Documenting heritage in the 21st century: the EAMENA project and its potential for 'big data' research in Levantine archaeology

Letty ten Harkel,[1] Michael Fradley,[1] Pascal Flohr[2], Azadeh Vafadari,[3] Sayantani Neogi[1]

Introduction

The Endangered Archaeology in the Middle East and North Africa (EAMENA) project (www.eamena.org) is among the most ambitious archaeological documentation projects ever to have taken place across the Middle East and North Africa. The project, under the current directorship of Bill Finlayson (2015–21 director Robert Bewley), is funded by the Arcadia Fund [Arcadia, a charitable fund of Lisbet Rausing & Peter Baldwin, grant number 4178]. It started as a collaboration between the universities of Oxford (PI: Andrew Wilson) and Leicester (Co-PI: David Mattingly), with Durham joining shortly afterwards (Co-PI: Graham Philip). In addition, extensive partnerships exist across the MENA region that were strengthened as a result of a series of training workshops funded by the UK's Cultural Protection Fund (CPF), administered by the British Council on behalf of the Department of Digital, Culture, Media and Sport. At the core of EAMENA's methodology is archaeological survey, primarily using remote sensing sources (satellite imagery and aerial photographs), stretching from Mauretania to Iran (Fig. 1). Funding for the project was obtained in the aftermath of the 2010–2011 Arab Spring and during the rise of the so-called Islamic State group, which resulted in deliberate destruction of cultural heritage as well as damage through, for example, neglect following political instability. As travel (and therefore reliance on field survey) to some parts of the MENA region had become more restricted or impossible, the remote sensing methodology of the EAMENA project offered an alternative approach.

Fig. 1: Map of EAMENA study area. Background imagery © Bing; external country boundaries forming the edges of the study area from diva-gis.org.

In addition to fulfilling an urgent heritage management need, the overarching research vision of the EAMENA project is to take the scale of archaeological analysis in the MENA region to a new level in terms of big data. As such, it fulfils a major aim of the Arcadia Fund, which is primarily interested in the documentation of endangered aspects of the natural and cultural world and the promotion of open-access information to preserve knowledge (https://www.arcadiafund.org.uk/). This vision fits perfectly with the general trend towards the generation of large digital datasets that lies at the heart of the emergence of big data approaches over the last two decades.

[1] School of Archaeology, University of Oxford.
Corresponding author: Letty ten Harkel, 1-2 South Parks Road, Oxford OX1 3TG, letty.tenharkel@arch.ox.ac.uk.
[2] Institute for Pre- and Proto-History and Excellence Cluster ROOTS, Christian Albrechts University of Kiel; School of Archaeology, University of Oxford.
[3] McDonald Institute for Archaeological Research, University of Cambridge.

The collection of papers in this Special Issue is the result of an initial exploration of the research potential contained within the project's open access Arches database. The EAMENA project builds on the success of previous remote-sensing research projects in the region — for example, the Fragile Crescent (e.g., Galiatsatos *et al*. 2009; Wilkinson *et al.* 2014) and the Trans-Saharan (e.g., Mattingly and Sterry 2013; Mattingly *et al.* 2013) projects — and other data collation initiatives across the globe that attempt to make sense of the ever-increasing mass of archaeological information that is available today. The increase in data quantity globally in the current 'Age of Information' is widespread. Within the discipline of archaeology, contributing factors include, but are not restricted to, improved heritage legislation and increased awareness of the importance of safeguarding knowledge of the past for future generations, which was further heightened during the period of political unrest in parts of the MENA region following the Arab Spring. The EAMENA project stands out for its supra-national perspective and crosses modern political boundaries, fully embracing the challenge of disentangling archaeological patterns from many different contexts of data generation.[4] At the same time, it collaborates closely with stakeholders in the MENA region through training and knowledge sharing.

Since its inception, EAMENA has had two main aims. The first of these is research: by recording all (possible) archaeological and heritage sites and by making this information freely available through its open access database, EAMENA aims to encourage informed debate on multiple spatial levels, ranging from local to international. By recording not only heritage resources (following EAMENA's definition, including much more than only 'listed' and/or protected sites; see below) but also linked information resources (publications and image material that provide information about the heritage resources), interpretations can be checked. Importantly, the project has adopted the recent trend towards landscape-based approaches, moving away from older site-based approaches that favoured important 'flagship' sites, which were often discussed with little reference to their surrounding landscapes.

The second aim is mitigation of threats to heritage, which the project aims to achieve — in collaboration with archaeologists and heritage professionals from the MENA region — through the enhancement (or, where previously absent, the creation) of comprehensive records of sites and monuments in MENA countries, which can be used to identify, understand and monitor the endangered archaeology of the region. To further the goal of mitigating threats to heritage, a condition assessment (including present and past disturbances as well as future threats) is also added to each archaeological assessment. Through a series of training workshops in remote sensing and digital documentation, EAMENA has furthermore contributed to the training and capacity building of heritage professionals in the region, creating international networks to promote the protection and conservation of MENA's archaeological heritage.

One of the unique strengths of the EAMENA database is thus that, unlike many other archaeological and heritage databases, it captures information that can be used *both* for research (e.g. focused on analysing detailed chronological and spatial patterns in the data) *and* heritage management (focused on managing and protecting existing sites in their landscape context). Although the two are (and should be) inherently related, this Special Issue is concerned primarily

[4] Since the initial success of the EAMENA project, its funder Arcadia has granted similar projects in different regions; see for example CAAL (Central Asia); MAHSA (South Asia); MAEASaM (sub-Saharan Africa); MarEA (MENA region, maritime); MAHS (South Asia, maritime). The EAMENA project remains one of the larger projects of its kind.

with the first aim — the research potential of this powerful resource — and includes six papers each representing an initial exploration of the emergent datasets, focusing on the Levant.

The EAMENA project aims to record all heritage and archaeological sites (in database terminology, 'heritage resources') predating *c*. 1950 CE, ultimately facilitating large-scale, comprehensive 'big data' analyses of the heritage and archaeology of the MENA region. As such, the EAMENA project exists at the cutting edge of archaeological knowledge production in the 21st century, even if it is still very much a work in progress. At the time of writing, the EAMENA database still contained less than 200,000 heritage records and generally not more than several thousand in any well-defined area that can serve as a case-study area. Too many gaps in its spatial coverage currently persist to enable the kinds of large-scale comparative analyses that it ultimately aims to facilitate. The initial explorations of the database's research potential presented in this Special Issue are therefore perhaps better understood as analyses of 'medium' rather than 'big' data (Flohr *et al*. 2021).

The significance of big data approaches for archaeological research — and the important role of the EAMENA project in driving big data research for the MENA region — will be explored in more detail in the next section, but one thing is important to emphasize at the outset. 'Big' data are not merely the same as 'a lot' of data; big data are also varied and dynamic, in the sense that they are generated continuously and structured flexibly so that new fields can be added easily and datasets can expand rapidly (Kitchin 2014: 2; also see Andreou *et al.* forthcoming; Cooper *et al.* 2021b; Cooper and Green 2016). In this respect, and even though it does not yet meet the quantitative criterion, the EAMENA database nonetheless fits the big data bill: it is very much a still-growing or 'living' record of archaeological information (Ten Harkel and Fisher 2021) and, as long as the project continues, will continue to do so with no predefined endpoint in sight.

The EAMENA database was developed in a UK-based academic environment, by a group of researchers with backgrounds in UK academic research and archaeological practice (Ten Harkel and Bewley 2018), albeit since 2017 in formal collaboration with heritage professionals in the MENA region. To acknowledge the research environment and intellectual context of the project, this introduction includes a brief overview of the history of UK-based initiatives aiming to document digital information about archaeological and heritage sites on a large scale, which can be traced back to the 19th century. In doing so, we explicitly acknowledge the continuation of relationships that were established between Europe and the MENA region during the colonial era, but with the aim — through information sharing — to avoid the continuation of inequality.

The processes and decisions that shaped the formation of the EAMENA dataset, including its methodology and the structure of the database, will be discussed next. An important milestone, achieved in 2017, was a large grant from the CPF to provide training in the project methodology and database, allowing the networks between the UK-based researchers and partners in the MENA region to come to fruition. Collaboration with MENA colleagues also led to improvements in database terminology and functionality, including the ability to record information on smaller (site component) levels (e.g., Mubaideen *et al.* 2021). A similar effect occurred after the start of collaboration with the Maritime Endangered Archaeology (MarEA) project, founded in 2019, whose emphasis on maritime heritage led to a significant expansion of the database glossaries and the development of new functionality to record palaeolandscapes (Andreou *et al.* 2020).

The final part of this chapter introduces the various papers that make up this Special Issue, starting with those that have helped to shape the project's most important achievement to date —

the database itself — and continuing with a few initial explorations of the research potential of the emerging dataset. All of these were written before this introductory chapter, and some were already published online when the ideas presented here were still being developed. Together, they have helped shape the vision of the EAMENA project in its current form, which Andrew Wilson reflects on more broadly in his Conclusion to this Special Issue.

The new age of archaeological big data

The EAMENA methodology is grounded in big data in several ways. On the one hand, the satellite imagery and aerial photos it uses for its analyses are big data in their own right. On the other hand, the EAMENA project exploits this big data resource to create another big dataset that arguably represents a form of intermediary big data, one that shares many characteristics with big data in terms of its abundance, complexity and dynamism. As stated above, we subscribe to a definition of big data that combines large volume with speed and the ongoing nature of data recording,[5] diversity in data structure, near-exhaustiveness in terms of scope and/or ambition, and which is fine-grained, relational and flexible, and generally too large and complex to be made sense of without specialist computer applications including databases, GIS packages and statistical programmes (e.g. Andreou *et al.* forthcoming; Kitchin 2013; 2014; Cooper *et al*. 2021b: 29; Cooper and Green 2016; De Mauro *et al.* 2015).

The multidisciplinary study of big data is an emerging field, only gaining in importance in the last two decades or so — with a substantial increase in publications since 2008 — as a result of the so-called digital revolution (Rieder and Simon 2017: 86; Halevi and Moed 2012). Compared to other disciplines, archaeological big data, especially for the MENA region, is still growing up, and the EAMENA project — in collaboration with the MarEA project, which makes use of the same database (Andreou *et al.* 2020; forthcoming) — seeks to take a leading part in this exciting development. Its open-endedness, inclusiveness and diversity in data structure set it apart from other — more common and arguably more straightforward, but equally important — collations of large numbers of data points pertaining to specific subsets of archaeological data (archaeozoological, archaeobotanical, radiocarbon data, etc.) that are collated for smaller or larger regions, generally to answer specific research questions (e.g., Arbuckle *et al*. 2014; Bradbury *et al*. 2015; Flohr *et al*. 2016; Palmisano *et al*. 2021; also see Andreou *et al.* forthcoming). A leading part is not without challenges however, and perhaps the biggest challenge facing the EAMENA project team is how successfully to make sense of a dataset that is, to all intents and purposes except perhaps its current size, 'big', and to do justice to its versatility and multivocality.

Big data research has been critiqued for representing a 'new empiricism' that spelled the 'end of theory', purporting to be 'more *comprehensive*, more *objective*, and more *predictive*' (Kitchin 2014: 3; Rieder and Simon 2017: 86–87). Uncritical proponents of big data argued that it had the potential to remove human bias and the need for hypotheses and models, in the process revealing answers to questions that scholars did not even know to ask (summarized in Kitchin 2014: 4–5). The inherent naivety in such a view is immediately apparent. As archaeologists, we

[5] Some scholars have described archaeological data as 'slow' instead of 'big' data based on the speed with which archaeological data is typically generated in comparison to other types of data (e.g., Bickler 2021: 186).

are all too well aware that any dataset, however comprehensive, is never more than a sample providing a view of past human behaviour; that the process of documenting the past is never objective, and decisions need to be made on how to record and represent archaeological remains in digital format; and that no interpretation can exist entirely free from the subjective circumstances of the interpreter, which may range from subject-specific knowledge to the socio-political environment of the researcher (also see Kitchin 2014: 4–5; Leonelli 2014). As archaeologists specializing in remote sensing, we can add to this list an awareness of obvious biases in our recording: variable suitability of different landscape types for site identification, and variable ease with which different site types can be identified on aerial or satellite imagery. In the MENA region our methodology favours positive features like standing structures, that often date to relatively recent time periods, or *tell* sites. It also favours features that are especially well visible and preserved in areas of sparse modern occupation, such as in current deserts (e.g. desert kites). Finally, we can add a shift in perspective that supplants local, ground-based knowledge with remote, top-down knowledge requiring a degree of technical expertise, resulting in a Western-dominant perspective (e.g., Andreou *et al.* forthcoming; Davis *et al.* 2021; Fernandez-Diaz *et al.* 2018; Fisher *et al.* 2021; Sanger and Barnett 2021).

Rather than the 'end of theory' and human bias, we envisage the EAMENA project ultimately fitting the big data paradigm that employs 'data-driven science' using 'guided knowledge discovery techniques to identify potential questions (hypotheses) worthy of further examination and testing' and integrating data with relevant socio-historical theories in order to obtain a more holistic understanding of the complex, dynamic and interconnected processes that have shaped — and continue to shape — human society (Kitchin 2014: 6–7). The EAMENA project does not seek to answer a single question or tell a straightforward narrative: instead, it aims to give voice to the many narratives that can be extracted from the wealth of rich data that are gathered in the database. This Special Issue can be seen as an initial step on the journey towards that goal.

Heritage databases for research: the European background

Attempts to inventory and organize archaeological data go back to the 19th century, well before the term 'big data' was ever used, with the more widespread use of digital systems for this purpose from the 1990s onwards (McKeague *et al.* 2019; Willems 2000). During the 19th and early 20th centuries, such initiatives were concentrated in Europe. During this period, natural and cultural landscapes across the Levant, and the MENA region in general, were classified and described (e.g., Bannister and Marsh 1844; Conder and Kitchener 1880; Gsell 1911; Jacobsen and Hodson 1999; Palestine Exploration Fund 1888; Palmer 1871; Wilson *et al.* 1869), imposing a western perspective and narrative that was to have a lasting impact. Even more attention was directed inwards, however, on the rapidly growing body of archaeological (or 'antiquarian') information from the UK and western Europe itself (e.g., Brand 1849; 1905; Britton and Taylor 1807; Del Mar 1900; Hearne and Byrne 1807; Wright 1845). In England, this period saw the inclusion of administration and protection of historic monuments within the governmental Office of Works (which later became English Heritage and is now Historic England).

Less immediately obvious, the initiatives that were founded to inventory and classify British and other European data were also set to shape the future of archaeological documentation worldwide. Even today, the Getty Conservation Institute's Arches platform's core capabilities draw on the accumulated experiences of mainly western initiatives and 'best practices' to develop digital inventories for the cultural heritage sector on an international scale. It is worth noting however, that the Arches pre-cursor MEGA-Jordan was developed in close collaboration with Jordanian heritage professionals and Jordanian Department of Antiquities staff to include local needs and requirements, and with the aim of being used by local archaeologists and heritage professionals (Myers and Dalgity 2012). What is more, in the case of the EAMENA database, the adaptation and customization of the Arches platform took on board the feedback of partners from the MENA region, although ultimate decision-making power nevertheless rests with a UK-based team with strong professional backgrounds in British/European archaeology (Ten Harkel and Bewley 2018).[6] This section takes a closer look at the historical trajectories of the production of these datasets, to place the EAMENA database and its spin-off versions in their historical context and to promote explicit awareness of the historical baggage of the methodology. In addition, we believe that they contain useful lessons for the present and future of the EAMENA database.

The history of compiling lists of archaeological sites in the UK is long and complex, and its relevance here is best explained by starting with a brief (and somewhat simplified) summary (also see Fig. 2). It originated with a host of antiquarian research societies, focused on particular time periods and/or regions, that were founded from the 19th century onwards, in an era when 'archaeology' as a professional discipline did not yet exist (Gosden *et al*. 2021: 5–6). It underwent a process of professionalization in the late 19th and early 20th century through the work of the (regional) Victoria County History groups and the (national) Royal Commissions for Historic Monuments of England (RCHME), which was mainly concerned with survey and recording, and the aforementioned Office of Works, as well as equivalent government bodies in Scotland, Wales and Northern Ireland (Currie 2001; Illsley 2019; Sargent 2001). Archaeological sites were also recorded spatially by the Ordnance Survey (OS) on map sheets from the mid-19th century (in Britain, but also in the MENA region; e.g., Conder and Kitchener 1880; Wilson *et al.* 1869), and more formally since the establishment of the OS Archaeological Division in the 1920s (Illsley 2019).

From the late 1960s onwards, local government took an increasingly active role in recording, documenting and controlling archaeology and heritage in the UK. The UK's first Sites and Monuments Record (SMR), subsequently renamed Historic Environment Record (HER), covering the entire county of Oxfordshire, was started in Oxford at this time (Gosden *et al*. 2021: 6). There are currently over 85 such HERs in England alone (Cooper *et al.* 2021b: 32). In the UK, devolvement to local government initially had a negative impact on the standardization of terminology and the coherence of recording strategies, and significant work was put into the retrospective development of unified standards such as the Forum on Information Standards in Heritage (FISH) (founded in 1998). A legal development took place in the 1990s that made archaeological recording a compulsory part of the planning process. In the UK, this was Planning Policy Guidance 16 (PPG16), followed, in 1992, by the EU-wide Valetta Treaty (sometimes

[6] Customized versions of the Arches platform are now also used to record and store digital records of archaeological and heritage sites in the UK, including in Greater London, Lincoln, the Isle of Man and Jersey (*Implementations of Arches* n.d.; Lee Enriquez 2016).

referred to as the Malta Convention). This development led to a major explosion in the amount of archaeological data available, which was largely dealt with at local government (HER) level (Gosden *et al*. 2021: 6–7; Cooper *et al*. 2021b: 30; O'Keefe 1993).[7]

In addition to local government safeguarding initiatives, the protection of ancient monuments was also systematized on a *national* scale during the second half of the 20th century in the different member countries of the United Kingdom (Gosden *et al*. 2021: 4; Illsley 2019). Focusing in on the example of England, in the 1980s the national body known as English Heritage (now Historic England) was founded out of the Office of Works (by then, the Ministry of Works), and in 1999 they merged with the RCHME. Until recently, English Heritage/Historic England maintained two systematic national inventories: the National Heritage List for England (NHLE) and the National Record of the Historic Environment (NRHE) (Gosden *et al*. 2021: 4; Illsley 2019). The NHLE is a comprehensive list of all 'statutory and protected sites in the country' (Illsley 2019), focusing on legal and administrative aspects, and, as such, has commonalities with for example UNESCO's World Heritage List. The NRHE, on the other hand, is an extensive record of historic sites, landscapes and other features of interest, bringing together a range of earlier digital repositories that were compiled since the late 19th century, and currently containing more than 400,000 entries (Illsley 2019; Historic England 2005). It is similar in concept to the EAMENA database in terms of its comprehensive and dynamic approach, incorporating new results as and when they emerge. However, as governmental priorities and budgets have once more changed, the latest development is that, in acknowledgement of the advantages of a local approach to archaeology and heritage, a programme of interlinked projects is now underway under the banner of 'Heritage Information Access Simplified' (HIAS) that aims to incorporate all of the national-level NRHE data into the UK's 80+ regional HERs and UADs (Urban Archaeology Databases) by 2026. Historic England will continue to maintain only the NHLE on a national level (Historic England 2021).

Fig. 2: simplified diagram showing the historic trajectories of different initiatives to record and protect archaeological sites and monuments in England, and the dynamic relationships and (partial) overlaps between different data repositories.

It is against this background that the EAMENA initiative was conceived, as a sort of NRHE that would coexist with existing, more static lists of protected sites. What is more, from an early stage there was the idea that spin-off versions of the database could be implemented across the MENA region at 'national' level (see below), thus resulting in a similar interplay between heritage databases on different spatial scales as exists in England. Finally, a UK-hosted 'international' version of the EAMENA database will also continue to exist, which is likely to contain different data as a result of different data entry and recording strategies, as well as the research questions and priorities of its users. Several issues should be raised in this light.

The first issue concerns the problems that can arise from multiple datasets covering the same areas, which can be a hindrance to effective research. Such are the complexities of working with parallel archaeological datasets in the UK that their study has become a field of research in

[7] Interpretation of the legislature varies significantly from country to country within the EU (e.g., Duineveld *et al*. 2013; Kristiansen 2008; 2009; Willems 2007: 65; Willems 2009).

its own right. For example, from 2011 to 2016, the differences between archaeological datasets in England were a focus of the University of Oxford-based 'English Landscapes and Identities' (EngLaId) project, which collated over 900,000 heritage records from 92 different data repositories (local as well as national) into a single database (Cooper *et al.* 2021a; 2021b: 29). Taking an explicit big data approach and highlighting the processes that underlie the formation of these datasets — from field practises to decisions about data curation and storage — it emphasized the complex nature of archaeological datasets and highlighted the importance of data standards, among other things. Chronological and spatial information contained within the English datasets, as well as overlap and consistency between datasets, was highly varied due to the cumulative effect of ongoing processes of human interpretation, reinterpretation and sometimes error, taking place at various stages during the archaeological recording process and beyond (Cooper *et al.* 2021b: 29-54; Cooper and Green 2016).

Working beyond the national level in Britain and the rest of Europe is more complex still, at least in part because attempts to catalogue and record archaeology and heritage invariably originated on national or subnational levels. Retrospective initiatives like the EU's INSPIRE requiring certain shared data standards for public archaeology datasets, and portals like ARIADNE, the Pelagios Network, or ArkeoGIS where archaeological 'site' datasets can be shared beyond the national level, have the potential of bringing datasets together (McKeague *et al.* 2019), although these are yet incomplete or have very narrow focus in terms of datasets captured. The most successful stories seem to have been those where a model from one country is adopted and adapted by neighbouring countries. A European example is the UK Portable Antiquities Scheme (PAS) that was set up to record archaeological objects found by members of the public and especially hobbyist metal-detectorists. This inspired similar schemes to be developed in other European countries, including in the Netherlands (Portable Antiquities of the Netherlands (PAN)), Flanders (the Dutch-speaking region of Belgium, which has the MEDEA platform), Denmark (Digitale Metaldetektorfund (DIME)) and Finland (SuALT / Find Sampo), united under the European Public Finds Recording Network. However, substantial differences in database structure, terminology and underlying legal frameworks, among other things, persist (Dobat *et al.* 2020).

Non-comparability of information in different data repositories is a problem for research questions that address time periods when modern political boundaries did not yet exist. In Europe, this becomes increasingly problematic the further back in time one goes. In the MENA region, many modern political boundaries are even more recent and were often imposed by European powers during the 19th and earlier 20th centuries.

When considering the research potential of the EAMENA database in the light of the UK experience, it is clear that a practical and realistic future vision for the EAMENA initiative, involving all stakeholders in the UK, the MENA region, and elsewhere in the world, is key. The UK experience of multiple, parallel datasets resulting from multiple, pluriform initiatives at documenting the nation's archaeology and heritage — most notably the existence of local and national-level archives, as well as differing systems used in England and in the devolved administrations of Wales, Scotland and Northern Ireland — complicated the UK system significantly. The fact that heritage professionals from partner countries in the MENA region have received extensive training in the EAMENA methodology and contributed to its continued

development should go some way to maintaining a degree of standardization, even when different countries may adopt their own, independent EAMENA-based recording systems in the future.

An important challenge concerns the languages of the database. The English-language EAMENA database vocabulary and the underlying Arches software partly draw on the Forum for Information Standards in Heritage (FISH) and the Art and Architecture Thesaurus (AAT), and an important challenge concerns the development of comparable Arabic-language standards across the entire MENA region. What is more, other pre-existing documentation efforts on local, national and international scales exist, including other international collaborations which should also be included. Initially, the driving force behind this was the UK-based EAMENA team (e.g., Izzo *et al.* 2021), but we wish to see the balance shift towards MENA-based researchers. The paper by Mubaideen *et al.* (2021) is an important step in this direction.

It is nevertheless unavoidable that 'national implementation' projects of the EAMENA system will lead to the existence of multiple, parallel databases. It is here that the concept of linked open data has much to offer. Linked open data will be implemented at some stage during the future development of the EAMENA database, although the legal and political implications, as well as the ethical issues, remain to be investigated in full (but see Fisher *et al*. 2021). Continued close collaboration between the different stakeholders is a must, and one route to achieve this is through the development of inclusive, international research agendas that result from level collaboration between the different parties (e.g. Europeana for Europe; DigitalNZ for New Zealand; see Smithies *et al*. 2021a; 2021b for a project working towards integrating Jordanian cultural heritage data). That way, the vision of the EAMENA undertaking, which is to enable cross-border and international research using big data, can be realized.

As stated previously, the EAMENA database has been developed as both a heritage management and a research tool. This dual functionality is, potentially, its most important aspect for the realization of its overarching vision. Referring to the European context once more, in the UK and elsewhere, the primary users of heritage inventories were — and still are — local government archaeologists and commercial archaeology units, who utilize the data as part of the planning process (Cooper *et al.* 2012b: 31; Illsley 2019). In the MENA region, as a result of the focus of the CPF-funded training programmes, the main users of the EAMENA platform are currently also government archaeologists and heritage specialists. In (parts of) Europe, the advent of digitized inventories in the 1990s and 2000s did not have an immediate, clear impact as a data source for research, but merely provided a more direct access point to these datasets for existing users. The issue of condensing data in the process of creating digital lists was even, sometimes, seen as a problem from a research perspective: the primacy of heritage management concerns often led to skeletal and highly simplified datasets that lacked the fine-grained detail required by some academic research (see Mubaideen *et al.* 2021). Physical access continued to be required in order to use many elements of these inventories, including unpublished grey literature (e.g. unpublished fieldwork reports and desk-based assessments, generally produced by professional archaeologists), from which (for example) English HER inventories are ultimately condensed. In acknowledgement of this problem, various digital repositories of more extensive 'original' archaeological datasets, where digitized field records, databases and lengthy unpublished reports can be stored, have been created in the UK and the rest of Europe: examples include the Archaeological Data Service (ADS) in the UK and Data Archiving and Networked Services (DANS) in the Netherlands, adding even more digital repositories to the already-dazzling range

of resources. The EAMENA database is perhaps less likely to cause a similar surge in 'supporting' digital repositories, as it contains functionality to record information resources — image material as well as text-based resources — including (where available) links to open-access resources. This will minimize problems caused by the condensing of data, although the fact that every database is the result of a process of human interpretation should never be overlooked.

The hurdles caused by the almost-impenetrable nature of the European datasets have already been discussed briefly. As a result, academic research making use of these European heritage management databases is a recent phenomenon, highlighting a long-lasting divide between professional and academic archaeology. In the UK, it was not until the 2010s that a flurry of projects explored the potential of digital data for new archaeological research, some focusing on long-term chronologies whereas others tackled comparisons across different (modern) political boundaries. A project synthesizing the results of development-led fieldwork producing prehistoric results across several countries in north-west Europe was published in 2015 (Bradley *et al*. 2015). The Roman Rural Settlement project explored the wealth of new, digital data for the Late Iron Age and Roman periods in England and Wales; again, mainly generated during routine archaeological investigations taking place as part of the planning process (e.g., Allen *et al.* 2017; Smith *et al*. 2016; 2018). The Roman to Early Medieval transition in England and Wales was investigated by Rippon *et al*. (2015), and the aforementioned EngLaId project, in addition to analyzing the character and organization of the data itself, was also able to use this data for an ambitious and large-scale analysis of changes and continuities in the English landscape over a period of 3500 years, from *c.* 1500 BCE to *c.* 1100 CE (Cooper *et al.* 2021a). Digitization of the data was a key factor in enabling these major digital data studies.

It is hoped that access to the EAMENA datasets will stimulate the jumping-off point for comparable research in the MENA region, and — with the 'lessons learnt' from the European experience in mind — will do so on a much quicker timescale. For one, it should be particularly well-placed to bridge the gap between heritage management and research. As the pressures on archaeological resources continue to increase through global processes including population growth, agricultural expansion, large-scale infrastructure developments, mass-tourism and climate change, it is now more important than ever to unite practitioners from different sub-sections of the discipline in the same debates.

The key, in all of this, is the long-term vision for the project. It is important to note that in the UK, while the various national and heritage datasets began to appear online from *c.* 2000, the major research projects discussed above were initiated a decade later and we are still in the process of critically responding and further developing these efforts. Similarly, the potential of the EAMENA project is immense, but its realization will require long-term investment, both financial and in terms of archaeological and technological expertise. Although its attachment to the UK academic sector had initial advantages in this respect, for a more permanent future the project (or projects) must have a steady base, ideally at least in part within the MENA region. When that has been achieved, it can tackle the 21st-century challenges facing the archaeological heritage of our planet head on and move the field of archaeology and heritage protection in the MENA region forwards in unprecedented ways.

With all of this in mind, the next section will provide a brief summary of the project's history since its foundation in 2015. It will focus specifically on its developing methodologies, which changed as a result of the CPF-funded training courses mentioned previously. Finally, the paper

will showcase some initial research to illustrate the practical uses of the database to date. It is hoped that these will serve as a source of inspiration for many future projects to come.

Background to the EAMENA database: developing methodologies, flexible datasets

The overarching objective of the EAMENA project is to document all known (published/documented or unpublished/undocumented) archaeological and heritage sites predating c. 1950 CE. The methodology of the EAMENA project relies heavily on the use of remote sensing imagery (Bewley *et al*. 2016), and builds on the successful use of imagery datasets that have been available online via platforms such as Google Earth since the 2000s (e.g., Kennedy and Bishop 2011). To allow systematic survey, the MENA region was divided into grid squares, with each grid measuring 0.25 degrees longitude/latitude (*c.* 25 km x 25 km or 625 km^2). Following a standard methodology, researchers use Google Earth Pro (sometimes in combination with the satellite imagery available in the aerial view of Bing Maps) to rapidly survey each grid square and 'pin' possible archaeological and heritage sites. At a practitioner-defined scale of zoom, dependent on the equipment being used and the type of landscape being investigated, the grid is systematically analyzed and potential heritage resources are identified visually. The emphasis on human-led data generation and its inherent biases (see above) may seem out of place in the context of big data research, which tends to emphasize the need for computerized applications (e.g., De Mauro *et al.* 2015: 99); however, this is only one step in a process of information production that starts with the capture of satellite imagery. The potential of Machine Learning (ML) to identify sites is being investigated by the University of Leicester team (see also Orengo *et al.* 2020; Berganzo-Besca *et al.* 2021), but although this may speed up the process of data recording significantly, the high variability of site types, landscapes and periods included in the EAMENA project creates a continued need for human interpretation (see Bickler 2021 for a useful brief discussion of ML for archaeology; also see Andreou *et al.* forthcoming).

The reliance on remote sensing imagery has implications for the resulting data, some of which have been highlighted above. Although not all landscape types are suitable, and not all site types are recognizable, or even visible, on remote imagery, satellite survey is still better suited than ground survey to the identification of larger-scale archaeological or heritage landscapes, such as the extensive relict field systems in Lebanon discussed by Flohr *et al.* (2021), thus allowing for a landscape-based approach from the outset. What is more, as the methodology allows researchers to survey huge areas of land in relatively short amounts of time, it has exceptional potential for data *quantity* (and thus for quantitative research). In terms of data *quality*, it could be argued that remote survey has a negative effect, as it causes more uncertainty in terms of site interpretation and dating (e.g., Ten Harkel and Bewley 2018: 43; Cowley and Palmer 2010: 130–31). This is unavoidable, as members of the team often (although by no means exclusively) work across different parts of the MENA region, including areas in which they may never have worked on the ground, and to which they do not have in-field access. As discussed below, however, acknowledgement of the resulting interpretative uncertainties is a standard aspect of the data recording methodology. In addition, the collaborative networks with partners from the MENA region mean that people with a first-hand familiarity with the local landscape can enhance

observations made by the EAMENA project team, resulting in quality control and providing information to guide future data collection.

To redress the low visibility of certain site types and in line with the data-driven science paradigm that promotes exhaustiveness and inclusivity, information from other online repositories, as well as the results of published surveys and information about previously documented archaeological sites, are also added. For the Levant, the site repositories of the Digital Archaeological Atlas of the Holy Land (DAAHL) and the Department of Antiquities in Jordan's online MEGA-Jordan database (megajordan.org; Myers and Dalgity 2012) and its precursor JADIS (Palumbo 1994; 2012) are examples of key sources of information that can be drawn on to add information to heritage resources and verify whether sites have been previously documented. Not all these data have yet been added to the EAMENA database: where data are already available in repositories the team prefers to add selective good-quality information, rather than copy *en-masse* existing data that may be variable in quality. In these cases, satellite imagery is used mainly to correct site location — using the certainty fields in the database to indicate where this is not possible (see below and Ten Harkel and Fisher 2021) — and update knowledge of site condition. Additional information can furthermore be obtained from modern or historical imagery, including ground photographs (e.g. the Historic Environment Information Record (HEIR); Manar al-Athar (McKenzie *et al.* 2013), recent aerial photographs (e.g. the Aerial Photographic Archive for Archaeology in the Middle East (APAAME), also available through the EAMENA database), historic aerial photographs (Fradley 2021) and historic satellite imagery (declassified 1970s imagery from the CORONA and KH-7/KH-9 missions available from the CORONA Atlas and the USGS Earth Explorer websites, e.g. Ten Harkel and Shams 2021; Tews *et al.* in prep.), all of which can be added to the database as information resources. Finally, information from ground visits can also be recorded, although this tends to occur on a limited scale for reasons of staffing and budgets. One of the things that has already become apparent through the application of such a comprehensive and inclusive approach to data generation is just how many gaps there are in our understanding of the past, including the fact that certain regions receive the repeated attention of scholars, while others remain un(der)explored. Pushing forwards, the EAMENA project is excellently placed to redress this imbalance.

The resulting data are stored in the EAMENA database, which uses the Arches platform of the Getty Conservation Institute (for more detail see Ten Harkel and Fisher 2021). General resource information (e.g. site name), geographical information (e.g. site location), archaeological information (e.g. type of features, cultural period, etc.), and information on site condition (e.g. disturbances, threats) can be entered. In order to reduce subjectivity, controlled vocabulary is used in the English-language version of the database (glossaries are available at https://eamena.org/cpf-training) based on, among others, FISH and AAT, but modified and extended to suit the specific requirements of the EAMENA work (Zerbini 2018). (As mentioned above, the Arabic version, as well as possible future translations in Farsi and French, remains to be perfected in this respect.) Information about the person creating the record, as well as the methods and materials used, are recorded under the assessment summary and/or added as information resources. The creation of the EAMENA dataset — like all archaeological datasets — nevertheless constitutes an extensive process of archaeological (re-)interpretation, which users of the database should be explicitly aware of.

Two ways in which the database structure is designed to alert users to the subjectivity of the data is by the addition of certainty values to a range of database fields, and for aspects of the archaeological or condition assessments to be recorded as "unknown". This also allows for research and heritage management interventions to be carried out on different levels. For example, someone wishing to obtain an overall impression of past land use over a large area may choose to include all records in their research, whereas a researcher of the Roman period might decide to only select records with a "probable" or "definite" indication of this period, or a heritage professional who needs to prioritize a course of action within a limited time frame might focus their attention on sites with "planned" threats.

In 2017, two years into their documentation work, the EAMENA team obtained additional funding from the CPF to provide training courses in remote sensing and digital documentation to heritage professionals in the MENA region, primarily members of the Departments of Antiquities of the countries concerned. Between 2017 and 2020, a number of training workshops were organized and offered in Egypt, Tunisia (for Tunisian and Libyan colleagues), Lebanon, Jordan (for Jordanian and Palestinian colleagues), Iraq and Ethiopia (for Yemeni colleagues), as well as online. Multiple additional training courses were offered to other archaeologists and heritage professionals outside the CPF framework, for example during an Oxford workshop for Iranian researchers funded by the UK-based Iran Heritage Foundation (IHF), and in the context of the Royal Commission of AlUla project in Saudi Arabia. In total more than 200 archaeologists and heritage professionals, mainly from departments of antiquities but also from other government departments, universities and institutions, received training in the use of the EAMENA database and methodology. During this period, the trainees created over 16,000 records in the EAMENA database alone.[8]

Initially the focus of the workshops was on the potential of remote sensing and use of satellite imagery to identify unknown sites and provide information about their location, archaeological remains and condition. Another aspect of the training was digital recording and the use of the EAMENA database. This resulted in a process of knowledge-exchange — detailed local archaeological knowledge for technical know-how — that went some way to balance out the western-dominant interpretations of remote sensing data (Fisher *et al*. 2021; Andreou *et al*. forthcoming). Taking into account the trainees' feedback, the focus of the workshops during subsequent years was on monitoring and analysis of the archaeological (site) data. These training sessions covered topics such as use of GIS, Google Earth Engine, and conducting field monitoring, damage and risk assessments. Towards the end of the training programme, a number of the trainees started to use the EAMENA database for their own research, some of which was presented at EAMENA's Protecting the Past international conference series (and see Mubaideen *et al.* 2021 for a project set up by former EAMENA trainees following the EAMENA methodology). The use of the EAMENA database as a springboard for many different multi-vocal research projects capitalizing on the holistic and inclusive nature of the data collection represents a vital step towards realizing its full research potential.

The CPF-funded workshops cemented relations between the UK-based researchers and partners in the MENA region and — through the Protecting the Past conferences — beyond, but had an important impact on the EAMENA database as well, underlining the dynamism and

[8] Not counting records created in other Arches databases by EAMENA workshop participants, using their knowledge gained during the EAMENA training.

flexibility of the EAMENA datasets (Mubaideen *et al.* 2021; Ten Harkel and Fisher 2021). Additional functionality was added capturing trainees' feedback, allowing for more detailed information obtained during ground visits and more in-depth monitoring of site conditions, significantly expanding the potential and versatility of the database. Attention to previously under-studied site types that are difficult to identify using remote sensing techniques, such as urban heritage (Mubaideen *et al.* 2021), led to a more inclusive approach to archaeological recording. At the time of writing in January 2022, the database also contains functionality to reach increasingly high levels of detail in a nested model starting at site level and potentially zooming in to the level of an individual feature, or even part of a feature. By making optimal use of this functionality, the database can be used for heritage management purposes, to undertake meta-data analyses of data character and distribution, but — once data entry is in a more advanced stage — also to focus more directly on grand historical questions on a larger geographical scale.

At the start of the CPF-funded part of the project, each of EAMENA's partner institutions were at different stages with regard to (the development of) digital inventory and database systems. A few of the countries where training was provided already had centralized and functioning digital database systems, one of which — the MEGA-Jordan platform in Jordan, www.megajordan.org — already had a public-facing interface, whereas the majority were not online and directly accessible to researchers and the wider public. The training workshops have encouraged an interest in some countries to adapt and implement a version of the EAMENA database for their national heritage database, leading to the question — alluded to in the previous section — of how these different databases will relate to each other. This is an ongoing process and discussion and implementation work are at different stages for each country (*The Cultural Protection Fund* n.d.).

In sum, as the EAMENA project has developed over the course of several years, the EAMENA database has developed with it (Mubaideen *et al.* 2021; Ten Harkel and Fisher 2021; Sheldrick and Zerbini 2017; Zerbini 2018). As such it is the result of a complex interplay between different requirements, including academic (with an emphasis on large-scale research potential) and governmental (with an emphasis on heritage management and protection). In an ideal situation, the two go hand in hand, with research stimulating heritage management strategies and *vice versa* (Winterburn 2021). Seen from a big data perspective, which emphasizes the flexible and dynamic (or 'characterful') nature of data, the process of creating the database holds many valuable lessons for the way in which we think about archaeological information production. These remain to be explored in full, although the importance of new and untapped sources of information is already highlighted by Fradley (2021) and Izzo *et al*. (2021; also see Fisher *et al.* 2021; Andreou *et al.* forthcoming). As data entry continues, new kinds of cross-regional comparative analyses of large, archaeological-historical questions will become possible; however, even now the emerging datasets can shed light on existing research practices (Ten Harkel and Fisher 2021) and patterns of past land use (Flohr *et al.* 2021). The next section will discuss these issues in more detail through a brief introduction of the research papers that make up this Special Issue.

Research and the EAMENA database

The papers in this issue are based on work undertaken to explore or enhance the research capabilities of the EAMENA database as the system stood in 2018, only a few years into the project. Early versions of most of the papers were first presented at a session of the British Association of Near Eastern Archaeology (BANEA) conference at Durham University in that year. The original BANEA session contained papers on the wider MENA region, but this volume only includes those that focus on the Levant. All the papers included here have been developed and some have substantially changed (Ten Harkel and Fisher 2021), while one is a new contribution (Mubaideen *et al*. 2021). In different ways, all were instrumental in developing the ideas presented in this introduction.

A common thread through most of the papers is the question of how the EAMENA database can be used for research. This is the case for Ten Harkel and Fisher, who take a sample of Crusader-era sites from the Levant with uneven spatial coverage — as a case study with wider applicability — to explore the usefulness and potential of the database as a work-in-progress for research and heritage management purposes. They show how the database can be queried to identify new research questions as well as how its existing data can be meaningfully investigated. As one of the defining characteristics of big data research is that there is no such thing as a 'complete' dataset (which logically implies that a dataset is meaningful at every stage of its development and population), this study helps to think through some of the potential and limitations of the kind(s) of data in the EAMENA database more broadly. Looking at the database from a user perspective, they also delve more deeply into the development of the EAMENA database as a 'living' document of archaeological datasets, highlighting its dynamic and adaptable character. Commenting on the mutually beneficial relationship between database design and data entry, they highlight how the experiences of database users provided guidance in the process of database development to date, clearly showing the value of responsive database development to enable research. These are important issues as they demonstrate how the EAMENA initiative *as a process* is a useful research tool.

The paper by Mubaideen *et al.* follows closely on this. The authors are Jordanian heritage professionals with a background in architecture and with an active project on Amman's 20th-century 'heritage houses'. As such they ask how the database, which was designed primarily by and for archaeologists (here taken to mean specialists in the study of pre-1750 CE remains), can be used — and what should be adapted — for recording a type of heritage characterized by standing and in-use architecture. As the authors note, such recent heritage is often left out of inventories in the region, at least in Jordan, as a result of a division in legislation between 'archaeology' and 'heritage' (in Jordan pre- and post-1750 CE). They use their Amman Heritage Houses project as a case study, which has the additional advantage of presenting this interesting dataset in an academic journal for the first time. The research underlying this paper had a real practical impact on EAMENA's developing methodologies, as it makes specific recommendations on how to adapt an earlier version of the EAMENA database to enable the recording of post-1750 CE built heritage, so that data can be included that are essential to the research on this kind of heritage. It usefully includes a table of recommendations for the development of the database, and specifies which were (and which were not) taken on board during the creation of the current version of the EAMENA database. In some respects, most notably the take on cultural period precision, their opinions differ somewhat from those presented by Ten Harkel and Fisher, bearing testimony to the multi-vocal input underlying EAMENA's database development. Representing a

much-needed conversation between the professional Jordanian heritage sector and UK-based academic archaeology, this paper is also a representative example of the more extensive and wide-ranging process of feedback and responsive database development that resulted from the EAMENA training programmes.

The research presented by Flohr *et al*. specifically investigates the kind of contribution that a comprehensive and 'medium-scale' study of data contained in the EAMENA database can make to academic research. They focus on the landscape archaeology of Lebanon, using recent satellite imagery to compare different transects extending from the coast to the mountains in the north and south of the country. The regions concerned have been subject to different patterns of research, as a result of their different socio-political trajectories. The study is of interest for the theme of the Special Issue as the authors incorporate data that were collected using only the EAMENA methodology, i.e. based primarily on satellite imagery analysis, for some of the study regions without adding extra, published information, without special focus on a specific time period, theme or landscape type. They then show how this type of data can be used to address existing hypotheses (such as differences between the landscapes of the coast and the mountains, and the absence of visible burial monuments) and to identify new research avenues (e.g. drawing attention to the multiple visible remains of Ottoman–20th-century agricultural landscapes). At the same time, they discuss some of the limitations of the methodology which they — as researchers who are also familiar with the archaeology of Lebanon on the ground — are in an excellent position to comment on. An interesting detail in this paper is the incidental assessment of the power of data standards, as a difference in the recording of field systems in the north and south transects is highlighted, which is shown to constitute a real difference rather than researcher bias.

The value of the EAMENA database as a repository for information resources that can stimulate future survey and research is covered by Fradley's paper on the efforts of EAMENA to access historic aerial photograph collections. Specifically, his contribution discusses the importance and huge potential of British inter-war (1918–39) imagery of the Levant in more detail. Importantly this data is not monopolized by the UK-based EAMENA project team, instead being made available to all database users. This is an important factor in making the database more than a heritage management tool, as it includes 'original' data that underlies the interpretive process of data formation in the database. Through the investigation of the history of aerial photographic missions in the MENA region, and the archival history of the resulting photograph missions, the project has been able to access and contextualize datasets that have not previously been available to researchers. The historic aerial images lend chronological depth to both archaeological (i.e. research) and condition assessments (i.e. heritage management), so that it is, for example, possible to identify remains that have long since been destroyed, thus making an important contribution to any landscape archaeology study in the region.

The value of historic aerial imagery as a part of the EAMENA survey methodology is taken further in the paper by Izzo *et al.* In this contribution, the well-researched Palestinian landscape of the Jericho Oasis is re-evaluated via a collection of German aerial photographs taken in the last years of the First World War (WWI). Capturing a landscape that has been subject to extensive agricultural intensification and urban expansion over the past century, the imagery allows the study of a range of archaeological features that have not been documented in previous archaeological studies, while the added dimension of the EAMENA damage assessment provides an account of the key factors that have impacted the heritage of the oasis. As such it represents

another example of how an ever-growing, ever-expanding dataset pertaining to a relatively well-studied area can continue to bring new results to light. It also highlights the importance of bringing together resources and researchers from different countries on an international scale to obtain the most comprehensive knowledge base.

Finally, Winterburn's contribution, which focuses on the WWI-era conflict landscapes of southern Jordan, showcases the potential of the EAMENA database both as a research repository and for heritage management purposes. Following fieldwork, remote sensing and extensive archive and literature research, the EAMENA database was used as a repository for the information gathered, thus allowing effective monitoring of the sites' conditions. Specifically, the comprehensive approach of the EAMENA methodology has given this important piece of research a place: even though Jordan already has a highly functioning heritage sector with a public-facing digital database system, many of these sites are nevertheless under significant threat because heritage sites of such recent date do not enjoy the same levels of protection as older sites (also see the contribution by Mubaideen *et al.*). It is hoped that in future, the work on recent heritage facilitated by the EAMENA project will have a positive bearing on legal protection mechanisms as well.

Together these six papers exemplify different aspects of the research potential of the EAMENA database, summarized and further expanded upon in the concluding statement by Andrew Wilson. By recording as great a range of data types as possible, it immediately becomes possible to carry out chronologically and/or regionally specific analysis, in this case with chronological case studies ranging from the Crusader era to the mid 20th century across the Levant, or with regional case studies in more specifically defined locales such as the Jericho Oasis. It also becomes possible to take arbitrary transects across different landscape types to investigate differences in archaeology and in the disturbances and threats to sites based on current land use, terrain, or relative altitude. The benefit of the remote-sensing methodology is that it facilitates a landscape-level perspective, literally viewing 'sites' in the wider context of the ever-changing surrounding landscape. To those familiar with the methodology, this has underlined the rapid acceleration of landscape change especially in the most recent decades, a realization that has important implications for heritage management professionals (for example, how to find the right balance between heritage preservation and economic development) as well as academic researchers (preserving a wealth of data that can accommodate scores of research questions, especially for the more recent time periods and 'traditional' heritage of kinds that have traditionally been un(der)recorded in archaeological inventories). As data entry continues, the questions — and their answers — will likely become ever more ambitious and sophisticated.

To create a truly 'big' dataset, a big effort is required, and the research presented here is only scratching the surface. These open datasets will allow ever bigger questions as the scope of the database continues to grow, and innovative methodological approaches facilitated by machine learning and linked data will further broaden the potential for big data analyses spanning the entire MENA region and incorporating all archaeological time periods from deep prehistory to the 20th century. This may seem a distant dream at present, and a fuller test of whether this potential has been met will be required in years to come. However, no dataset can ever be 'complete', and it is therefore vital that explorations of data and themes continue and that they, like the studies presented here, will continue to impact on the development of the EAMENA project and its database. The EAMENA project has drawn extensively on the accumulated experiences of its

staff and users. As a result, both the longer-term historical trajectories of heritage databases in Europe and the specific requirements of heritage professionals from the MENA region have impacted on its development. Given time and future investment, the vision of the EAMENA project has the power to change the nature of archaeological research in the Levant, and in the MENA region more broadly.

Acknowledgements

The EAMENA project and its training programmes have been generously funded by the Arcadia Fund, a charitable fund of Lisbet Rausing & Peter Baldwin, and the UK's Cultural Protection Fund (CPF). Thanks to Bill Finlayson, Georgia Andreou, Andrew Wilson, and Graham Philip for helpful comments; to Victoria Donnelly, Chris Gosden and especially Roger Thomas for their thoughts on English archaeology; and to Pieterjan Deckers for information about the European Public Finds Recording Network.

References

AAT (Art and Architecture Thesaurus). [online thesaurus] Available at:<https://www.getty.edu/research/tools/vocabularies/aat/> [Accessed 1 October 2021].

ADS (Archaeology Data Service). [online data repository] Available at: <https://archaeologydataservice.ac.uk/> [Accessed 19 October 2021].

Arbuckle B.S., Kansa, S.W., Kansa, E., Orton, D., Çakırlar, C., Gourichon, L., Atici, L., Galik, A., Marciniak, A., Mulville, J., Buitenhuis, H., Carruthers, D., De Cupere, B., Demirergi, A., Frame, S., Helmer, D., Martin, L., Peters, J., Pöllath, N., Pawłowska, K., Russell, N., Twiss, K., Würtenberger, D. Data sharing reveals complexity in the westward spread of domestic animals across Neolithic Turkey. *PLoS One* 9(6):e99845. doi: 10.1371/journal.pone.0099845. Erratum in: PLoS One. 2014;9(9):e107824. PMID: 24927173; PMCID: PMC4057358.

Allen, M., Lodwick, L., Brindle, T., Fulford, M. and Smith, A. 2017. *The Rural Economy of Roman Britain: New Visions of the Countryside of Roman Britain Volume 2*. Britannia Monographs Series Vol 30. London: Society for the Promotion of Roman Studies.

Andreou, G., Blue, L., Breen, C., El Safadi, C., Huigens, H., Nikolaus, J., Ortiz Vazquez, R., and Westley, K. 2020. Maritime endangered archaeology of the Middle East and North Africa: the MarEA project. *Antiquity* [e-journal] 94 (378): E36. Available at: <https://doi.org/10.15184/aqy.2020.196> [Accessed 8 October 2021].

Andreou, G. M., Nikolaus, J., Westley, K., El Safadi, C., Blue, L., Smith, A. and Breen, C. forthcoming. Big data in maritime archaeology: challenges and prospects from the Middle East and North Africa. *Journal of Field Archaeology*.

APAAME (Aerial Photographic Archive of Archaeology in the Middle East). [online data repository] Available at: <www.humanities.uwa.edu.au/research/cah/aerial-archaeology> [Accessed 9 October 2021].

ARIADNE. [online datasets repository]. ARIADNEplus Project. Available at: <http://ariadne-infrastructure.eu/Portal/> [Accessed 19 November 2021].

ArkeoGIS. [online data repository]. Mapping, conserving, studying, managing: ArkeoGIS, sharing and using spatialized data about the past. Available at <https://arkeogis.org/en> [Accessed 19 November 2021].

Bannister, J. T. and Marsh, W. 1844. *A Survey of the Holy Land: Its Geography, History, and Destiny*. London: Binns and Goodwin.

Berganzo-Besga I., Orengo H. A., Lumbreras F., Carrero-Pazos M., Fonte J., Vilas-Estévez B. 2021. Hybrid MSRM-Based deep learning and multitemporal Sentinel 2-based machine learning algorithm detects near 10k archaeological tumuli in north-western Iberia. *Remote Sensing* 13(20): 4181. https://doi.org/10.3390/rs13204181.

Bewley, R., Wilson, A. I., Kennedy, D., Mattingly, D., Banks, R., Bishop, M., Bradbury, J., Cunliffe, E., Fradley, M., Jennings, R., Mason, R., Rayne, L., Sterry, M., Sheldrick, N. and Zerbini, A. 2016. Endangered Archaeology in the Middle East and North Africa: introducing the EAMENA project. In, Campana, S., Scopigno, R., Carpentiero, G., and Cirillo, M. (eds), *CAA2015. Keep the revolution going: Proceedings of the 43rd annual conference on computer applications and quantitative methods in archaeology*: 919–32. Oxford: Archaeopress.

Bickler, S. H. 2021. Machine learning arrives in archaeology. *Advances in Archaeological Practice* 9 (2):186–91.

Bradbury, J., Davies, D., Jay, M., Philip, G., Roberts, C., and Scarre, C. 2015. Making the dead visible: problems and solutions for 'big' picture approaches to the past, and dealing with large 'mortuary' datasets. *Journal of Archaeological Method and Theory* 23: 561–91.

Bradley, R., Haselgrove, C., Van der Linden, M. and Webley, L. 2015. *The Later Prehistory of North-West Europe: the evidence of developer-led fieldwork*. New York: Oxford University Press.

Brand, J. 1849. *Observations on the Popular Antiquities of Great Britain: Chiefly Illustrating the Origin of our Vulgar and Provincial Customs, Ceremonies, and Superstitions*. England: Bohn.

Brand, J. 1905. *Popular Antiquities of Great Britain. Faiths and Folklore; A Dictionary of National Beliefs, Superstitions and Popular Customs, Past and Current, with their Classical and Foreign Analogues described and Illustrated. Forming a New Ed. of "The popular antiquities of Great Britain" by Brand and Ellis, Largely Extended, Corrected, Brought Down to the Present Time, and Now First Alphabetically Arranged. By W. Carew Hazlitt*. England: Reeves and Turner.

Britton, J. and Taylor, J. 1807. *The Architectural Antiquities of Great Britain: Represented and Illustrated in a Series of Views, Elevations, Plans, Sections, and Details, of Various Ancient English Edifices: With Historical and Descriptive Accounts of Each*. London: Longman, Hurst, Rees, and Orme.

CAAL (Central Asia Archaeological Landscapes). [online] Available at: <https://uclcaal.org/> [Accessed 29 September 2021].

Conder, C. R. and Kitchener, H. H. K. 1880. *Map of western Palestine in 26 sheets: from surveys conducted for the committee of the Palestine exploration fund by Lieutenants C. R. Conder and H. H. Kitchener R. E. during the years 1872–1877.* Southampton: Ordnance Survey.

Cooper, A., Creswell, M., Donnelly, V., Franconi, T., Glyde, R., Gosden, C., Green, C., Kamash, Z., Mallet, S., Morley, L., Stansbie, D. and Ten Harkel, L. 2021a. *English Landscapes and Identities: Investigating Landscape Change from 1500 BC to AD 1086*. Oxford: Oxford University Press.

Cooper, A., Donnelly, V., Green, C. and Ten Harkel, L. 2021b. Characterful data: its character and capacities. In, Cooper, A., Creswell, M., Donnelly, V., Franconi, T., Glyde, R., Gosden, C., Green, C., Kamash, Z., Mallet, S., Morley, L., Stansbie, D. and Ten Harkel, L. *English Landscapes and Identities: Investigating Landscape Change from 1500 BC to AD 1086*: 30-54. Oxford: Oxford University Press.

Cooper, A. and Green, C. 2016. Embracing the complexities of archaeological 'big data': the case of the English Landscape and Identities project. *Journal of Archaeological Method and Theory* 23 (1): 271–304.

CORONA Atlas. n.d. [online] Available at: <https://corona.cast.uark.edu> [Accessed 5 September 2021].

Cowley, D. and Palmer, R. 2010. Interpreting aerial images – developing best practice. In, Campana, S., Forte, M., and Liuzza, C. (eds), *Space Time and Place: Proceedings of the III International Conference on Remote Sensing in Archaeology, 17–21 August 2009, Tirucirapalli, India.*: 129–35. Oxford: BAR.

Currie, C. R. J. 2001. The Victoria County History. *History Today* 49 (12): 28–30.

DAAHL (Digital Archaeological Atlas of the Holy Land). [online data repository]. Geo-Archaeological Information Applications Lab, webmaster S. Savage. Available at: <https://daahl.ucsd.edu/DAAHL/> [Accessed 5 September 2021].

DANS. [online data repository] Available at: <https://dans.knaw.nl/nl> [Accessed 29 September 2021].

Davis, D. S., Buffa, D., Rasolondrainy, T., Creswell, E., Anyanwu, C., Ibirogba, A., Randolph, C., Ouarghidi, A., Phelps, L. N., Lahiniriko, F. and Chrisostome, Z. M. 2021. The aerial panopticon and the ethics of archaeological remote sensing in sacred cultural spaces. *Archaeological Prospection* 28 (3): 305–20.

Del Mar, A.1900. *Ancient Britain in the Light of Modern Archæological Discoveries*. New York: The Cambridge Encyclopedia Company.

De Mauro, A., Greco, M. and Grimaldi, M. 2015. What is big data? A consensual definition and a review of key research topics. *AIP conference proceedings* 1644 (1): 97–104.

DigitalNZ. [online linked open data collections]. National Library of New Zealand. Available at <https://digitalnz.org/>. [Accessed 19 Novermber 2021].

DIME (Digitale Metaldetektorfund). [online data repository] Available at: <https://www.metaldetektorfund.dk> [Accessed 29 September 2021].

Dobat, A. S., Deckers, P., Heeren, S., Lewis, M., Thomas, S., Wessman, A., 2020. Towards a cooperative approach to hobby metal detecting: the European Public Finds Recording Network (EPFRN) Vision Statement. *European Journal of Archaeology* 23: 272–92.

Duineveld, M., Van Assche, K. and Beunen, R. 2013. Malta's unintentional consequences: archaeological heritage and the politics of exclusion in the Netherlands. *Public Archaeology* 12 (3): 139–54.

Europeana. [online linked open data collections]. Available at <https://www.europeana.eu/en/>. [Accessed 19 January 2021].

Fernandez-Diaz, J. C., Cohen, A. S., González, A. M. and Fisher, C. T. 2018. Shifting perspectives and ethical concerns in the era of remote sensing technologies. *SAA Archaeological Record* 18 (2): 8–15.

FISH (Forum on Information Standards in Heritage). [online] Available at: <http://www.heritage-standards.org.uk> [Accessed 19 October 2021].

Fisher, M. Fradley, M., Flohr, P., Rouhani, B. and Simi, F. 2021. Ethical considerations for remote sensing and open data in relation to the endangered archaeology in the Middle East and North Africa project. *Archaeological Prospection* [online] 28 (3). Available at: <10.1002/arp.1816> [Accessed 19 October 2021].

Flohr, P., Fleitmann, D., Matthews, R., Matthews, W., Black, S. 2016. Evidence of resilience to past climate change in Southwest Asia: Early farming communities and the 9.2 and 8.2 ka events. Quaternary Science Reviews 136: 23-39. http://dx.doi.org/10.1016/j.quascirev.2015.06.022

Flohr, P., Bradbury, J. and Ten Harkel, L. 2021. Tracing the patterns: fields, villages, and burial places in Lebanon. *Levant*.

Fradley, M. 2021. British Inter-war Aerial Photogrammetric Mapping in the MENA Region: Archives, Access and Research Potential. *Levant.*

Galiatsatos, N., Wilkinson, T. J., Donoghue, D. N. M. and Philip, G. 2009. The Fragile Crescent Project (FCP): analysis of settlement landscapes using satellite imagery. *CAA 2009: Making history interactive.* Williamsburg, Virginia, USA, 22–26 March 2009. Available at: <https://dro.dur.ac.uk/6909/1/6909.pdf?DDD6+DDD14+dgg1ng+d67a9y> [Accessed 19 October 2021].

Gosden, C., Franconi, T. and Ten Harkel, L. 2021. Introduction. In, Cooper, A., Creswell, M., Donnelly, V., Franconi, T., Glyde, R., Gosden, C., Green, C., Kamash, Z., Mallet, S., Morley, L., Stansbie, D. and Ten Harkel, L. *English Landscapes and Identities: Investigating Landscape Change from 1500 BC to AD 1086*: 1–26. Oxford: Oxford University Press.

Gsell, S. 1911. *Atlas archéologique de l'Algérie*. Alger et Paris: Adolphe Jourden et Fointemoing et Companie.

Halevi, G., and Moed, H. 2012. The evolution of big data as a research and scientific topic: overview of the literature. *Research Trends* 30: 3–6.

Hearne, T. and Byrne, W. 1807. *Antiquities of Great Britain, Illustrated in Views of Monasteries, Castles and Churches, Now Existing*. London: T. Cadell and W. Davies.

HEIR (Historic Environment Image Resource). [online data repository] Available at: <http://heir.arch.ox.ac.uk> [Accessed 19 October 2021].

Historic England 2005. *Historic England National Record of the Historic Environment (NRHE).* Available at: <https://archaeologydataservice.ac.uk/archives/view/398/> [Accessed 23 July 2021].

Historic England 2021. *The National Record of the Historic Environment (NRHE) to Historic Environment Record (HER) Transfer.* Available at: <https://historicengland.org.uk/research/support-and-collaboration/heritage-information-access-simplified/data-transfer-to-historic-environment-records/> [Accessed 23 July 2021].

Illsley, W. R. 2019. Problematising the Historic Environment Record: comments on persistent issues in England and Sweden. *Conservation and Management of Archaeological Sites* 21 (2) 113–34.

Implementations of Arches n.d. [online] Available at: <https://www.archesproject.org/implementations-of-arches> [Accessed 27 Sept 2021].

Izzo, P., Fradley, M. and Zerbini, A. 2021. Snapshots from the past: discoveries and destruction in the Jericho Oasis. *Levant*.

Jacobson, D. M. and Hodson, Y. 1999. *Survey of Western Palestine, Including a Survey of Eastern Palestine 1881: Essays to Accompany the Archive Editions Facsimile*. Slough: Archive Editions in association with the Palestine Exploration Fund.

Kennedy, D. and Bishop, M. 2011. Google Earth and the archaeology of Saudi Arabia: a case study from the Jessah area. *Journal of Archaeological Science* 38 (6): 1284–93.

Kitchin R. 2013. Big data and human geography: opportunities, challenges and risks. *Dialogues in Human Geography* 3 (3): 262–67.

Kitchin R. 2014. Big Data, new epistemologies and paradigm shifts. *Big Data & Society* [online]. Available at: <doi:10.1177/2053951714528481> [Accessed 5 August 2021].

Kristiansen, K. 2008. Do we need the 'archaeology of Europe'? *Archaeological Dialogues* 15 (1): 5–25.

Kristiansen, K. 2009. Contract archaeology in Europe: an experiment in diversity. *World Archaeology* 41 (4): 641–48.

Lee Enriquez, A. 2016. *Arches to be implemented in Lincoln and Greater London.* [online] Available at: <https://www.archesproject.org/arches-to-be-implemented-in-lincoln-and-greater-london> [Accessed 27 September 2021].

Leonelli, S. 2014. What difference does quantity make? On the epistemology of Big Data in biology. *Big Data and Society* [online]. Available at: <https://doi.org/10.1177/2053951714534395> [Accessed 5 August 2021].

MAEASaM (Mapping Africa's Endangered Archaeological Sites and Monuments). [online] Available at: <http://maeasam.org/> [Accessed 29 September 2021].

MAHS (Maritime Asia Heritage Survey). [online] Available at: <https://maritimeasiaheritage.cseas.kyoto-u.ac.jp/> [Accessed 29 September 2021].

MAHSA (Mapping Archaeological Heritage in South Asia) [online] Available at: <https://www.arch.cam.ac.uk/research/projects/current-projects/mapping-archaeological-heritage-south-asia> [Accessed 29 September 2021].

Manar al-Athar. [online] Available at: <http://www.manar-al-athar.ox.ac.uk/> [Accessed 29 September 2021].

MarEA Maritime Endangered Archaeology in the Middle East and North Africa) [online] Available at: <https://marea.soton.ac.uk/> [Accessed 29 September 2021].

Mattingly, D. J. and Sterry, M. 2013. The first towns in the central Sahara. *Antiquity* 87 (336): 503–18.

Mattingly, D. J., Sterry, M. and Leitch, V. 2013. Fortified farms and defended villages of Late Roman and Late Antique Africa. *Antiquité Tardive* 21: 99–119.

McKeague, P., van 't Veer, R., Huvila, I., Moreau, A., Verhagen, P., Bernard, L., Cooper, A., Green, C. and van Manen, N., 2019. Mapping our heritage: towards a sustainable future for digital spatial information and technologies in European archaeological heritage management. *Journal of Computer Applications in Archaeology* 2 (1):89–104.

McKenzie, J. *et al*. 2013–. Manar al-Athar Photo-Archive. Available at: <http://www.manar-al-athar.ox.ac.uk> [Accessed 15 November 2021].

Mubaideen, S., Salameen, D. and Al Momani, R. 2021. Using the EAMENA database to document modern heritage: the Amman Heritage Houses, Jordan, case study. *Levant.*

Myers, D., and Dalgity, A. 2012. The Middle Eastern Geodatabase for Antiquities (MEGA): an open source GIS-based heritage site inventory and management system. *Change Over Time* 2 (1): 32–57.

O'Keefe, P. J. 1993. The European Convention on the Protection of the Archaeological Heritage. *Antiquity* 67: 406–13.

Palestine Exploration Fund 1888. *The Survey of western Palestine*. England: Palestine Exploration Fund.

Palmer, E. H. 1871. *The Desert of the Exodus: Journeys on Foot in the Wilderness of the Forty Years' Wanderings, Undertaken in Connexion with the Ordnance Survey of Sinai and the Palestine Exploration Fund*. Cambridge: Deighton, Bell and Co.

Palmisano, A., Lawrence, D., de Gruchy, M.W., Bevan, A., Shennan, S. 2021. Holocene regional population dynamics and climatic trends in the Near East: A first comparison using archaeo-demographic proxies. Quaternary Science Reviews 252: 106739. https://doi.org/10.1016/j.quascirev.2020.106739.

Orengo, H. A., Conesa, F. C., Garcia-Molsosa, A., Lobo, A., Green, A.S., Madella, M., Petrie, C.A. 2020. Automated detection of archaeological mounds using machine-learning classification of multisensory and multitemporal satellite data. *Proc. Natl. Acad. Sci. USA* 117: 18240–18250.

Palumbo, G. 1994. *JADIS: the Jordan Antiquities Database and Information System: a summary of the data.* Amman: Department of the Antiquities of Jordan.

Palumbo, G. 2012. Twenty-five years of archaeological site inventories in the Middle East: challenges and perspectives. *Change Over Time* 2 (1): 20–31.

PAN (Portable Antiquities Netherlands). [online data repository] Available at: <https://portable-antiquities.nl/pan/#/public/about> [Accessed 5 July 2021].

PAS (Portable Antiquities Scheme). [online data repository] Available at: <https://finds.org.uk/> [Accessed 5 July 2021].

The Pelagios Network. Available at: <https://pelagios.org/> [Accessed 19 November 2021].

Rieder, G. and Simon, J., 2017. Big data: a new empiricism and its epistemic and socio-political consequences. In, Pietsch, W.; Wernecke, J.; and Ott, M. (eds), *Berechenbarkeit der Welt? Philosophie und Wissenschaft im Zeitalter von Big Data*: 85–105. Wiesbaden: Springer.

Rippon, S., Smart, C. and Pears, B. 2015. *The Fields of Britannia: Continuity and Change in the Late Roman and Early Medieval Landscape.* Oxford: Oxford University Press.

Sanger, M. C. and Barnett, K., 2021. Remote sensing and indigenous communities: challenges and opportunities. *Advances in Archaeological Practice* 9 (3): 194–201.

Sargent, A. 2001. 'RCHME' 1908-1998 – A history of the Royal Commission on the Historical Monuments of England. *Transactions of the Ancient Monuments Society* 45: 57–80.

Sheldrick, N. and Zerbini, A. 2017. A heritage inventory for documenting endangered archaeology in the Middle East and North Africa. *ISPRS Annals of the Photogrammetry, Remote Sensing, and Spatial Information Sciences* IV-2/W2: 237–41.

Smith, A., Allen, M., Brindle, T. and Fulford, M. 2016. *The Rural Settlement of Roman Britain: New Visions of the Countryside of Roman Britain Vol. 1.* Britannia Monographs Series Vol 29. London: Society for the Promotion of Roman Studies.

Smith, A., Allen, M., Brindle, T., Fulford, M., Lodwick, L., and Rohnbogner, A. 2018. *Life and Death in the Countryside of Roman Britain: New Visions of the Countryside of Roman Britain Vol 3.* Britannia Monographs Series Vol 31. London: Society for the Promotion of Roman Studies.

Smithies, J., Bala'awi, F., Flohr, P., Mubaideen, S., Esposito, A., Idwan, S., Palmer, C., Mahasneh, I., Rababeh, S. 2021a. MaDiH (مديح): Mapping Digital Cultural Heritage in Jordan Policy White Paper (English Version) (1.0). Zenodo. https://doi.org/10.5281/zenodo.4866975

Smithies, J., Bala'awi, F., Flohr, P., Mubaideen, S., Esposito, A., Idwan, S., Palmer, C., Mahasneh, I., Rababeh, S. 2021b. MaDiH (مديح): Mapping Digital Cultural Heritage in Jordan Technical White Paper (English Version) (1.0). Zenodo. https://doi.org/10.5281/zenodo.4867531.

SuALT / Find Sampo. [online] Available at: <https://blogs.helsinki.fi/sualt-project/about-the-project/> [Accessed 4 July 2021].

Ten Harkel, L. and Bewley, R. 2018. An incredible journey? Understanding ancient landscapes from England to the Middle East and North Africa. *Medieval Settlement Research* 33: 40–59.

Ten Harkel, L. and Fisher, M. T. 2021. The EAMENA database and its potential impact on research and heritage management: a case study of Crusader heritage in Lebanon. *Levant*.

Ten Harkel, L. and Shams, A., forthcoming. Declassified satellite imagery (KH-9): analysing the impact of mass tourism on the St Catherine region, South Sinai. EAMENA project blog [tbc].

Tews, S., Ten Harkel, L. and Shams, A. in prep. From the coast to the High Mountains: a remote sensing survey of disturbances and threats to the archaeology and cultural heritage of South Sinai. *Journal for Maritime Archaeology*.

The Cultural Protection Fund n.d. Available at <https://eamena.org/cultural-protection-fund> [Accessed 22 October 2021].

USGS Earth Explorer. n.d. [online] Available at: <https://earthexplorer.usgs.gov/> [Accessed 20 September 2021].

Wilkinson, T. J., Philip, G., Bradbury, J., Dunford, R., Donoghue, D., Galiatsatos, N., Lawrence, D., Ricci, A. and Smith, S. L. 2014. Contextualizing early urbanization: Settlement cores, early states and agro-pastoral strategies in the Fertile Crescent during the fourth and third millennia BC. *Journal of World Prehistory* 27(1): 43-109.

Willems, W. 2000. The management of the archaeological heritage in the Netherlands. *Archaeologia Polona* 38: 153-168.

Willems, W. 2007. The work of making Malta: the Council of Europe's Archaeology and Planning Committee 1988–1996. *European Journal of Archaeology* 10 (1): 57–71.

Willems, W. 2009. Archaeological resource management and academic archaeology in Europe: some observations. In, D'Agata, A. L. and Alaura, S. (eds), Quale futuro per l'archeologia?: 89–99. Rome: Gangemi Editore.

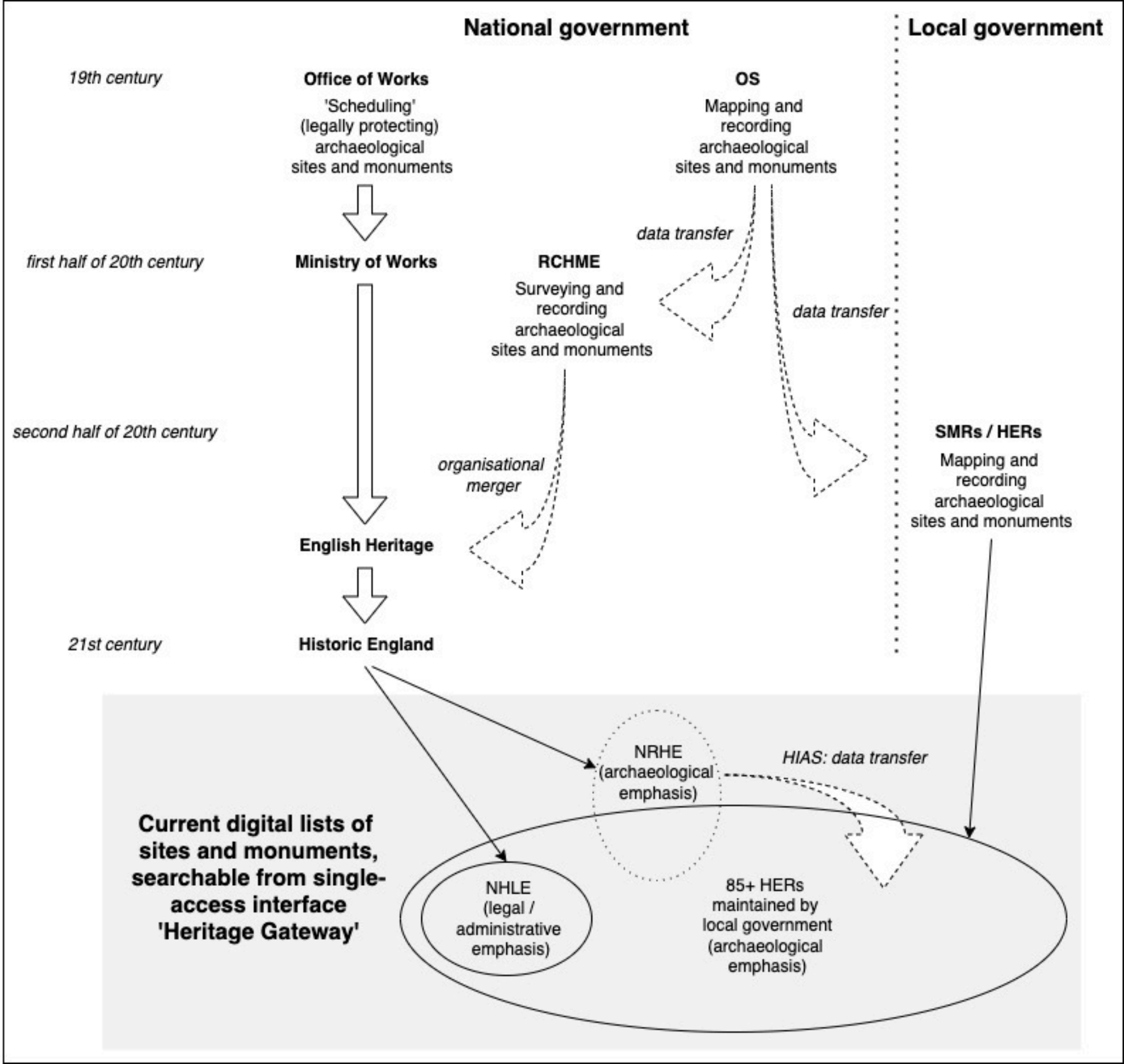
Wilson, C. W., Palmer, H. S. and James, H. 1869. *Ordnance Survey of the Peninsula of Sinai*, Southampton and London: Ordnance Survey Office and H.M. Stationery Office.

Winterburn, J. 2021. *Levant. The late Ottoman defended landscapes along the Hijaz railway in southern Jordan.*

Wright, T. 1845. *The Archaeological Album, or, Museum of National Antiquities.* England: Chapman and Hall.

Zerbini, A. 2018. Developing a heritage database for the Middle East and North Africa. *Journal of Field Archaeology* 43: Supplement 1, 9–18.


1000 km


National government
Local government
19th century
Office of Works
'Scheduling' (legally protecting) archaeological sites and monuments
OS
Mapping and recording archaeological sites and monuments
data transfer
first half of 20th century
Ministry of Works
RCHME
Surveying and recording archaeological sites and monuments
data transfer
second half of 20th century
organisational merger
SMRs / HERs
Mapping and recording archaeological sites and monuments
English Heritage
21st century
Historic England
NRHE (archaeological emphasis)
HIAS: data transfer
Current digital lists of sites and monuments, searchable from single-access interface 'Heritage Gateway'
NHLE (legal / administrative emphasis)
85+ HERs maintained by local government (archaeological emphasis)