The whole is other than the sum of its parts: the untapped potential of spatial data from archaeological fieldwork

[version 1; peer review: 2 approved with reservations]

Peter McKeague

Heritage Directorate, Historic Environment Scotland, Edinburgh, Scotland, EH8 9NX, UK

Abstract

Spatial data is fundamental to documenting our past, underpinning research questions and informing the decisions we – and others – take in the protection, understanding and stewardship of the historic environment. Investing in organising and sharing spatial data typically delivers a benefit to cost ratio of up to sixteen times the outlay. In contrast, existing practices are saturated with inefficiency. Even if the spatial data reaches an archive, it can be hard to find, access and use. It is often in the wrong format and incompatible with similar data from different projects.

Through the European Union Infrastructure for Spatial Information in Europe (INSPIRE) Directive curators of protected sites data (scheduled monuments, listed buildings etc.) are required to publish metadata, view and download services. In contrast to these ‘curated’ datasets there is neither the mandate, nor the mechanisms to coordinate, curate and share, data created through archaeological fieldwork and research. The degree of standardisation in documenting fieldwork recording and archival deposition varies considerably, posing challenges to the reuse of data. Both are key factors in not realising the potential of ‘collated data’ from multiple sources. The opportunities to develop a consistent approach for making greater use of data routinely created through fieldwork and research are explored using case studies from Scotland.

Despite the obvious benefits of developing a consistent approach to spatial data from fieldwork, the framework, standards, specifications, guidance and infrastructure to realise that potential are absent. Archaeological data can and should contribute to delivering wider societal benefits, including environmental monitoring, digital twins and climate change. To contribute meaningfully to these and other societal challenges, archaeological data needs to be accessible and consistent.
Keywords
Archaeology, Data Standards, Environment, Geospatial data, Heritage, Societal benefit

This article is included in the New Digital Archaeologies collection.

Corresponding author: Peter McKeague (peter.mckeague@hes.scot)

Author roles: McKeague P: Conceptualization, Investigation, Project Administration, Resources, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: I am very grateful to Historic Environment Scotland for covering the costs of publishing this article. The paper draws on research undertaken through an Arts and Humanities Research workshop funding award (2019) from the Royal Society of Edinburgh.

Copyright: © 2023 McKeague P. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: McKeague P. The whole is other than the sum of its parts: the untapped potential of spatial data from archaeological fieldwork [version 1; peer review: 2 approved with reservations] F1000Research 2023, 12:112
https://doi.org/10.12688/f1000research.126361.1

First published: 30 Jan 2023, 12:112 https://doi.org/10.12688/f1000research.126361.1
Introduction: reasons for and benefits of harmonising data at scale

The benefits of collating and sharing spatial data systematically from multiple providers, to address a range of environmental issues or to underpin research, have long been recognised. Developed in response to ever increasing environmental challenges such as flooding and pollution on a trans-national scale, the European Union Infrastructure for Spatial Information in Europe (INSPIRE) Directive (2007) provides the mandate and framework for public organisations to share environmentally related datasets within their own and neighbouring countries to improve public access, inform decision making and improve the stewardship of the environment. Initiatives like One Geology, a partnership of 190 organisations from 121 countries, have developed their own network developing the framework, infrastructure and standards for their discipline. Collating data from multiple sources and many countries is also at the heart of the multi-disciplinary The Role of Culture in Early Expansion of Humans (ROCEEH) project.¹

For marine data, The Marine Environment Data Information Network (MEDIN) provides the coordination and acts as a data hub for a broad range of offshore datasets managed by accredited data archive centres (DACs). The seven DACs cover Bathymetry, Water Column Oceanography, Geology and Geophysics Flora, Fauna and Habitat, Meteorology, Fisheries and the Historic Environment. Each DAC contributes key metadata about project datasets to the MEDIN portal. The portal enables users to search across over 17,225 marine datasets from over 600 UK organisations.

Benefits of harmonising and sharing spatial data from multiple providers

A number of studies demonstrate the value of harmonising data across multiple providers. A cost benefit analysis² undertaken by environmental economists and marine consultants on behalf of MEDIN demonstrated a benefit to cost ratio (BCR) of 8.2. A review of a range of marine spatial data infrastructures (SDIs) in New Zealand found the BCR to range from 2 to 18.² Benefits include efficiency of data collection and reuse, supporting marine science and more effective marine spatial planning.

There is already a lot of good practice in the stewardship of historic environment data. National Heritage Agencies and Local Authority Historic Environment Records manage inventories for their respective jurisdictions. A recent study³ of the impact of local authority archaeologists in Britain, informed by information in Historic Environment Records (HERs), estimates that archaeology in development management had a BCR of 15, contributing £218m to local economies from commercial archaeology and saving up to £1.3 billion in delay and emergency excavation costs. Project archives resulting from fieldwork and research may be deposited in a relevant archive, including Core Trust Seal accredited archives at Historic Environment Scotland (HES) and the Archaeology Data Service. An analysis⁴ of the value and impact of the Archaeology Data Service in 2013 estimated a two- to eight-fold return through additional reuse of data deposited in their digital archive, though not everything reaches an archive, or is deposited in a reusable format.

Despite the seemingly impressive benefits for the stewardship of data, are we really realising the value and potential of historic environment data, particularly that captured from primary archaeological fieldwork and research? The following sections explore the different approaches to publishing spatial data from organisations responsible for creating and maintaining spatial data - curated data – and the range of data created through primary field work by many organisations for a broad range of purposes, data that needs to be collated to realise real benefit.

Spatial Data and the historic environment - curated data

Good data is at the heart of decision making and stewardship of the historic environment. It should also be an important contributor to wider societal challenges including climate change, automation and complex modelling of our environment through eco-system services and digital twins. Good data should deliver efficiencies to a broad range of users outside of, but interacting with, heritage.

Through The INSPIRE (Scotland) Regulations 2009 key protected sites datasets, including scheduled monuments and listed buildings, are readily accessible through national and organisational metadata portals. These are curated datasets where organisations are responsible for the creation and maintenance of content to deliver their business objectives. Publication of these datasets includes Discovery Metadata, View and Download services.

The Historic Environment Scotland portal provides access to the statutory data HES is responsible for, as well as access to point data from the National Record of the Historic Environment, published through Canmore. For most datasets users may download data as zipped shape files or access Web Map and Web Features services or an Atom feed. Historic land use data may also be accessed through a more restrictive licence. Two web map browsers, The Designations Map Search and PastMap provide location-based services. PastMap also allows users to view and search Canmore and most HER data providers, to query, select and download point data from across searchable datasets in csv or kml format, and to view content on linked web pages hosted by the data contributor. Canmore offers both text and map based searches.
HES key datasets are also added to the SpatialData.gov.scot metadata portal, which enables users to find, share and reuse spatial data provided by Scottish public sector organisations. These records are harvested by and added to the Data.gov.uk open data portal.

Most HERs maintain or have access to databases recording information about monuments and events (specific activities that lead to the recognition of or add to the understanding of a monument). As well as contributing to PastMap, monument and event data is provided to The Spatial Hub. Established by The Improvement Service (the national improvement organisation for Local Government in Scotland), The Spatial Hub provides the coordination and infrastructure for harmonising and transforming 138 datasets across Scotland’s 32 local authorities and other partner organisations, for use by the Scottish public sector bodies and their contractors under the Public Sector Mapping Agreement. Academic researchers may access the data through EDINA.

Spatial data and the historic environment - collated data

Primary archaeological fieldwork, research and analysis, often referred to as ‘events’, are unique activities that happen once over a discrete period of time – often through research, stewardship of the historic environment or in discharge of a condition required under planning consent. Fieldwork routinely creates a wealth of spatial data whose reuse potential is largely unrealised beyond short-term project delivery. Results are typically presented in a pdf report or more traditional publication such as the Royal Commission on the Ancient and Historical Monuments of Scotland (RCAHMS) survey of the Mar Lodge Estate, Aberdeenshire (Figure 1),6 where the spatial data is fossilised to the page orphaned from its broader landscape context.

The underlying data might be deposited with a formal archive such as the Historic Environment Scotland Archives or The Archaeology Data Service (ADS): both Core Trust Seal certified repositories. A study by the Association of...
Local Government Archaeology Officers (ALGAO) in England, working with the ADS, estimated that at best, 2-3% of commercial projects had been digitally archived with the ADS. The proportion of archives from fieldwork in Scotland deposited with HES remains unquantified.

Even within the archive, the data may not be easily found and retrieved. Archival discipline necessarily compartmentalises data by collection and project (Figure 2) greatly understating the value of the spatial component.

Notable exceptions include the HES National Collection of Aerial Photography and the National Library of Scotland Map Library who both publish online map browsers as finding aids to their respective collections. All too often it falls to individual initiatives or individual enthusiasm and project-based solutions to realise the potential of spatial data preserved in archives. Creating these collated datasets requires considerably more investment to collect, preserve, harmonise and publish than curated data. Critically, the mandate provided for curated data to provide that coordination is lacking.

Record enhancement is a key function of both the HERs and HES. Existing practice is for the curator to enhance the record through extracting key data from project reports produced as part of the discharge of a planning condition, or from summary (often interim) reports published in *Discovery and Excavation in Scotland*, an annual digest of fieldwork. Since 2007, fieldwork can be reported through an online application: OASIS (Online AccesS to the Index of archaeological investigations). The OASIS form ensures that fieldwork reporting is standardised, meeting the MIDAS Heritage Standard by drawing on controlled vocabularies published on Heritagedata.org to ensure a degree of consistency at project level. Additionally, the project reports should be uploaded and shared with the relevant HER and HES before release to the Archaeology Data Service. The latest iteration, OASIS V, encourages users to upload project extents, albeit with restricted geometries, to accurately locate projects and share with curators under an open CC-BY licence.

Nevertheless, project extents are still often digitised from raster images designed for inclusion in a data structure report or publication – a process ingrained with inefficiency and inaccuracy. OASIS addresses standardisation and harmonisation at project level but major barriers remain in routinely accessing and sharing more detailed spatial data including interoperability across datasets from multiple providers.

The Scottish HERs published specifications for documenting the basic ‘event’. These specifications provide a baseline, or vanilla, record for an event (the who, what, why, and when) enough to meet HER requirements in documenting where
activities have taken place. The specifications do not and are not intended to address the increasing sophistication of primary datasets from a range of non-invasive and invasive techniques to systematically document content within data types or act as finding aids or spatial indexes to associated archives. For instance, as well as seeing the extent of a geophysical survey, a user should be able to assess the usefulness of the dataset through viewing the technical metadata (instrumentation, resolution of survey etc) – all captured through the geophysical module in OASIS with a link to the related archive, where known.

With data created by a broad range of professional organisations, in academia and through community led projects for a wide-range of purposes, the absence of relevant data standards, specifications and reference vocabularies presents a challenging environment for data harmonisation or homogenisation. Even if data is internally consistent from a single provider – and that is not always the case - interoperability across data providers remains challenging.

Case studies
Scotland’s past is a rich and diverse resource ranging from the lowlands to the highlands, from rural to urban and from subtle archaeology revealed as cropmarks and increasingly recognised through a range of remote sensing analysis to upstanding archaeological sites and landscapes. The following examples explore the potential for sharing spatial data from separate projects and a range of techniques – to build something other than the sum of the parts.

Field survey
Excepting a small number of university research projects, for many decades systematic archaeological field survey was the almost exclusive preserve of two national, public-sector organisations. RCAHMS compiled inventories of monuments and Britain’s National mapping agency, The Ordnance Survey, surveyed and published depictions of archaeological monuments as topographic on their paper maps. Both organisations established robust mapping conventions to interpret and present individual site depictions consistently. With the need for environmental impact assessments and planning requirements, walk over surveys are now an important part of many infrastructure projects. Spatial data standards have not, however, kept pace with the growth of a thriving private sector.

The RCAHMS Afforestable Land Survey was established in 1989 to undertake strategic targeted surveys in areas where the archaeological potential had not been recognised through low levels of archaeological recording. The survey

![Figure 3. Extract of archaeological survey data from Mar Lodge viewed within a Geographic Information System (GIS). Background mapping © Crown Copyright HES (2023). Ordnance Survey licence number 100057073.](image)
programme evolved to use a range of digital equipment from electronic distance measurement equipment (EDMs) to differential global positioning systems (dGPS) to map the archaeological landscape. With the advent of Geographic Information Systems (GIS), data initially gathered to produce maps of the archaeological landscape for archive and publication (Figure 3) could be combined across survey projects to present a seamless view of mapped survey areas.

The approach is extensible. It could and should form the basis of specifications to realise the potential of mapped data fossilised in project reports from the ever-increasing number of commercial, community and research projects.

**Airborne mapping and remote sensing analysis**

Transcribing or mapping archaeological features from oblique and vertical aerial photography on a site-by-site basis is a long-established technique. What was a slow and laborious manual process has been transformed through computerised rectification programmes. In Scotland, an Airborne Mapping programme has been undertaken by RCAHMS/HES for over thirty years. Again, the potential to combine individual site transcriptions within the GIS to present a seamless view contextualising individual site transcriptions within a wider archaeological landscape (Figure 4) was quickly realised through development of consistent data specifications similar to, but not identical to, that for field survey mapping.

Geophysical survey techniques reveal anomalies which can be interpreted as archaeological features and sites. The techniques of airborne laser scanning or LiDAR (light detection and ranging) analysis offer fresh opportunities for identifying and mapping archaeology, beyond the rich arable lowlands, into the upland and penetrating the tree canopy. Techniques and methodologies may differ, but the outputs from the interpretation need to be consistent across approaches, with particular attention to the terminology used. Local vocabularies need to reference formal, accessible, shared, and broadly applicable language for knowledge representation to improve interoperability (The ‘I’ of the FAIR Data Principles).

**Taking the holistic approach**

Consistent application of spatial data specifications is an essential prerequisite to develop a holistic understanding of complex archaeological landscapes from a range of non-invasive and invasive techniques and researchers. The rich
archaeological landscape at Inveresk, East Lothian has been mapped through the Airborne Mapping programme and subject to numerous small-scale excavations (Figure 4). Airborne mapping recorded the complex multi-period cropmark landscape, including features identified as Roman temporary camps and a Roman field system overlying a Neolithic Cursus monument. Excavation extents and key features were digitised from the interim report12 to map the archaeology at Inveresk. Excavations on the line of a new bypass in 1984 were designed to confirm that the linear cropmarks were part of the Roman temporary camp. The results confirmed the interpretation of the western linear features as Roman, but also that the eastern linear features, previously interpreted as a Roman road, were Neolithic in date.

**Presenting results**

A common product of any research is the presentation of the analysis or interpretation of the results. A prerequisite is a standardised terminology. Although HES airborne mapping and field survey mapping data both acknowledge the Scottish Monument Thesaurus, the poly-hierarchical relationship between a concept, broader terms and top-level terms introduces enough uncertainty and inconsistency to group and present related feature types in a meaningful way cartographically. Development and promotion of micro-thesauri to group the many hundreds of monument terms succinctly and inform the map legend would help address this challenge and help deliver interoperability across a range of observed data.

**Untapped potential**

Those best placed to document the extents and detail from fieldwork are those undertaking that work. Technological advances help. Widespread availability of high precision surveying equipment coupled with a marked growth in both developer-funded archaeology from the 1990s and community-led projects has democratised survey, be it on excavation, from survey or analysis.13 In keeping with the INSPIRE principle that data should be collected once and maintained at a level where this can be done most effectively,14 primary spatial data should be routinely shared with data curators. For instance, a tabulated gazetteer of sites recorded from a large walkover survey embedded in a pdf is far less useful than the equivalent data presented as a csv file, or preferably as a GIS dataset mapping site extents supported by detailed attribution. Similarly, a plan fossilised as an image in a pdf has less reuse potential than the original georeferenced data.

Sophisticated digital datasets require much richer attribution than ‘entry level’ event data required for established record purposes and basic information discovery. Appropriate attribution should enable a potential user to explore the character of the data15 at landscape (GIS layers) not project level (pdf).
There are already many highly skilled surveyors and geomatics officers working in archaeology. However, approaches are organisation focused rather than industry based. This is starting to change with large scale infrastructure projects like High Speed 2 which require consistent approaches across multiple contractors and sites. The profession needs to invest in developing and promoting the necessary data standards and demonstrate the benefits of a coordinated approach across techniques and organisations.

The bigger picture
With the establishment of the Geospatial Commission in 2018, the UK government aims to accelerate the delivery of economic, social and environmental benefits derived from geospatial data, products and services across the private and public sector. A key infrastructure project is the National Underground Asset Register, with the Scottish Community Apparatus Data Vault (VAULT) in Scotland, which is coordinating the complexity of buried services from multiple utilities to save time in decision making and save lives on site. In the same vein, archaeologists need to move beyond sensitivity mapping of known areas of high archaeological interest and locations of excavations, to routinely map and share buried archaeological contexts, the location of significant features and key contexts to deposit modelling. If the highly competitive construction industry can coordinate data sharing at scale, so can archaeology.

Archaeological and built heritage data needs to engage with digital twins: virtual models that accurately represent the physical city above and below ground. To do so requires access to detailed, accurate models of our heritage. Without those models, and as remote access to geospatial data for decision-making increases, the absence of archaeological data poses a risk to our past. If that data is invisible or inaccessible within an archive in the wrong format, it has no value in the emerging geospatial ecosystem. Developing consistent archaeological datasets not only makes managing that data more efficient, but also unlocks economic value through saving time and money.16

Spatial data is essential for the stewardship of the historic environment, not only in terms of managing specific risks at sites but informing the management of the wider landscape. Using a range of spatial data techniques, HES monitors the condition of and risk to properties in its care. In 2010 HES surveyors created a digital model of Skara Brae, a well-preserved Neolithic settlement and UNESCO World Heritage Site on Orkney, to inform site conservation, interpretation and engagement.17 A cyclical monitoring programme every two years enables comparison of the data over time against the 2010 benchmark to gain a better understanding of the effects of coastal erosion on the environs of the site.

Figure 5. (A) Fieldwork creates a wealth of archaeological data but without an effective blueprint in place the value of that data tin contributing to the wider archaeological landscape remains unrealised. (B) The sum of the parts: the temptation is to sort and organise data based on similarity – of technique, or by provider. (C) Something other than the sum of the parts: the true value of data can only be realised through an effective blueprint, acknowledging the FAIR Data principles, to build something other than the sum of the parts. (© Crown Copyright HES: Open Government Licence (nationalarchives.gov.uk)).
The data is not only essential for the stewardship of the site but also forms a case study for Dynamic Coast, a pan-government partnership developing an evidence base of national coastal change across Scotland.

To remain relevant in an increasingly digital society and realise the potential of expensively gathered data, archaeology needs to develop a shared infrastructure, to capture, share and publish a broad range of spatial data adhering to the FAIR Data principles. The manifesto is straightforward:

**Vision**

We will create an environment in which spatial data from archaeological research is shared openly, maximising its contribution to the study and stewardship of the past, and engages positively with the broader geospatial environment.
**Mission**

To develop a sustainable approach to collecting and sharing spatial data from archaeological research that increases efficiency within our discipline and releases the full potential of that data to the broader geospatial environment.\(^\text{16}\)

The value of curated data is already recognised through statutory requirements for spatial data, but the potential of primary archaeological data remains largely unrealised, compartmentalised by project and archive. With advances in geospatial technologies, it is increasingly essential to standardise and share data. Most of the building blocks needed to realise the potential of spatial data from primary archaeological data are in place\(^\text{16,17}\) but we lack agreed standards and infrastructure to harmonise and share that data. Without the relevant blueprints, we cannot transform data into information and knowledge effectively (Figure 5).

User requirements and specifications can be addressed through a collaborative approach between data curators, archivists and those creating data through fieldwork and research supported by robust training resources. The substantive challenge is developing, hosting and maintaining the relevant infrastructure to manage the long-term delivery of that spatial data to realise the both the societal and financial benefits of good data stewardship.

**Acknowledgements**

The arguments for developing a Spatial Data Infrastructure for archaeological data have benefited from discussions with colleagues at Historic Environment Scotland, the wider archaeological community in Scotland and from participants at conference and workshop sessions over many years. I would like to thank Susan Hamilton for reviewing an earlier draft of the paper and supporting the funding request.

**References**

In this paper, Peter McKeague highlights the ongoing disjuncture between the capabilities of spatial infrastructure with regard to the universal archival and sharing of archaeological data on the one hand, and our lack of actual exploitation of that potential on the other. The author presents a series of succinct examples of the significant and ongoing challenges: chief among these are, first, the continued production of field data in static document form including spatial data that is “fossilised” (in the author’s words); a byproduct of compliance-centered regimes that limits the future use of the data.

A second major challenge McKeague correctly highlights are the obstacles faced in developing universal archaeology data standards for what is still a very heterogeneous metadata landscape. Looking towards the future, projects aimed at creating a universally applicable ontology for capturing cultural heritage data, such as CIDOC CRM, may point towards the most comprehensive solution, but even if such ambitious undertakings meet their goals, they still require broad adoption across many frontiers to achieve their promise, something that not even the most optimistic visions see as a short-term likelihood.

McKeague rightly points out that such a level of data sharing is being achieved elsewhere (the INSPIRE spatial data infrastructure, for example); additionally, the ‘building blocks’ for a major paradigm shift in archaeology are already available. Indeed, the primary spatial data being generated by most current archaeological fieldwork encompasses a relatively small set of geospatial standards and formats, and this data could be easily and quickly collated into national and international archives. This would leave the thornier questions of metadata compatibility to the next generation of development of such an infrastructure but, as an alternative to the proliferation of parallel efforts, it is certainly attractive.

My own collaborative work (Trepal, Lafreniere, and Stone (2021)1) has attacked the problem from the opposite end, building a discrete community-focused heritage project incorporating data from the archaeological and historical records using standard GIS-based tools. This is driven by a desire to foreground the community heritage value of the data to boost the profile of the archaeological
data within community heritage regimes. The chief danger of this bottom-up approach in the long term is of course the proliferation of idiosyncratic small-scale digital projects. However, once again and, I think, to McKeague's point in this paper - the primary spatial 'building blocks' of the archaeological data within such projects are still readily compatible with most other standard archaeological primary geospatial data being collected worldwide, and could be collated into an international-scale spatial data infrastructure, with all the 'value-added' interpretive content and digital middleware reserved for the community-focused project that generated it.

The general thrust of the paper's argument, that most archaeological data remains 'silied', both within and without the discipline, even though the basic primary data is (in principle) easily interconnected using current off-the-shelf approaches, is therefore sound and articulates an urgent challenge given the increasing pervasiveness of geospatial data. With that being said, the author's examples offer a relatively narrow slice of what is truly a global problem within archaeology and cultural heritage. The situation in North America, despite the long-term development and maintenance of the Digital Archaeology Record (tDAR), is arguably worse than in Europe, where EU-funded work on digital ontologies and infrastructures for archaeology and heritage more broadly are at least grappling with the challenges inherent in creating international-scale data archives from the first. There is even less discussion within the discipline regarding developments outside Europe and North America, suggesting that even the bluntest critiques may be understating the scale of the challenge.

The case studies section does lose focus somewhat, with the subheadings not providing the same sharp focus as the introductory and concluding sections, even if the case studies bring forth some illuminating details. This paper's impact could be substantially enhanced if it were to include a critical review of the patchwork of archives currently in existence that attempt to curate digital, spatial archaeological data. Literature covering digital archaeology and heritage topics ages quickly, and the steady stream of digital initiatives being launched year on year means that reviews of the state of the field cannot come too often if we are all to collectively keep up. What is perhaps most lacking in the paper is a prescriptive roadmap towards a solution, even if only in outline form. For example, to what extent could INSPIRE serve as a direct model for an archaeological counterpart?

Critiques aside, this paper represents a cogent call for action in an area within archaeology that needs far more attention in the published literature than it receives at present, and to which the author has made, and continues to make, an outsized and well-appreciated contribution.

References

Is the topic of the opinion article discussed accurately in the context of the current literature?
Yes

Are all factual statements correct and adequately supported by citations?
Yes
Are arguments sufficiently supported by evidence from the published literature?
Yes

Are the conclusions drawn balanced and justified on the basis of the presented arguments?
Yes

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Industrial Archaeology, Digital Humanities, GIS

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 20 February 2023

https://doi.org/10.5256/f1000research.138763.r161843

© 2023 Novák D. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

David Novák
Institute of Archaeology of the Czech Academy of Sciences, Prague, Czech Republic

Peter McKeague's paper addresses the very important issue of archaeological spatial data management. There is no doubt that spatial data are far from fulfilling their potential to be the *lingua franca* linking individual archaeological datasets and data from other disciplines involved in landscape research. The author shows very well and with appropriate examples the difference between reporting the location of archaeological sites or monuments using maps and plans, and the creation of digital corpora of harmonized spatial data rich in information potential.

To the author's point of view, it can be noted that the harmonization of descriptive metadata is easier and, in a way, a secondary problem. The main challenge is to set up an archiving system for archaeological datasets that fully integrates spatial data describing archaeological research and its results, ideally in 3D space. Although there are already some attempts to grasp this problem, the barriers to realistic scaling up are countless. Even with appropriate standards in place, the actual competence of archaeologists and related positions to produce valid spatial data remains a significant problem. This is alluded to between the lines by the author at the end of the article, but it would be useful to emphasize this line a bit more.

The issues of data models, thesauri and especially coordinate systems are very complex and require the involvement of appropriate data management and GIS specialists from the very beginning of archaeological projects. Again, this is an issue of organisation and planning on the one hand, and of resources to fund similar positions and activities on the other. It is questionable whether, for example, current trends emphasizing the intrinsic value of FAIR scientific data (e.g.
activities related to the building of EOSC in Europe) could not also be used as a suitable way to promote better care of spatial data in archaeology.

In principle, it still appears that the fundamental problem is the approach to archiving the fieldwork results as such, which is primarily archiving PDF reports instead of archiving the primary data. This is confirmed by a recent survey carried out in collaboration between SEADDA and EAC (publication in preparation). It can certainly be argued that the letter "R" is the least fulfilled in the FAIR principles in the case of archaeological datasets, as the straightforward application of general archiving and descriptive standards leads to suppression of re-usability. Only a combined effort on various levels, especially organizational, competence, technical, and financial, can lead to change.

I am convinced that unless central organisations and key players in the industry fundamentally lower the threshold for other organisations, systemic change will not happen. It can be lowered both by providing support and education, but especially through widely applicable tools and services, with a very well-developed user interface and a quality backend. Until recording quality spatial data in archaeology is as simple as producing PDF reports, change will be very difficult to achieve.

It is unfortunate that the actual organizational issues and the problem of quality tools are more or less avoided by the author, and it would be interesting to add some of this insight. It would also have been helpful to focus more on the overall vision of how to achieve change, as the criticisms are undoubtedly valid, but without setting out a clearer path, the discussion will continue to stagnate or only be difficult to translate into practice.

In terms of readability of the text, the loose transition between general descriptions of the issues and specific case studies is sometimes problematic. This applies to roughly the second half of the paper. In this respect, better structuring and a slight reorganisation of the text would have helped. The positioning of the chapter 'Presenting results' is also not very clear, as it completely breaks out of the flow of the text. Also, the headings tend to be somewhat redundant and not very instructive ('Taking the holistic approach' / 'The bigger picture', etc.).

However, the article is undoubtedly a valuable call to action, and I fully agree with the views presented. The comments above are merely suggestive for the author's consideration, to improve readability and increase the potential impact. The proposed modifications are not of a fundamental nature.

Is the topic of the opinion article discussed accurately in the context of the current literature?
Yes

Are all factual statements correct and adequately supported by citations?
Yes

Are arguments sufficiently supported by evidence from the published literature?
Yes

Are the conclusions drawn balanced and justified on the basis of the presented arguments?
Yes

**Competing Interests:** In the last few years, I have been occasionally working together with the author (Peter McKeague) in the framework of SEADDA COST Action. Despite this connection, I honestly declare that this review has been produced in an impartial and unbiased manner.

**Reviewer Expertise:** Archaeological data archiving, GIS, research infrastructure management, landscape archaeology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com