accessed 14 November 2022].

Bowden, M. 2001. Mapping the past: O.G.S. Crawford and the development of landscape studies. *Landscapes* 2(2): 29–4. DOI: https://doi.org/10.1179/lan.2001.2.2.29

Boyd, D. and Crawford K. 2012. Critical questions for big data. Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society* 15(5): 662–679. DOI: https://doi.org/10.1080/1369118X.2012.678878

Brather, S. 2011. Siedlungs-, Umwelt- und Landschaftsarchäologie: Entdeckung und Analyse des Raumes. *Freiburger Universitätsblätter* 50/Heft 192: 123–136.

Brughmans, T. et al. 2019. Formal modelling approaches to complexity science in Roman studies: a manifesto. *Theoretical Roman Archaeology Journal* 2(1): 4. DOI: https://doi.org/10.16995/traj.367

CHREP. 2022. Coin Hoards of the Roman Empire. Available at https://chre.ashmus.ox.ac.uk [Last accessed 14 November 2022].

Clarke, D. 1973. Archaeology: The loss of innocence. *Antiquity* 47(185): 6–18. DOI: https://doi.org/10.1017/S0003598X0003461X

Cooper, A. and Green, C. 2016. Embracing the complexities of 'big data' in archaeology: the case of the English Landscape and Identities project. *Journal of Archaeological Method and Theory* 23(1): 271–304. DOI: https://doi.org/10.1007/s10816-015-9240-4

Cooper, A. and Green, C. 2017. Big questions for large, complex datasets: approaching time and space using composite object assemblages. *Internet Archaeology* 45. DOI: https://doi.org/10.11141/ia.45.1

Couisson, P. 1931. *Carte archéologique de la Gaule romaine F.O.R. Partie orientale du département des Alpes-Maritimes*, fasc. I. Paris: E. Leroux

Cristofani M. and Francovich, R. 1990. Editoriale. *Archaeologia e Calcolatori*: 1: 7–8.

Crummy, N. 1981. Bone-working at Colchester. *Britannia* 12: 277–285. DOI: https://doi.org/10.2307/526252

DARIAH-EU 2012. 'ARIADNE'. Available at https://www.dariah.eu/activities/projects-and-affiliations/ariadne/ [Last accessed 14 November 2022].

DARIAH-EU 2018, 'Digital Numismatics'. Available at https://www.dariah.eu/activities/working-groups/wg-digital-numismatics/ [Last accessed 14 November 2022].

Davies, R. 1971. The Roman military diet. *Britannia* 2: 122–41. DOI: https://doi.org/10.2307/525803

Déchelette, J. 1904. *Les vases céramiques ornés de la Gaule romaine (Narbonnaise, Aquitaine et Lyonnaise)*. Paris: A. Picard.

Desbat, A. 1990. Les bons comptes font les bons amis ou la quantification des céramiques. In: L. Rivet (ed.) *Actes du Congrès de Mandeure-Mathay*, 24–27 mai 1990. Marseille: Société française d'étude de la céramique antique en Gaul: 131–134.

Diebolt, C. 2011. The stakes of cliometrics in ancient history. *Historical Social Research / Historische Sozialforschung* 36(137): 350–361.

Diebolt, C., and Haupert, M. 2021. Cliometrics: Past, Present, and Future. In: J. Hamilton (ed.) *Oxford Research Encyclopedia of Economics and Finance*. Oxford: Oxford University Press: 1–20. DOI: https://doi.org/10.1093/acrefore/9780190625979.013.552

Djindjian, F. 2015. A short history of the beginnings of mathematics in archaeology. In: J.A. Barcelo and I. Bogdanovic (eds) *Mathematics, and Archaeology*. Boca Raton, FL: CRC Press: 65–85. DOI: https://doi.org/10.1201/b18530-4

Doyen, J-M. 2011. Cliométrie et numismatique contextuelle : compter et quantifier le passé? Petite histoire de la méthode (1960–2011). *The Journal of Archaeological Numismatics* 1: 9–46.

Doyen, J.-M. 2019. 'Big is beautiful?' Faut-il VRAIMENT étudier les ‹mégadépôts› monétaires?. In: B. Callegher (ed.) *Too Big to Study? Troppo grandi da studiare*? Trieste: EUT Edizioni Università di Trieste: 305–321.

Dragendorff, H. 1895. *Terra sigillata. Ein Beitrag zur Geschichte der griechischen und römischen Keramik*. Bonn: Verein von Altertumsfreunden im Rheinlande.

Duval, P.-M. 1989. Avant-Propos. Travaux sur la Gaule. *Publications de l'École Française de Rome* 116(1): 13–15.

Dyson, S.L. 1981. A classical archaeologist's response to the 'new archaeology'. *Bulletin of the American Schools of Oriental Research* 242: 7–13. DOI: https://doi.org/10.2307/1356543

Eggers, H. J. 1959. *Einführung in die Vorgeschichte*. München: R. Piper.

Ehret, R. 1964. *Tierknochenfunde aus der Stadt auf dem Magdalensberg bei Klagenfurt in Kärnten*: *2. Carnivora, Lagomorpha, Rodentia und Equidae.* Klagenfurt: Verlag des Landesmuseums für Kärnten.

ERC 2022. European Research Council. Scientific Coucil. *Open Research Data and Data Management Plans. Information for ERC Grantees.* European Comission: 1–23.

Ettlinger, E. 1971. Review of: Fouilles de l'Ecole française de Rome à Bolsena (Poggio Moscini), 1962–1967: 4. La céramique arétine lisse. *Germania* 49: 264–271.

Feugère, M. et al. 2021. Typologie 2.0 – Datenbanken in der Archäologie: das europäische Projekt Artefacts.mom.fr. *Germania* 98: 194–236. DOI: https://doi.org/10.11588/ger.2020.85216

Fox, C. 2010. *The Archaeology of the Cambridge Region: A Topographical Study of the Bronze, Early Iron, Roman and Anglo-Saxon Ages, with an Introductory Note on the Neolithic Age.* Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511697739

Fulford, M. and Holbrook, N. 2011. Assessing the contribution of commercial archaeology to the study of the Roman period in England, 1990-2004. *Antiquaries Journal* 91: 323–345. DOI: https://doi.org/10.1017/S0003581511000138

Fulford, M. and Holbrook, N. 2018. Relevant beyond the Roman period: approaches to the investigation, analysis and dissemination of archaeological investigations of the Rural Settlements and Landscapes of Roman Britain. *Archaeological Journal* 175(2): 214–230. DOI: https://doi.org/10.1080/00665983.2017.1412093

Garland, N. 2021. TRAC at 30: A bibliometric analysis of the TRAC community. *Theoretical Roman Archaeology Journal* 4(1): 1. DOI: https://doi.org/10.16995/traj.4344

Gifford-Gonzalez, D. 2018. *An Introduction to Zooarchaeology*. (1st ed). Cham: Springer International Publishing. DOI: https://doi.org/10.1007/978-3-319-65682-3

GO FAIR Initiative. 2017. FAIR Principles. Available at https://www.go-fair.org/fair-principles/ [Last accessed 14 November 2022].

Graham, S., Milligan, I., and Weingart, S. 2016. *Exploring Big Data. The Historian's Macroscope*. London: Imperial College Press.

Greene, K. 2006, Archaeological data and economic interpretation. In: P.F Bang, M. Ikeguchi, and E. Ziche (eds) *Ancient Economies, Modern Methodologies: Archaeology, Comparative History, Models and Institutions*. Bari: Edipuglia: 109–136.

Gruber, E., Bransbourg, G., Heath, S., and Meadows, A. 2014. Linking Roman coins: current work at the American Numismatic Society. In: G. Earl et al. (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: Amsterdam University Press: 249–258. DOI: https://doi.org/10.1515/9789048519590-027

Habermehl, K.-H. 1959/1960. Die Tierknochenfunde im römischen Lagerdorf Butzbach. *Saalburg-Jahrbuch* 18: 67–108.

Habermehl, K.-H. 1962. Die Tierknochenfunde des augusteischen Lagers Rödgen, Kreis Friedberg, Hessen. *Saalburg-Jahrbuch* 20: 46–63.

Hacıgüzeller, P. 2012. GIS, critique, representation and beyond. *Journal of Social Archaeology* 12(2): 245–263. DOI: https://doi.org/10.1177/1469605312439139

Hacıgüzeller, P., Taylor, J., and Perry, S. 2021. On the emerging supremacy of structured digital data in archaeology: a preliminary assessment of information, knowledge and wisdom left behind. *Open Archaeology* 7(1): 1709–1730. DOI: https://doi.org/10.1515/opar-2020-0220

Heising, A. Unpubl. Forschungsgeschichte, aktuelle Situation und Zukunftsstrategien der Provinzialrömischen Archäologie in Deutschland.

Hodder, I. and Hassall, M. 1971. The non-random spacing of Romano-British walled towns. *Man* 6(3): 391–407. DOI: https://doi.org/10.2307/2799028

Hopkins, K. (1980). Taxes and trade in the Roman empire (200 B.C.–A.D. 400). *Journal of Roman Studies* 70: 101–125. DOI: https://doi.org/10.2307/299558

Huggett, J. 2015. A manifesto for an introspective digital archaeology. *Open Archaeology* 1(1): 86–95. DOI: https://doi.org/10.1515/opar-2015-0002

Huggett, J. 2020. Is big digital data different? Towards a new archaeological paradigm. *Journal of Field Archaeology* 45: sup1, S8-S17. DOI: https://doi.org/10.1080/00934690.2020.1713281

Johnson, M. 2006. On the nature of theoretical archaeology and archaeological theory. *Archaeological Dialogues* 13(2): 117–132. DOI: https://doi.org/10.1017/S138020380621208X

Jongman, W. (2014). The new economic history of the Roman Empire. In: F. de Callataÿ (ed.) *Quantifying the Greco-Roman Economy and Beyond*. Bari: Edipuglia: 169–188.

Jourdan, L. (1976). *La faune du site gallo-romain et paléo-chrétien de la bourse (Marseille): Espèces domestiques et espèces sauvages : l›élevage et l'alimentation à Marseille du IIème au Vème siècle*. Paris: C.N.R.S.

Kaltenthaler, D. and Lohrer, J.-Y. 2018. *The Historic Development of the Zooarchaeological Database OssoBook and the xBook Framework for Scientific Databases*. arXiv. DOI: https://doi.org/10.1007/978-3-319-92270-6_9

Kamash, Z. 2021. Rebalancing Roman archaeology: from disciplinary inertia to decolonial and inclusive action. *Theoretical Roman Archaeology Journal* 4(1): 4. DOI: https://doi.org/10.16995/traj.4330

Kenrick, P. 2014. The OCK Database. Available at https://blog.fautores.org/the-ock-database/ [Last accessed 14 November 2022].

King, A.C. 1999. Diet in the Roman world: a regional inter-site comparison of the mammal bones. *Journal of Roman Archaeology* 12(1): 168–202. DOI: https://doi.org/10.1017/S1047759400017979

King, A. 2005. Animal remains from temples in Roman Britain. *Britannia* 36: 329–369. DOI: https://doi.org/10.3815/000000005784016964

Kitchin, R. and McArdle, G. 2016. What makes big data, big data? Exploring the ontological characteristics of 26 datasets. *Big Data & Society* 3(1). DOI: https://doi.org/10.1177/2053951716631130

Klein, L.F. and Gold, M.K. (eds). 2019. *Debates in the digital humanities 2019*. Baltimore, Maryland: Project Muse.

Kortüm, K. 1998. Zur Datierung der militärischen im obergermanischen-rätischen Gebiet: Chronologische Untersuchung anhand der Münzfunde. *Saalburg-Jahrbuch* 49(1): 5–65.

Kraft, K. 1955/56. Das Enddatum des Legionslagers Haltern. *Bonner Jahrbücher* 155/56(1): 95–111.

Kristiansen, K. 2014. Towards a new paradigm? The third science revolution and its possible consequences in archaeology. *Current Swedish Archaeology* 22(1): 11–34. DOI: https://doi.org/10.37718/CSA.2014.01

Laflin, S. 1982. Computer Applications in Archaeology 1982. Conference Proceedings. Birmingham: Centre for Computing and Computer Science.

Laflin, S. 2013. *Computer Applications in Archaeology 1973–1995*. Lulu.com.

Lau, H. and Kansa, S.W. 2018. Zooarchaeology in the era of big data: contending with interanalyst variation and best practices for contextualizing data for informed reuse. *Journal of Archaeological Science* 95: 33–39. DOI: https://doi.org/10.1016/j.jas.2018.03.011

Laubenheimer, F. 2018. Comptages, comparaisons, évaluations: l'exemple des amphores en Gaule. In: J. Remesal Rodríguez, V. Revilla Calvo, and J.M. Bermúdez Lorenzo (eds) *Cuantificar las economías antiguas: Problemas y métodos* (1st ed). Barcelona: Universitat de Barcelona Edicions: 57–70.

Leveau P. 1984. *Caesarea de Maurétanie: une ville romaine et ses campagnes*. Rome: Ecole française de Rome.

Lloyd, J. 1991. Forms of rural settlement in the early Roman Empire. In: G. Barker and J. Lloyd (eds) *Roman Landscapes: Archaeological Survey in the Mediterranean Region*. London: British School at Rome: 233–238.

Lo Cascio, E. and Malanmina, P. 2014. Ancient and pre-modern economies GDP in Roman Empire and early modern Europe. In: F. de Callataÿ (ed.) *Quantifying the Greco-Roman Economy and Beyond*. Bari: Edipuglia: 229–251.

Lockyear, K. and Gruber, E. 2013. From DbaseIII+ to the semantic web: twenty-five years of the coin hoards of the Roman Republic database. In: A. Traviglia (ed.) *Proceedings of the Computer Applications and Quantitative Methods in Archaeology Conference*, Perth on 25–28 March 2013: 336–346.

Lodwick, L. 2019. Sowing the seeds of future research: data sharing, citation and reuse in archaeobotany. *Open Quaternary* 5(1): 7: 1–15. DOI: https://doi.org/10.5334/oq.62

Malone, C. and Stoddart, S. 1998. Special section: David Clarke's 'Archaeology: the loss of innocence' (1973) 25 years after. *Antiquity* 72(277): 676–677. DOI: https://doi.org/10.1017/S0003598X00087081

Maltby, M. 2019. From the appendix to integration? A review of the contribution of zooarchaeology to Romano-British studies since 1970. In: M.G. Allen (ed.) *The Role of Zooarchaeology in the Study of the Western Roman Empire*. Portsmouth: Journal of Roman Archaeology: 11–46.

Mattingly, D.J. 2014. *Imperialism, Power, and Identity: Experiencing the Roman Empire*. Princeton, NJ: Princeton University Press. DOI: https://doi.org/10.23943/princeton/9780691160177.001.0001

Mattingly, D.J. 2017. Conclusion and final discussion: a view from the Far South. In: S. González Sánchez and A. Guglielmi (eds) *Romans and Barbarians Beyond the Frontiers: Archaeology, Ideology and Identities in the North*. Oxford: Oxbow Books: 152–156.

Mazzilli, F. and Aldred, O. 2022. *What happens when we dig big: Recent work on huge datasets in Roman archaeology*. Available at http://trac.org.uk/2021/11/new-call-for-papers-workshops-trac-2020-postponed-2022-split/?fbclid=IwAR2Lkry2gNG5Nk4s-R7qxd9pxun6yiKJZsMrzGwxxp0VwB78XY_rJbvHC30#4-what-happens-when-we-dig-big [Last accessed 25 October 2022].

Méniel P. et al. 2009. Une base de données sur les études de faunes des établissements ruraux en Gaule. In: I. Betrand et al. (eds) *Habitats et paysages ruraux en Gaule et regards sur d'autres régions du monde celtique. (actes XXXIe coll. AFEAF, Chauvigny, mai 2007)*. Chauvigny: Association des Publications Chauvinoises: 417–446.

Millett, M. 2015. Roman Britain since Haverfield. In: M. Millett, L. Revell, and A. Moore (eds) *The Oxford Handbook of Roman Britain*. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/oxfordhb/9780199697731.013.003

Moscati, P. 2016. Jean-Claude Gardin and the evolution of archaeological computing. *Les nouvelles de l›archéologie* 144: 10–13. DOI: https://doi.org/10.4000/nda.3457

Moschek, W. 2011. *Der römische Limes: Eine Kultur- und Mentalitätsgeschichte*. Speyer: Kartoffeldr.-Verl.

Nieto-Espinet, A. et al. 2021. Resilience and livestock adaptations to demographic growth and technological change: a diachronic perspective from the Late Bronze Age to Late Antiquity in NE Iberia. *PLOS ONE* 16(2): e0246201. DOI: https://doi.org/10.1371/journal.pone.0246201

Orton, C. 1980. *Mathematics in Archaeology.* London: Collins.

Oxé, A. and Comfort, H. 1968. *Corpus Vasorum Arretinorum: A Catalogue of the Signatures, Shapes and Chronology of Italian Sigillata*. Bonn: R. Habelt Verlag.

Oxé, A., Comfort, H. and Kenrick, P. 2000. *Corpus Vasorum Arretinorum: A Catalogue of the Signatures, Shapes and Chronology of Italian Sigillata* (2nd ed). Bonn: R. Habelt Verlag.

OxRep 2012 The OXREP Databases - The Oxford Roman Economy Project Available at; http://oxrep.classics.ox.ac.uk/databases/ [Last accessed 14 November 2022].

Perry, S. and Taylor, J. 2018. Theorising the digital: a call to action for the archaeological community. In: M. Matsumoto and E. Uleberg (eds) *Oceans of Data: Proceedings of the 44th Conference on Computer Applications and Quantitative Methods in Archaeology.* Oxford: Archaeopress: 11–22.

Pinon, P. 2015. La découverte de la céramique sigillée gallo-romaine : une approche moderne dès les xviiie et xixe siècles. In: A. Fenet and N. Lubtchansky (eds) *Pour une histoire de l'archéologie xviiie siècle - 1945: Hommage de ses collègues et amis à Ève Gran-Aymerich*. Bordeaux: Ausonius Éditions. DOI: https://doi.org/10.4000/books.ausonius.5828

Pitts, M. 2019. *Roman Object Revolution: Objectscapes and Intra-Cultural Connectivity in Northwest Europe*. Amsterdam: Amsterdam University Press. DOI: https://doi.org/10.1017/9789048543878

Project Mercury. 2018. Computational Modelling in Roman Studies - Datasets. Available at https://projectmercury.eu/datasets/ [Last accessed 14 November 2022].

RDF Working Group. 2014. Research Description Framework. Available at https://www.w3.org/RDF/ [Last accessed 14 November 2022].

Reddé, M. 2017. *Gallia rvstica: les campagnes du nord-est de la Gaule, de la fin de l'âge du Fer à l'Antiquité tardive*. Bordeaux: Ausonius éditions.

Reece, R. 1993. Theory and Roman archaeology. In: E. Scott (ed.) *Theoretical Roman Archaeology: First Conference Proceedings*. Aldershot: Avebury: 28–38. DOI: https://doi.org/10.16995/TRAC1991_29_38

Reece, R. and Casey, J. 1974. *Coins and the Archaeologist*. Oxford: British Archaeological Reports.

Richards, J.D. 2017. Twenty years preserving data: a view from the United Kingdom. *Advances in Archaeological Practice* 5(3): 227–237. DOI: https://doi.org/10.1017/aap.2017.11

Robert, S. 2017. Archéologie du paysage et géographie: entre observation, transferts et co-construction. *Palethnologie* 9. DOI: https://doi.org/10.4000/palethnologie.297

Samian Research. 2016. Names on Terra Sigillata. Corpus Vasorum Arretinorum / OCK. Available at https://www1.rgzm.de/samian/home/frames.htm. [Last accessed 14 November 2022].

Schibler, J. 1998. OSSOBOOK: A database system for archaeozoology. In: P. Anreiter, L. Bartosiewicz, E. Jerem, and W. Meid (eds) *Man and the Animal World*. Budapest: Archaeolingua Alapítvány: 491–510.

Schmidt, S.C., Thiery, F., and Trognitz, M. 2022. Practices of linked open data in archaeology and their realisation in Wikidata. *Digital* 2(3): 333–364. DOI: https://doi.org/10.3390/digital2030019

Schnapp, A. 1981. Les Annales et l'archéologie: une rencontre difficile. *Mélanges de l›École française de Rome, Antiquité* 93(1): 469–478. DOI: https://doi.org/10.3406/mefr.1981.1281

Smith, A., Allen, M., Brindle, T., and Fulford, M. 2016. *New Visions of the Countryside of Roman Britain*, vol 1. London: Society for the Promotion of Roman Studies.

Spielmann, K. and Kintigh, K. 2011. The digital archaeological record: the potentials of archaeozoological data integration through tDAR. *The SAA Archaeological Record* 11(1): 22–25.

Storey, G. R. 1999. Archaeology and Roman society: integrating textual and archaeological data. *Journal of Archaeological Research* 7(3): 203–248. DOI: https://doi.org/10.1007/BF02446112

tDar 2011. The digital archaeological record. A service of digital antiquity. Available at https://www.tdar.org [Last accessed 14 November 2022].

Terrenato, N. 2002. The innocents and the sceptics: ANTIQUITY and classical archaeology. *Antiquity* 76(294): 1104–1111. DOI: https://doi.org/10.1017/S0003598X00091997

Theune, C. 2006. Nachwort. In: H.J. Eggers and C. Theune (eds) *Einführung in die Vorgeschichte: Mit dreissig Karten und Bildtafeln (5. Auflage.)*. Berlin: scrîpvaz-Verlag Christof Krauskopf: 322–330.

Thiery, F., Mees, A., and Gottwald, D. 2020. Linked open Samian Ware – unveiling the hidden data dragons and uncovering temporal vagueness with the help of little minions. In: *Joint Chapter Meeting 2020 of Computer Applications and Quantitative Methods in Archaeology by the CAA Chapters CAA-NL-FL and CAA-DE (JCM2020)*. [last accessed 11 November 2022]. DOI: https://doi.org/10.5281/zenodo.4306642

Trigger, B.G. 1967. Settlement archaeology: its goals and promise. *American Antiquity* 32(2): 149–160. DOI: https://doi.org/10.2307/277900

van der Leeuw, S. 1998. Archaeomedes, un programme de recherces européen sur la désertification et la dégradation des sols. *Nature Sciences Sociétés* 6(4): 53–58. DOI: https://doi.org/10.1016/S1240-1307(98)80142-8

Van Valkenburgh, P. and Dufton. J.A. 2020. Big archaeology: horizons and blindspots. *Journal of Field Archaeology* 45, Supplement Issue 1: S1–S7. DOI: https://doi.org/10.1080/00934690.2020.1714307

Verhagen P. 2018. Spatial analysis in archaeology: moving into new territories. In: C. Siart, M. Forbriger, and O. Bubenzer (eds) *Digital Geoarchaeology. Natural Science in Archaeology*. Springer: Cham. DOI: https://doi.org/10.1007/978-3-319-25316-9_2

Verhagen, P. 2022. *Simulating the Limes*. Available at https://limes2022.org/timetable/31-simulating-the-limes/ [Last accessed 2 March 2022].

Von Carnap-Bornheim, C. 2001. Hans Jürgen Eggers und der Weg aus der Sackgasse der ethnischen Deutung. In: D. Hakelberg and H. Steuer (eds) *Eine hervorragend nationale Wissenschaft: deutsche Prähistoriker zwischen 1900 und 1995*. Berlin: W. de Gruyter: 173–197.

Wigg-Wolf, D. 2018. 9 or not 9?: New thoughts on Waldgirmes, Haltern, Kalkriese and co. In: S. Burmeister and S. Ortisi (eds) *Phantom Germanicus: Spurensuche zwischen historischer Überlieferung und archäologischem Befund*. Rahden/Westf: VML Verlag Marie Leidorf GmbH: 237–252.

Williamson, S. and Whaples, R. 2003. Cliometrics. In: J. Mokyr (ed.) *The Oxford Encyclopedia of Economic History*. Oxford: Oxford University Press.

Woolf, G. 2004. The present state and future scope of Roman archaeology: a comment. *American Journal of Archaeology* 108(3): 417–428. DOI: https://doi.org/10.3764/aja.108.3.417

Zabehlicky-Scheffenegger, S. 2013. Historia Fautorum. A (partly personal) illustrated history of the RCRF. Available at https://www.fautores.org/pages/historia-fautorum.htm [Last accessed 14 November 2022].

ZoomWest 2017. Zooarchaeology and Mobility in the Western Mediterranean: Animal husbandry from the Late Bronze Age to Late Antiquity. Available at https://zoomwest11.wixsite.com/zoomwest [Last accessed 14 November 2022].