One Archaeology: A Manifesto for the Systematic and Effective Use of Mapped Data from Archaeological Fieldwork and Research

Peter McKeague 1,*, Anthony Corns 2, Åsa Larsson 3, Anne Moreau 4, Axel Posluschny 5, Koen Van Daele 6 and Tim Evans 7

1 Historic Environment Scotland, John Sinclair House, 16 Bernard Terrace, Edinburgh, Scotland EH8 9NX, UK
2 The Discovery Programme, 6 Mount Street Lower, D02 T670 Dublin, Ireland; anthony@discoveryprogramme.ie
3 Swedish National Heritage Board, Storgatan 41, SE-114 84 Stockholm, Sweden; asa.larsson@raa.se
4 INRAP, 121 Rue d’Alesia, CS 20007, 75685 Paris, France; anne.moreau@inrap.fr
5 Keltenwelt am Glauberg, Am Glauberg 1, 63695 Glauberg, Germany; axel.posluschny@keltenwelt-glauberg.de
6 Flanders Heritage Agency, Flanders Heritage Agency, Herman Teirlinckgebouw, Havenlaan 88 bus 5, 1000 Brussels, Belgium; koen.vandaele@vlaanderen.be
7 Archaeology Data Service, Department of Archaeology, University of York, The Kings Manor, York YO1 7EP, UK; tim.evans@ads.ac.uk

* Correspondence: peter.mckeague@hes.scot

Received: 27 February 2020; Accepted: 14 April 2020; Published: 17 April 2020

Abstract: The Infrastructure for Spatial Information in Europe (INSPIRE) Directive (2007) requires public organisations across Europe to share environmentally-related spatial datasets to support decision making and management of the environment. Despite the environmental focus of INSPIRE, it offers limited guidance for archaeological datasets. Most primary data is created outside, but ultimately curated within, the public sector. As spatial evidence from fieldwork activities is not considered by the Directive, it overlooks a range of barriers to sharing data, such as project-based fieldwork, a lack of data standards, and formatting and licencing variations. This paper submits that these challenges are best addressed through the formalised management of primary research data through an archaeological Spatial Data Infrastructure (SDI). SDIs deliver more efficient data management and release economic value by saving time and money. Better stewardship of archaeological data will also lead to more informed research and stewardship of the historic environment. ARIADNE already provides a digital infrastructure for research data, but the landscape and spatial component has been largely overlooked. However, rather than developing a separate solution, the full potential of spatial data from archaeological research can and should be realised through ARIADNE.

Keywords: archaeology; geospatial data; spatial data infrastructures; data management

1. Introduction

Every year across Europe archaeologists create a wealth of primary data from fieldwork and research—irreplaceable records of the past. This information records the geographical location, extent, characteristics and relationships between data. Most of this data is born digital or manipulated digitally during analysis and report production. Despite considerable investment in data collection, and realization of the short-term value of that data through analysis and publication, its long-term potential remains unrealised. Barriers to sharing archaeological spatial data greatly hinder its potential
to positively impact research and society at large. The need to harmonise this data becomes all the more pressing with the growth of digital economies through government and private sector investment in geospatial technologies in decision making. Despite the need to keep pace with broader societal geospatial developments, there has been little coordinated engagement with the spatial value of archaeology within the discipline and no infrastructure to coordinate and share that data.

These challenges are not unique to archaeology but are common to any discipline working with spatial data. For example, geologists addressed these challenges by developing a thematic Spatial Data Infrastructure: One Geology [1,2], whose mission statement is simply, “Make web-accessible the best available geological and other geoscience data worldwide at the best possible scales, starting with at least 1:1 million scale” [3]. Building on the success of OpenStreetMap, open-source collaborative, community-based projects recognise the need for consistency in mapping approaches, and help address the need for high-quality mapping in developing countries [4]. The need for an archaeological, historical and heritage Spatial Data Infrastructure (SDI) has long been recognised [5,6], but little progress has been made to date.

2. Background and Context

The European Convention for the Protection of the Archaeological Heritage (revised), known as The Valletta Convention (1992) [7], established a high-level series of principles aimed at the conservation and enhancement of archaeological heritage through sharing knowledge and experience. The Convention requires that signatory states maintain “an inventory of its archaeological heritage and the designation of protected monuments and areas” (Article 2i); that “for the purpose of facilitating the study of, and dissemination of knowledge about, archaeological discoveries, each Party undertakes . . . to make or bring up-to-date surveys, inventories and maps of archaeological sites in the areas within its jurisdiction” (Article 7i); and that the importance of public awareness is recognised in valuing and “understanding the past and the threats to this heritage” (Article 9). Although the convention sets out a framework for best practice in archaeology, implementation is determined by the legislative and organisational structures of individual countries or even regionally.

Archaeological inventories typically record information about sites and monuments within a defined jurisdiction. Originally established to provide the evidence base for a range of decision making activities about the protection and stewardship of the archaeology [8–11], online access to inventory data has greatly increased public engagement with heritage.

Acknowledging that an increasing number of environmental policies across Europe have a strong spatial dimension, much of which is trans-national, the European Commission published the INSPIRE Directive in 2007 [12]. The Directive was subsequently transposed into the national legislation of member states in 2009. INSPIRE sets down the general rules for sharing environmentally-related spatial datasets to support environmental policies and decision making across Europe through an SDI. An SDI may be defined as ‘the relevant base collection of technologies, policies and institutional arrangements that facilitate the availability of and access to spatial data’. The SDI provides a basis for spatial data discovery, evaluation and application for users and providers within all levels of government, the commercial sector, the non-profit sector, academia and by citizens in general’ [13]. Data should be discoverable through metadata catalogues and published as view (Web Map/WMS) and download (Web Feature/WFS) services.

INSPIRE identifies 34 thematic datasets [14] relevant to the stewardship of the environment of which only one, The Protected Sites theme, specifically addresses historic environment data. Protected Sites may be formally designated or managed specifically through international or national legislation, or informally through other effective means [15]. Despite the mandate, there is no consistent approach to implementing the Directive for archaeological sites, or indeed the built heritage, across Europe. This is in part due to legislative differences. In some countries, particularly in Scandinavia, any sites older than a certain date are automatically protected through legislation. Elsewhere, only a small proportion of sites recorded in inventories are formally designated and
published as INSPIRE services, whilst there is inconsistent coverage at the regional level for undesignated sites. For instance, in Scotland, although about 7% of known archaeology is formally protected, Planning Advice Note 2/2011 [16] recognises the significance of undesignated sites as non-renewable resources containing unique information about how Scotland’s historic and natural environments developed over time, and its contribution to local, regional and national identities. The guidance empowers local authority archaeology services to advise on the protection of archaeology within the planning process. In 2016–2017, 8% of all planning applications in Scotland monitored by the Association of Local Government Archaeological Officers (ALGAO) members required archaeological mitigation [17].

The technical specifications for the Protected Sites theme define a very limited set of attributes and values covering a broad range of datasets for both the historic and natural environment which are suitable for general use but do not capture the full complexity of archaeological data. More tailored approaches are required to model the complexity of archaeological data. Germany developed the Archaeological Data eXport Standard (ADeX) to address the need to share common attributes across the culturally independent Federal States for large research and infrastructure projects [18,19]. In recognising the benefits offered by an SDI for rendering data interoperable, and the limitations of the existing technical specifications within INSPIRE [15], Fernández Freire et al. [20] proposed a much richer data model for Spanish archaeological datasets. Many countries routinely publish richer attribution in their spatial data than required by INSPIRE, or provide links to further online resources such as webpages or structured data in JSON, GeoJSON or XML format [21].

3. Archaeological Events: Thinking Beyond the Monument

Traditional approaches to archaeological inventories, often originating from paper records, reflect the need to document the monument. They typically record the location, age and type of site, and are supported by descriptive notes summarising the history of investigation, notable characteristics and selected artefactual evidence. These individual pieces of fieldwork or observation are the activities or ‘events’ that led to the initial recognition or subsequent investigation of a monument. An event represents “a single episode of primary data collection over a discrete area of land. This single recording event can only consist of one investigative technique and is therefore a unique entity in time and space” [22]. Each event should be uniquely and persistently identified in the record system. Events, such as excavations or surveys, are implicit within early inventory systems. With the development of computerised records and Geographic Information Systems (GIS), the need for more sophisticated data models to explicitly document an event becomes apparent. Events may relate to one or more records in the database. Occasionally an event may not lead to the definition of a monument record.

‘Events’ is a catch-all term for activities that document our past, including a range of remote sensing techniques, aerial photograph transcriptions of archaeological features revealed as cropmarks, field survey and field walking, and full-scale excavation. Whereas the core values for documenting a monument include location, age and classification, event documentation should include as a minimum what type of work was undertaken, by whom and when it took place, as well as the location and extent of the work.

The Event–Monument model is well established across Historic Environment Records in Britain. Commercially available software packages include HBSMR [23] and ARCHES [24], to manage Event data within the context of Monument records. In Wales, the four Archaeological Trusts use Archwilio [25] to manage and publish their inventories in a single portal. However, the natural focus of online publication remains the monument (the Protected Site) rather than the activity. The value of primary data are left in supporting role informing interpretation rather developed as separate but related geospatial time series resources.
3.1. Organisational Variables in Archaeological Fieldwork

There is no single model for the organisation of fieldwork across Europe. In some countries, fieldwork is restricted to universities, museums and state archaeology services, while elsewhere there are thriving competitive commercial archaeology practices and community-led projects. With so many agencies working in a deregulated environment, there is an overreliance “on users to self-police when it comes to quality and report critical information that others might use to evaluate data as metadata” [26] (p. 78). Lack of standardization in fieldwork recording and reporting practices between projects and organisations, compounded by proprietary software, also forms a considerable barrier to reuse, assuming one can find the data in an archive.

Information flow from field to record to archive can be erratic. There is often no requirement to deposit project results in an archive. Even where best practice encourages deposition, there is almost no monitoring of the lifecycle of a project, even in regulatory systems. Many archives do not specialize in archaeological data with information added to general collections. They are often ill-equipped to manage digital data. Less than five European countries “have repositories with the required specialist knowledge and mechanisms in place to ensure archaeological data will be freely and openly available for re-use by future generations of researchers” [27].

3.2. Reporting Variables in Archaeological Fieldwork

The amount of fieldwork undertaken across Europe each year is not easily quantified, often with no legal requirement to report each investigation. Summary accounts of fieldwork projects may be published in yearbooks such as Discovery and Excavation in Scotland [28] or Das archäologische Jahr in Bayern [29]. These volumes provide prompt descriptive accounts—and in some cases the only account of a project—but they represent little more than an interim statement, often lacking site plans.

Several European countries have established digital workflows to manage the transfer of project reports from those undertaking the fieldwork to the relevant curatorial records. In England and Scotland, most local authority archaeologists require work undertaken as a condition of planning consent to be reported through OASIS (the Online AccesS to the Index of archaeological investigationS) [30,31]. The use of OASIS is voluntary for other research. OASIS users can upload a digital copy of the project report with an option to transfer a digital boundary file. The project report and data may be downloaded by the relevant local authority archaeological service and the national record. A copy of the project report may be available online through these records or through the Archaeology Data Service Library [32] and ArchSearch catalogue [33]. In Northern Ireland completion of an OASIS form is part of granting an excavation licence [34].

In Flanders [35] and Sweden [36] it is a legal requirement to provide digital copies of project reports and other data. Since 2016, it has been a legal requirement in Flanders for all archaeologists (both commercial and academic) to submit intermediate and final reports, including the digital extents of the project, to the Flanders Heritage Agency. Failure to submit has serious consequences for spatial planning initiatives [35]. Two publicly accessible digital layers are created: archeologienota’s and nota’s (reports of preliminary research) and eindverslagen (final reports), offering a geographic view of the data [37], a more traditional alphanumeric search interface [38] or geospatial geometries as downloads or WMS services [39]. The Swedish National Heritage Board require that all archaeological projects ordered by the County Administrative Boards, in effect all excavations and most of the surveys, must now be registered in the Historic Environment Record as downloadable spatial data with the perimeter of the project and excavated trenches, as well as certain administrative metadata. Submitted project reports are available to the public through an e-archive.

Where implemented, digital reporting forms have greatly improved visibility and accountability of fieldwork ensuring that project reports are available online promptly. Online forms referencing controlled terminologies provide a structure to capture project and event metadata consistently. Users can upload and transfer digital boundary files alongside copies of the project report, but a
wealth of spatial data essential to our understanding of the archaeological landscape lies fossilised and inaccessible in PDF files.

4. Spatial Data and the Technique-Dependent Challenges within Archaeological Fieldwork

Archaeologists should as a matter of practice provide accurate spatial data from their research for reuse in archaeological inventory systems. This event data should be collated and shared online to act as a spatial finding aid to material deposited in archives.

Minimum documentation should include who undertook what type of work, when and where. In record systems, the location (where) is often recorded as a coordinate rather than represented by the spatial extent of the activity. Additional metadata, which can be recorded as part of a spatial dataset, is technique dependent. For more sophisticated techniques such as remote sensing surveys, essential documentation should also include the instrumentation used to collect data and the processes applied to transform that data. For vector datasets, the terms of reuse should also be clearly recorded against the data as part of the dataset’s attribution to help others assess if the data is fit for purpose and understand its reuse constraints. For raster data, the information should be captured in an accompanying index file.

4.1. Invasive Techniques

A range of invasive techniques may be used to investigate archaeological sites. Approaches include test pitting to sample the nature of the topsoil and subsoil, trial trenching to evaluate the presence/absence, nature, preservation, age and extent of any buried archaeological features to full-scale excavation. Watching briefs may be conducted during any operation carried out for non-archaeological reasons that may damage archaeological deposits whilst borehole investigations can also inform our understanding of stratigraphy.

Excavation

As a destructive process excavation and related approaches are argued to be ‘preservation by record’ [40,41], in which case it is essential that the record is firstly preserved but also findable, accessible and reusable through record and archival systems. Excavations may be undertaken for research and community engagement, but particularly in response to planning applications by a range of organisations often using their own recording systems. Knowing exactly where investigations took place and the character of the archaeological remains are fundamental to the recording process. The project area, location of trenches opened and features recorded are provided as illustrations within the report. However the raster format of most pdf reports hinders easy reuse in a GIS. The project archive may contain plans, databases and spreadsheets in reusable formats but the archive is often not easily accessible. Where the data can be accessed, individual recording methodologies can further frustrate reuse.

The benefits of a standardised approach have been recognised in France where most ‘preventive’ excavations are undertaken by the Institut National de Recherches Archéologiques Préventives (INRAP), although since 2003 a number of private companies also offer archaeological services. In 2018, INRAP undertook 1934 assessments and 225 excavations across France and its overseas departments using a standardised recording system. As well as documenting the excavation and satisfying planning requirements, INRAP recognised the value of data as contributing to scientific research and public welfare [42].

Since 2011, INRAP has invested in an extensive GIS training programme to record fieldwork. Moving from a localised recording tool to an organisation-wide application highlighted several problems, including how to define and harmonise data collection to ensure data could be easily shared [42]. The result is a tightly-defined specification for six separate layers of spatial data (Table 1) (Figure 1), all using Lambert93 (EPSG code 2154) based on the legal geodetic reference system RDF93—the French version of ETRS89.
Table 1. Six layers of excavation record (Figure 1) (after [36]).

<table>
<thead>
<tr>
<th>INRAP Layer</th>
<th>Form</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum requirement for Planning purposes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perimeter Polygon</td>
<td>Polygon</td>
<td>Defines the full extent of the area to be investigated</td>
</tr>
<tr>
<td>Opening Polygon</td>
<td>Polygon</td>
<td>Defines the actual extents of excavation trenches, test pits</td>
</tr>
<tr>
<td>Polygonal observations</td>
<td>Polygon</td>
<td>The location and extents of features recorded during excavation</td>
</tr>
<tr>
<td>Punctual observations</td>
<td>Point</td>
<td>The location of archaeological remains and find spots</td>
</tr>
<tr>
<td>Samples</td>
<td>Point</td>
<td>The locations of samples and geomorphological observations</td>
</tr>
<tr>
<td>Sections</td>
<td>Line</td>
<td>Location of sections</td>
</tr>
</tbody>
</table>

Figure 1. Spatial representation of excavation datasets (credit: INRAP-Caviar, 2020) described in Table 1.

Each feature has a unique ID corresponding to the ID in the related archaeological database. Where possible, standardised terminologies reference the PACTOLS thesaurus [43,44] thereby meeting Semantic Web and Open Science standards. A hyperlink from the GIS takes the user to related resources on Dolia—the online catalogue and digital library of INRAP [45]. Currently, public access is restricted to locational data, rather than the full spatial extents on INRAP’s Archaeozoom platform [46], or as one of the datasets on the ARIADNE portal [47].

At present, the INRAP approach is not used universally across France. Their system does not include project data from commercial companies or data from university research. Data from 20th-century projects, captured in the PATRIARCHE system will be available through the Pleade Portal [48,49] but again not integrated into a national excavation dataset. Nevertheless, the INRAP approach provides a model for the organisation of geospatial data from excavation greatly reducing the challenges of integrating non-conforming data and legacy data into geospatial databases.

By investing in data standards and a robust training programme, there is a consistent approach to data recording and management from INRAP’s excavations. Elsewhere in Europe, a deregulated approach to excavations has resulted in multiple approaches to documentation, creating barriers to the reuse of data. Such barriers can be overcome through semantic approaches [50], but these techniques require a high level of understanding of Linked Data to maximise the value of spatial data. Developing
consistent methodologies, standardising vocabularies and investing in training is required to develop a consistent dataset.

4.2. Non-Invasive Techniques

Archaeologists employ a range of survey techniques to discover and document archaeological sites. Although some surveys target a specific site, others adopt a landscape approach to document archaeological evidence over a large area. Technological advances in geophysical survey techniques and the availability of LiDAR now enable investigation of much larger areas than traditionally undertaken. Despite the need for, and advantages of, consistent symbology, as used by every single national mapping agency, there is a lack of guidance and standards for clearly presenting archaeology observed from survey techniques in print or as a GIS layer. Consequently, there is little consistency in data visualisation between projects or across techniques.

4.2.1. Airborne Mapping

In many countries, there has been a long tradition of recognising and mapping (transcribing) individual archaeological sites from oblique and vertical aerial photographs. Computerised rectification programmes transformed the laborious manual process of transcribing archaeological features from aerial imagery. The need for consistent representation of mapped features emerged when data from these processes were combined within a GIS to move from a site-focused understanding to a seamless view of the archaeological landscape (Figure 2) [51].

During the 1990s, Historic England, The Royal Commission on the Ancient and Historical Monuments of Scotland (RCAHMS) (now part of Historic Environment Scotland) and The Royal Commission on the Ancient and Historical Monuments of Wales (RCAHMW), embarked on large-scale mapping programmes of cropmarks. Each organisation developed their own approaches, resulting in similar but subtly different specifications. For the Scottish data (Figure 2) [51], cropmarks are presented using a simple map legend largely informed by the top-level classifications from the Scottish Monuments Thesaurus [52] but customised for cartographic presentation. Depending on the underlying attribution, it would be just as possible to symbolise the mapped archaeological landscape by archaeological component or by period [53].

Data from these national mapping programmes has been shared with local authority archaeological services to assist with managing the archaeological resource. The data is not easily accessible for others working in archaeology, such as researchers and consultants. It could be easily published as View and Download Services to help these users work more efficiently.

4.2.2. Field Survey

National mapping agencies publish selected archaeological sites as topographic features on their map products. With paper maps only a selection of nationally important archaeological sites could ever be published (Figure 3) [54]. Through GIS View and Download Services, it is now possible to present a more complete view of the archaeological landscape spatially (Figure 4) [55] yet we are failing to capitalise on that potential.

The availability of digital surveying tools, including Electronic Distance Measurement Total Stations (EDM) and differential Global Positioning Systems (dGPS), enables archaeologists to undertake site surveys and map extensive archaeological landscapes systematically. Data gathered forms the basis of illustrations in project reports and may be deposited as part of the project archive. Once archived, data from separate survey projects is difficult to retrieve, let alone harmonise with data from other projects. Fundamental challenges include data discovery and copyright issues, whilst the lack of consistent data standards in documentation presents an intellectual hurdle to overcome.

In Britain, archaeological survey had long been the preserve of the respective Royal Commissions on Ancient and Historical Monuments in England, Scotland and Wales, who dutifully produced inventories describing and illustrating individual monuments. Survey plans were generally divorced
from their landscape context in these volumes. In response to afforestation programmes in 1980s Scotland, the RCAHMS commenced a strategic mapping programme, the Afforestable Land Survey, which developed in-house data standards for the display of mapped detail [56] aligned with the top-level classification of the Scottish Monument Thesaurus [52].

Figure 2. In Scotland, common attribution and a simple symbology enables data from individual cropmark transcriptions (top right) to be combined seamlessly. The data may then be viewed as an archaeological landscape in a Geographic Information Systems (GIS) and published online through Canmore [51] (bottom), the online portal of The National Record of the Historic Environment of Scotland and potentially as View and Download services. Background mapping © Copyright and database right 2020 Ordnance Survey licence number 100057073.
Following the introduction of Planning Guidance in the 1990s, the near monopoly of archaeological field survey in Scotland held by RCAHMS was broken when commercial archaeology companies began to undertake surveys on behalf of their clients. The deposition of the project report with, and acceptance
by the relevant local authority service, satisfies their clients need to discharge a planning condition. If projects are reported through the OASIS form they are usually available through the Archaeology Data Service [30,31]. Reusability of the data is often not considered, compromised by the format of the survey product, lack of access to the underlying data and the absence of relevant data standards.

4.2.3. Legacy Surveys and Site Plans

Whilst the focus of this paper is data created through born digital processes, the value in translating analogue plans and drawings to digital formats for use in a GIS should not be overlooked. After all, these surveys and plans form a substantial part of the archaeological record and provide unique insights into the understanding of a site at the time of survey.

By means of an example, research at the Mayan city of Copán, Honduras, drew heavily on archived plans from previous investigations [57]. Challenges included combining plans from a range of scales, from different survey methodologies for reuse in a GIS. The researchers successfully combined 1:200 photogrammetrically-derived excavation plans with 1:2,000 plane table and alidade surveys augmented by details from the associated collection of original field notes, to inform their understanding of the city. The legacy data was georeferenced and combined with a Digital Terrain Model derived from LiDAR data to improve the archaeological map of the study area and inform surface modelling of the archaeology in virtual environments. An important part of the datification of analogue plans is the paradata describing the transformation process. Whilst not essential for the immediate use of the digitised product, documentation enables a critical analysis of project results in the future.

4.2.4. Remote Sensing Techniques

With ever-increasing computer processing power, the range of digital recording techniques has expanded and become increasingly more sophisticated. Compared to the targeted site surveys of just over a decade ago, geophysical surveys now cover large landscapes. Structure from Motion (SfM) enables the creation of 2.5D surface models from multiple images. Laser scanning produces incredibly detailed models from point clouds, not only of archaeological objects and sites but also extensive landscapes as well as the built heritage. Although there is guidance on the metadata required to document these techniques [58,59], the spatial extents of these surveys are generally overlooked.

Geophysicists employ a range of techniques, including magnetometry, earth resistance and Ground Penetrating Radar to identify sub-surface anomalies. The technical metadata requirements for documenting these surveys are well established and are outlined in the Archaeology Data Service/Digital Antiquity Guides to Good Practice for Remote Sensing [58]. The metadata is published as part of the project report. However, in England and Scotland metadata may be entered through the OASIS form by the practitioner. The metadata can then be exported to the Historic England Geophysics database [60] ensuring that the database is updated without additional rekeying or miskeying of data. The database enables users to find where surveys took place but does not record the extent of those surveys. Simply linking the metadata already provided as attribution for the spatial extents would create a powerful map index, enabling viewers to easily discover the area that has been surveyed and to what resolution.

Geophysical surveys produce spatial datasets documenting the location and extent, results and interpretation for each method used (Table 2). In Britain, some local authority archaeologists receive a georeferenced raster of the survey products that can be used directly in their GIS, though any vector and raw data is often retained by the practitioner.
Table 2. Expected spatial data products from geophysical surveys.

<table>
<thead>
<tr>
<th>Output</th>
<th>Preferred Format</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Survey extent</td>
<td>Vector Polygon</td>
<td>Extent of survey grids</td>
</tr>
<tr>
<td>Grey-scale raster plots</td>
<td>Raster grid data and associated images surrogates</td>
<td>May include both the unfiltered and filtered rendering of the survey data</td>
</tr>
<tr>
<td>Geophysical interpretation</td>
<td>Vector polygon and line features</td>
<td>Analysis of survey results highlighting the interpretation of geophysical anomalies</td>
</tr>
<tr>
<td>Archaeological interpretation</td>
<td>Vector polygon and line features</td>
<td>Archaeological interpretation of the geophysical anomalies</td>
</tr>
</tbody>
</table>

Where vector data is accessible, practitioners are rarely consistent in their transcription and categorisation of either the geophysical or archaeological anomalies. Again, the lack of standardised vocabularies hinders interoperability between data provided by different practitioners. Although there is guidance on the practice [61] and archiving [58] of geophysical survey data, there is no consistent approach to labelling interpretation data with agreed terminologies.

Many countries have made publicly funded LiDAR or Airborne Laser Scanning (ALS) data conforming to INSPIRE Annex II Elevation Theme specifications [62], available under a range of licencing conditions. Examples of portals where data may be downloaded under Open licences include England [63], Scotland [64], Wales [65], Ireland [66] and Flanders [67]. Elsewhere, access and reuse is more closely managed. Although LiDAR derived data is available for the whole of Germany, Federal States adopted different approaches to releasing data. In Northrhine-Westfalia the data may be freely downloaded at 1m resolution, but in Hesse, a comparable dataset may be used within the governmental bodies like the State Heritage Service, whereas the public pay to access the data. A list of available Open LiDAR data sources was compiled by Kokalj and Hesse in 2017 with the caveat that “Host sites will inevitably change and new datasets are becoming available regularly, so use your favourite search engine” [68], (p.13). These sources are discoverable through the INSPIRE Geoportal [69] but, as invaluable research data for archaeologists, could be signposted from the ARIADNE portal [47].

LiDAR data may be manipulated to create a series of visualised datasets including hillshade and Sky View Factor (SVF). Visualisations enable archaeologists and others without extensive technical knowledge to visually examine the data under different conditions. Visualisations may be created for project purposes through tools such as the Relief Visualisation Toolkit [68,70], or created once and shared with both archaeologists and the wider geospatial community.

In Sweden, hillshade and SVF visualisations are available as layers in the public search portal of the Historic Environment Record [71]. However, access to the LiDAR data for analysis and processing requires purchasing it from the government’s Land Survey Agency. In Belgium, The Flanders Heritage Agency created multidirectional hillshade and SVF visualisations from Open LiDAR data commissioned by another Flemish Government agency, Information Flanders. In turn, these visualisations have been published under an Open Data license allowing free reuse, even for commercial purposes, as long as attribution is respected [67,72].

For many users it is often sufficient to know the date of capture, resolution, ownership and licencing conditions. For visualisations of the raw data as Digital Surface Models (DSMs) or Digital Terrain Models (DTMs), Doneus and Briese [73] identified several factors that “have a considerable impact on the quality and usability of the final result” for archaeological research in areas of afforestation (Table 3). “Without that knowledge, ALS becomes a kind of ‘black box’, where the derived DTM is used without further knowledge of underlying technology, algorithms, and metadata. Consequently, there is a certain risk that the data used will not be suitable for the archaeological application”. This data is typically included within the project documentation, usually as a PDF accompanying the survey data. It would be much more effective as part of the attribution of a searchable spatial index.
Table 3. Core metadata and extended paradata for documenting a LiDAR survey (after [73]).

<table>
<thead>
<tr>
<th>Minimum metadata required to document a survey</th>
<th>Name of survey Brief description of survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td>20mm</td>
</tr>
<tr>
<td>Brief description</td>
<td>Fixed-wing aircraft;</td>
</tr>
<tr>
<td>Sensor</td>
<td>Drone</td>
</tr>
<tr>
<td>Platform</td>
<td>Minimum number of points per metre</td>
</tr>
<tr>
<td>Point Density per metre(Resolution)</td>
<td>Date survey flown</td>
</tr>
<tr>
<td>Capture Date</td>
<td>Copyright statement and terms for reuse</td>
</tr>
<tr>
<td>Licence</td>
<td>Who undertook the survey</td>
</tr>
<tr>
<td>Originator</td>
<td>Place of deposit</td>
</tr>
<tr>
<td>Custodian</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Additional paradata required by Doneus and Briese [73]</th>
<th>Purpose</th>
<th>Planning condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purpose</td>
<td>Topography; Archaeology</td>
<td></td>
</tr>
<tr>
<td>Scanner type</td>
<td>Planning condition Conventional; Full-Waveform</td>
<td></td>
</tr>
<tr>
<td>Scan Angle(Whole field of view)</td>
<td>Value, e.g., 15°</td>
<td></td>
</tr>
<tr>
<td>Flying Height above Ground</td>
<td>1100m;</td>
<td></td>
</tr>
<tr>
<td>Speed of aircraft</td>
<td>65 m/s</td>
<td></td>
</tr>
<tr>
<td>Laser Pulse Rate</td>
<td>36 m/s</td>
<td></td>
</tr>
<tr>
<td>77.500 Hz</td>
<td>100.000 Hz</td>
<td></td>
</tr>
<tr>
<td>Scan Rate</td>
<td>60 Hz</td>
<td></td>
</tr>
<tr>
<td>Strip Adjustment</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Filtering</td>
<td>Robust Interpolation (SCOP++)</td>
<td></td>
</tr>
</tbody>
</table>

LiDAR visualisations are used increasingly in archaeology [74–76]. For example the State Heritage Service of Baden-Württemberg uses LiDAR data to systematically investigate unknown archaeological sites across the state, particularly in forested areas, with great success [77]. Of course, data used in the planning system is also of considerable value to research and vice versa. The Keltenwelt am Glauberg Research Centre, part of the State Heritage Service of Hesse, uses LiDAR visualisations as a tool to research hillforts and burial mounds across the state, and shares the results with the state record.

Manual identification of features from LiDAR is now being augmented by machine learning techniques. Deep Learning techniques—Convolutional Neural Networking [78] and Regions-based Convolutional Neural Networking [79]—have successfully trained computers to recognise archaeological features from the data. With the development of machine learning approaches to feature recognition, clear and transparent paradata is essential to inform potential users of the data and processes applied to recognise features. It is suggested here that this attribution should be managed and viewed as part of the spatial data for that event, as well as for any descriptive account in a site record.

5. The Case for an Archaeological Spatial Data Infrastructure

The fundamental issues of long-term data preservation, archival challenges and accessibility are being addressed through the SEADDA (Saving European Archaeological Data from the Digital Dark Ages) COST Action [27]. European initiatives do address aggregating data from multiple organisations, but the emphasis is firmly on the cultural significance of an object or site rather than its spatial value in a landscape context. Europeana [80] provides access to a wealth of digital imagery and virtual collections. ARIADNE [47] and its successor project ARIADNE Plus [81] is building a digital infrastructure that pools data from multiple contributors to create a European wide interoperable dataset. ARIADNE recognizes that there is a wealth of accumulated knowledge of individuals, teams and institutions that forms a vast and fragmented corpus but that their potential has been constrained by difficult access and non-homogenous perspectives. The technological solutions applied
enable users to cross-search multiple datasets through gazetteer and map-based searches to find information about a site or object held in a collection. Users can drill down to data but then cannot explore the spatial relationships between results. ARIADNE provides a digital infrastructure and uses some location-based searching but it is not an SDI; that is the full spatial value of the data is not realised. Development of an archaeological SDI, complementing existing infrastructures, would address a knowledge gap and add value to existing research data initiatives through map-based indexes.

5.1. Challenges

Digital technologies have transformed data collection and processing from primary research. They have also democratised those processes by opening up survey and interpretation beyond established institutions in many countries. McCoy [26] identifies volume, velocity, variety, veracity, visualization and visibility as characteristics of Geospatial Big Data; characteristics that challenge the capacity of established record systems to address or many specialist and non-specialist users to digest. These challenges can be partially mitigated through improving regulation of data collection through definition and adherence to standards and a commitment to archive in an appropriate digital repository or cultural memory institution. Data needs to be findable, not only through online catalogues but as View and Download services which should act as finding aids for richer associated data in the archives.

5.1.1. Data Themes

Inventories represent the collected knowledge and interpretation of locations derived from evidence (the archive) provided by events and other sources. Through online database and Web-GIS systems, public perception, and the understanding of many archaeologists, is that the archaeological record is site-focused, with individual event descriptions contributing to the narrative or biography of individual records.

This perspective undervalues the importance of primary event data as separate but related resources. The event data represents time-series information—unique observations in time and place which may relate to one or more sites. With the landscape approach being developed for remote sensing surveys, it is all the more critical to recognise data from these projects as resources in their own right rather than as an adjunct to the site record. As spatial data, it must be discoverable, accessible and reusable in its own right.

The potential of these spatial datasets remains largely unrealised, with few systematic efforts to collate and publish data from different sources at a regional or national scale. Datasets for consideration should include, but not be restricted to, the range of invasive and non-invasive techniques reviewed above. The SDI also needs to define approaches to 3D data generated from surface modelling, excavations and the built heritage. Each dataset should be supported by the minimum technical specifications documenting the character of the data. These specifications should not constrain innovation in reuse and analysis. Development of the SDI must be supported by investment in training and promotion including celebrating best practice. Although primarily addressing the challenges of integrating, manipulating and analyzing data from primary research, the SDI also needs to engage broader concepts surrounding spatial thinking [82].

5.1.2. Framework

Through INSPIRE, there is already a robust infrastructure for delivering public sector environmental data. Opensource (e.g., Geoserver and Mapserver) and proprietary (e.g., ArcGIS Server and ERDAS Apollo) software platforms provide for managing data internally and publishing online. The technical guidance for building and publishing View and Download Services is well established for both vector and raster datasets. What is lacking is the mandate or requirement to standardise and share data between different branches of the profession. To make the case for a step-change in how we manage spatial data from primary research, we need exemplar datasets to demonstrate how
standardising and sharing spatial data can deliver efficiencies across the profession for the benefit of the wider society.

5.1.3. Hosting

To compile data from multiple sources, several factors must be considered. First and foremost is who should publish the data. Archival structures are designed to organise projects within collections reinforcing the project-based character of fieldwork. Emphasis is on preservation and reuse at a project level. Archives accrue data but generally have no requirement to develop additional geospatial geometries from separate projects through integration into a single dataset. A notable exception is the National Library of Scotland Map Library whose online map catalogues provide access to georeferenced historic mapping in addition to individual images of maps in their collection [83,84]. Creation of searchable spatial indexes to collections through integrating key data could act as an invaluable finding aid in web browsers, signposting users to additional resources in the archives. Such initiatives are often project-driven or the inspiration of driven individuals. For instance, the Canadian Archaeological Radiocarbon Database (CARD), maintained by the Canadian Museum of History and the Laboratory of Archaeology University of British Colombia, collates data from multiple sources into a single database [85]. However, the long-term viability of such initiatives needs to be supported beyond project funding or enthusiasm.

Under INSPIRE, the organisation managing mandated Protected Sites data has a clear duty to publish that data. There is no such requirement for primary data from archaeological research. Efficiency dictates that similarly-themed data should be collated and published as a single dataset rather than published by multiple data owners. Currently, there is no legal requirement to collate and publish archaeological event data spatially. As this data informs our understanding of the archaeological landscape, it would seem a natural extension for organisations already publishing Protected Sites services to host additional services for primary datasets.

5.1.4. Licencing

The success of an archaeological SDI depends on dismantling the barriers to sharing data between different branches of the profession. Building on the principles of Open Access publications, archaeologists should adopt a culture of Open Data so that data created by one part of the profession can be reused without undue restriction by others. However, retention of sensitive data because of concerns over commercial sensitivity, ongoing research or embargoed data must be addressed through robust access rights management. Spatial data may include third-party content, typically from national mapping agencies, whose intellectual property rights may restrict reuse and needs to be acknowledged in any licencing.

5.1.5. Formats

As discussed above, archaeologists still insist on condemning rich data formats to raster formats, which severely limits interoperability [86]. An archaeological SDI should promote the use of appropriate open data formats. Data captured or created in a vector format should be available for reuse but despite the flexibility of vector digital formats, the final outputs are often visual renderings of the data as images that lack intelligence. These formats restrict reuse with the redigitisation of plans and rekeying of gazetteers being both inefficient and inaccurate.

Providing data in reusable formats referencing open data standards would be far more efficient for data curators charged with integrating results into record systems, and make the data more accessible to researchers.

In 2013, the G8 summit published The Open Data Charter which recognised that although governments and businesses collect a wide range of data, they do not always share it in ways that are easily discoverable, useable, or understandable by the public [87]. Acknowledging the missed opportunities caused by data retention in institutional silos, the Charter calls for greater transparency
through Open Data to increase awareness, innovation and efficiencies. However, data can take many formats, as illustrated by the incremental Five Star schema for Linked Open Data defined by Tim Berners Lee (Figure 5) [88]. This schema is usually expressed for textual data where One Star data is simply data available on the web but with an Open Licence. Two Star data is available as machine-readable structured text in proprietary formats (e.g., Excel), with Three Star data held in non-proprietary formats (csv). Four Star data conforms to open standards from WC3 (e.g., Resource Description Frameworks) and SPARQL to identify things so that people can point to your resources, and Five Star data helps you link your data to other people’s data to provide content.

5.1.5. Formats

As discussed above, archaeologists still insist on condemning rich data formats to raster formats, which severely limits interoperability [86]. An archaeological SDI should promote the use of appropriate open data formats. Data captured or created in a vector format should be available for reuse but despite the flexibility of vector digital formats, the final outputs are often visual renderings of the data as images that lack intelligence. These formats restrict reuse with the redigitisation of plans and rekeying of gazetteers being both inefficient and inaccurate.

Providing data in reusable formats referencing open data standards would be far more efficient for data curators charged with integrating results into record systems, and make the data more accessible to researchers.

In 2013, the G8 summit published The Open Data Charter which recognised that although governments and businesses collect a wide range of data, they do not always share it in ways that are easily discoverable, useable, or understandable by the public [87]. Acknowledging the missed opportunities caused by data retention in institutional silos, the Charter calls for greater transparency through Open Data to increase awareness, innovation and efficiencies. However, data can take many formats, as illustrated by the incremental Five Star schema for Linked Open Data defined by Tim Berners Lee (Figure 5) [88]. This schema is usually expressed for textual data where One Star data is simply data available on the web but with an Open Licence. Two Star data is available as machine-readable structured text in proprietary formats (e.g., Excel), with Three Star data held in non-proprietary formats (csv). Four Star data conforms to open standards from WC3 (e.g., Resource Description Frameworks) and SPARQL to identify things so that people can point to your resources, and Five Star data helps you link your data to other people’s data to provide content.

5.1.6. Data Transfer

An efficient workflow to streamline reporting from field to record is key to delivering an archaeological SDI. As discussed above, online applications already exist in some countries to manage the transfer of data, including spatial extents, on project completion. However, data standards and harmonisation are often only enforced for Protected Sites and other ‘managed’ datasets created, maintained and published by public sector organisations. The relevant data standards are often not promoted to those creating data through primary research. Nor is there a requirement to collate data from separate projects or organisations into datasets, available as view or download services (Figure 6).
Promoting relevant data standards, such as controlled terminologies, during the reporting stage would help improve the quality and consistency of the data and ease the task of aggregating data from multiple contributors. The flow line would ensure that data is collected only once, transferred to the relevant curatorial organisation for use in their systems, and potentially published through an archaeological SDI (Figure 7).

**Figure 6.** Data harmonisation often only applies to public sector data mandated through INSPIRE. There is no requirement to adhere to data standards or create homogenous data from primary fieldwork undertaken by a range of bodies. (© Peter McKeague CC-BY 4.0).

5.1.7. Access and Sensitivity

An archaeological SDI needs to collate and share data between those undertaking primary research and those curating the data for the long term so that it may inform future research. We must acknowledge the symbiotic relationship between the data creator and data curator as equal stakeholders.

**Figure 7.** Developing an archaeological Spatial Data Infrastructure (SDI) would streamline reporting of primary research releasing the value of expensively gathered data. This can be achieved through the Framework and guidance supported by promotion and training (blue) and data aggregation tools (red). (© Peter McKeague CC-BY 4.0).
in building the record. Data should not be hidden behind a paywall but be freely exchanged under open licences permitting access and reuse. Spatial data about archaeologically-sensitive sites, such as investigations of hoard findspots, as well as commercially-sensitive data such as speculative and pre-planning applications, should still be reported, but onward access should be managed through rights management.

Attitudes to sharing data widely vary from country to country and by organisation, but the value of better access to archaeological information is encouraged by the Valletta Convention [7]. Online portals already provide access to site information and raise awareness and foster stewardship of the past whilst recognising the risk of misuse through illegal activities, such as metal detecting.

Archaeological resource managers are also concerned that the complexity of the record and how it was accumulated is open to misinterpretation. Areas lacking investigation are not necessarily absent of the archaeological resource. Unlike formally protected sites, which are either designated or not, the wider archaeological record represents the summation of our current understanding of the record updated through new pieces of research. Imperfect data should not act as a barrier to sharing data, but be used to help promote a better understanding of the archaeological resource to developers and the wider public.

Approaches to data access range from full open access datasets available online and for download to freemium or tiered data models. For instance, the Norwegian Cultural Heritage Database provides tiered access to data from their national inventory [92]. Alternatively, web-GIS portals may restrict the level of detail users can zoom into or the accuracy of the published coordinates. Both approaches are adopted by CSIC—The Spanish Superior Council of Scientific Investigations—in their IDEArq Archaeological Research SDI for Levantine rock art, radiocarbon dates and isotope analysis data [93].

Many public portals only display the location rather than the spatial extent of records. Point data is useful for navigation and generalised distribution maps but does not present a true representation of our knowledge [56]. For example, analysis of the LiDAR DTM in the Federal State of Hesse, Germany, discovered previously unknown burial mounds preserved in the forest. The locations of these prehistoric burial grounds are typically represented by a single marker representing the location of the site (Figure 8A) rather than defining the extent of the site, or marking the position of each mound (Figure 8B).

(A)

Figure 8. Cont.
Figure 8. (A). Burial fields in selected forested areas in the federal state of Hesse in Germany represented as one dot per burial ground (1m resolution LiDAR data courtesy of Hessisches Landesamt für Bodenmanagement und Geoinformationen). (B). Burial mounds in selected forested areas in the federal state of Hesse in Germany represented by one dot per mound. The mounds have been mapped with the help of LiDAR data visualisations. (1m resolution LiDAR data courtesy of Hessisches Landesamt für Bodenmanagement und Geoinformationen).

Subsequent analysis of the data produced different results depending on whether the site location or mound location data was used (Figure 9A,B). The resultant heatmaps reflecting site density show significantly differing distribution centres, resulting in a significantly differing perception of the landscape—and of the meaning of gaps in between the dots.
5.1.8. Skills and Raising Awareness

With an over-reliance on descriptive data and locational approaches to publishing spatial data, the onus is on the developers of an archaeological SDI to raise awareness of its potential and build capacity within the sector to use it effectively. As INRAP have shown [36], investing in people through training in GIS creates a flexible workforce and improves the quality and consistency of data recording.

Improving spatial literacy also requires investing in metadata and paradata so that potential data users can understand how that data was created. As the heatmap visualisations of the density of burial grounds in Hesse (Figure 8A,B) demonstrates, different processes can produce different results. Without the relevant paradata, it is impossible to evaluate the results correctly.

The availability of online data means there is currently a constant need to explain the character of the archaeological record. Online access to primary spatial data will help explain how our knowledge has accumulated and highlight the limitations of our current knowledge.

5.1.9. Recording Systems

As previously discussed, approaches to documenting fieldwork vary from technique to technique and, depending on the organisational structure of field archaeology, country by country. High-level attribution common to all datasets must document what took place where and when, and be supported by more detailed technique-specific attribution. For more technical datasets, such as geophysical surveys and laser scanning [58,59], metadata requirements are already identified for archival purposes, but the simple step of joining the descriptive with the spatial is overlooked.

For invasive techniques such as test pit surveys, watching briefs and excavations, separate metadata is required to document the project extents and trenches (typically who, what and when) from that needed to document each feature or context revealed. Standardising recording practices for excavations delivers a range of benefits organisationally. Aside from developing a flexible, skilled workforce, consistent data standards enable efficient data management and intra- and inter-site analysis. Systems may be standardised on an organisational basis (e.g., INRAP) or use an available software package. Options for the latter include INTRASIS [94], a commercially available software package used...
in much of Scandinavia and by Historic England, or ARK, a bespoke open source, standards-compliant, web-based toolkit for the collection, storage and dissemination of archaeological data [95]. However, many smaller organisations and projects have standalone bespoke recording systems.

5.1.10. Vocabularies

Formalised vocabularies, or thesauri, are now in common usage in archaeological inventories. They provide a structure to the data and encourage a uniform use of vocabularies to improve querying and information retrieval [96]. Originally developed for internal databases, many organisations have now published their key vocabularies online in SKOS Linked Data formats. Examples include a range of vocabularies from Britain, published through Heritagedata.org [97] under the auspices of the Forum for Information Standards in Heritage (FISH) [98], from The Flanders Heritage Agency [99], and from PACTOLS in France [43,44]. The challenge of multilingual thesauri is being addressed through ARIADNE Plus [81], which is mapping resources published in their native language to the hub vocabulary, the Getty Art and Architectural Thesaurus [100], and developing the tools to cross-search multiple datasets in the ARIADNE Portal [47,101].

The application of controlled vocabularies to spatial datasets from primary research is less mature than in online databases. The exciting potential generated by new technologies and the democratisation of data creation has resulted in a wealth of data but little adherence to consistency in interoperability, presentation or reuse. Consequently, there is little control or standardisation over terminologies used, many of which have not been formally defined or are incomplete. Quality is further compromised through the level of interpretation applied or the granularity of the terminology used, with more generic or broader terms competing alongside narrower interpretations, occasionally in the same dataset.

5.2. Societal Benefits

Spatial data is seen as integral to digital transformation processes and growing the digital economy by most governments. Through INSPIRE, a wide range of public sector data is already available through national SDIs, often under an Open Licence. Adoption of the Open Data Charter [89] by national governments encourages greater transparency and accessibility.

The systematic value of data in an SDI delivers both cost savings and benefits to both data curators and users. Although not directly comparable with archaeology, analysis of marine SDIs found the cost-benefits ranged from 1:2 to 1:18. Benefits included: efficiency of data collection; improved risk assessment for navigation; more effective marine spatial planning; supporting marine science; reduced mineral exploration costs; and disaster management [102].

The United Kingdom government established the Geospatial Commission to coordinate and facilitate the generation and use of domestic geospatial data, products and services across the private and public sectors. It aims to accelerate the delivery of economic, social and environmental benefits derived from that data [103]. The Geospatial Commission has already invested in developing an Underground Asset Register to bring together data from the privatised utilities across Britain [104]. In this context, the absence of a coordinated approach to archaeological spatial data should be alarming. If buried infrastructure data is available (if not openly accessible), decisions will be made on the basis of that data. 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 economy. Developing consistent archaeological datasets not only makes managing that data more efficient, but also unlocks economic value through saving time and money.

Archaeological data is routinely used in the stewardship of the historic environment through curated spatial datasets which inform a wide range of planning activities. A recent study of the impact of local authority archaeologists in Britain [105] estimated that archaeology in development management returned £15 for every £1 spent on planning archaeologists, and contributed £218 million to local economies from commercial archaeology. Up to £1.3 billion was saved in delay and emergency
excavation costs. For digital archives, analysis of the value and impact of the ADS in 2013 estimated a two- to eight-fold return through additional reuse of data [106]. While there are no comparative figures for the use of spatial data in archaeological research, easier access to better structured data can deliver efficiencies in data gathering and analysis, leading to an improved understanding of the past.

Archaeology is ideal for engaging the public with the past through citizen science initiatives. In the last twenty years a number of initiatives, including the Portable Antiquities Scheme in England and Wales [107], Portable Antiquities of the Netherlands [108] and Digitale Metaldetektorfund (DIME) in Denmark [109], have engaged hobbyists, including the metal detecting community to report finds through online portals. Development of the European Public Finds Recording Network (EPFRN) [110], aims to encourage trans-national collaboration for finds reporting, including data sharing through ARIADNE.

Studies of the usefulness of archaeological data to non-specialists are scarce. A survey of land developers in Finland and Sweden [111] found that the spatial location of sites was the most useful information. Analysis showed that respondents used both online portals and archaeological consultants to access data. Respondents appreciated having direct access to digital spatial information and having it available in one place.

Several localised initiatives demonstrate the benefits of a systematic approach to collating and publishing spatial archaeological data at a local or regional level. For instance, integration of spatial and documentary information in a GIS created a management tool for the historic city of Santiago de Compostela, Spain. Wider public benefit was realised through an online portal [112] enabling researchers and the wider public to explore the rich content spatially [113]. Similar approaches have been developed to manage the architectural heritage of Sardinia, Italy [114] and the archaeology of Rome [115,116]; the latter incorporating semantic mappings to the CIDOC-CRM and extensions. The challenges and value of a national approach to the ‘homogenization of the Archaeological Cartographic Data on a National Scale’ is also recognised in Italy [117].

Historical and archaeological data is an essential component in interdisciplinary approaches. Past archaeological observations help provide evidence for climate change affecting both archaeological sites and the wider environment. Collaboration demonstrates the relevance of the discipline to wider societal challenges and can attract investment for archaeological research. The potential for archaeology to contribute to the long term monitoring and measurement of coastal change may be illustrated by two projects.

Funded through the European Regional Development Fund Ireland-Wales Programme 2014–2020, Climate Change and Coastal Heritage (CHERISH) project, aims to raise awareness and understanding of the past, present and near-future impacts of climate change, storminess and extreme weather events on the rich cultural heritage of the Irish and Welsh regional seas and coast. The project is developing methodologies and skills through shared experiences and knowledge exchange between the two countries and the wider expert community of geographers and geologists [118].

Historic Environment Scotland uses a range of spatial data to monitor the condition of and risk to properties in its care. A highly accurate digital model of the well-preserved Neolithic settlement of Skara Brae, Orkney was created in 2010 to inform site conservation, interpretation and engagement [119]. 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 [120]. The project is also a case study for Dynamic Coast [121,122] a pan-government partnership developing an evidence base of national coastal change across Scotland.

We need to think beyond the immediacy of project or organisational goals to consider the long-term reuse of that data by archaeologists and non-specialists alike. Effective reuse of archaeological data beyond the discipline requires a consistent approach as to how we capture, describe and present that data. Improved organisation and standardisation not only benefits archaeological research and stewardship of the historic environment but also delivers wider societal benefits. The potential is there—it just needs to be realised.
5.3. Mandate

INSPIRE mandates the release of Protected Sites data as discovery metadata, View and Download services. Through the European Convention on the Protection of the Archaeological Heritage (revised) (1992) (the Valletta Convention) [7], archaeologists have a duty to develop and maintain inventories of archaeological sites and to promote public awareness. Inventory systems have been transformed from paper records to an online database and Web-GIS. The INSPIRE Protected Sites theme makes inventory data (Valletta Convention Article 2i [7]) available as spatial View and Download Services but primary research data—Article 7i of the Convention [7]—remains largely inaccessible.

The potential for using emerging digital technologies to share, connect and provide access to archaeological information is recognised by The European Archaeological Consilium’s (EAC) Amersfoort Agenda [123]. The Amersfoort Agenda acknowledges that this will require improved collaboration and development of, and participation in, European networks. The Agenda also promotes the greatest possible access to digital archaeological resources for various user groups, and the exploitation of digital databases to their full potential, including reaching the wider public. Without dedicated working groups or resources at a European level through the EAC or a relevant Community of Interest in the European Archaeological Association, there will be no concerted impetus to help deliver this vision.

In 2015, the Horizon 2020 Expert Group on Cultural Heritage [124] found that spatial data about the historic environment should be at the heart of good decision making but is noticeable by its absence. The group recognised the need to shift from “an object-orientated approach towards a spatial approach in heritage planning”, and to consider “cultural landscapes early as part of land use and spatial planning processes”. Addressing this need by developing a common framework for creating and sharing archaeological data from multiple sources would create something greater than the sum of its parts.

5.4. Realising the Potential

The authors propose the following vision and mission statements for this endeavour.

**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.

In the absence of legislative requirements, advancing the case for a structured approach to spatial data requires promoting existing initiatives to make the case for routinely capturing spatial data through an archaeological SDI. The SDI would provide the utilities, methods and best practice to boost strategic vision and leadership in the discovery, evaluation and use of archaeological spatial data. This would help all data creators—public bodies, the private sector, academia and the general public.

The following principles, adopted from INSPIRE [125], would underpin the delivery of the vision:

- Data should be collected once and maintained at a level where this can be done most effectively (Infrastructure).
- It should be possible to seamlessly combine data from different sources and share it between many users and applications (Interoperability).
- Data should be openly shared so that users can discover which spatial data is available, evaluate its fitness for purpose and understand the conditions for reuse (Findability).

There is also a need to:

- build capacity within the profession to use and share spatial data;
- ensure that datasets are routinely and promptly updated;
- promote best practice;
- make data discoverable, available as View and Download Services, preferably under an Open Licence. INSPIRE technical specifications already promote good data stewardship, encapsulated by the FAIR Data principles \cite{126,127} for scientific data. In particular, metadata should facilitate interoperability and reuse \cite{128};
- enable users to discover what the dataset is about. Users should also be able to explore the data to ensure that the data is fit for purpose. Finally, users also need to understand how the data can be used through exploitation metadata \cite{128}.

5.4.1. Implementation

Realising the potential of archaeological spatial data should follow the INSPIRE model. It should be viewed as an extension to established data themes and broaden its constituency to include both the research and private sectors.

5.4.2. Infrastructure Goals and Priorities

As with INSPIRE, which sets down the general rules for spatial information across Europe, implementation at a national level needs to reflect, but not be constrained by, the organisational structures, regulations and concerns of each jurisdiction. Archaeologists must agree who is responsible for collating and publishing data from multiple sources. The roles of the record system and the archive must complement each other, with spatial data services built on top of established archiving requirements.

The ability to transfer data easily between different parts of the sector is key to delivering benefits from spatial data. Immediate goals should be to improve information flow from field to record, and promote and adopt agreed thesauri. The long-term ambition is to move beyond the map to develop the Semantic Web capabilities of the spatial data.

We need to agree and prioritise what datasets should be created and, for more technical datasets, define (in conjunction with domain experts) and publish the relevant minimum technical specifications so they do not constrain innovation. Online applications should also be developed or improved to standardise data capture and ease transfer from creator to curator.

5.4.3. Promoting Open Licencing

Unlike INSPIRE, which promotes the sharing of data across the public sector, archaeological data is created through research, commercial archaeology and community projects. With multiple data creators, there is no consistent approach to the reuse of spatial data. Where datasets are available, the licencing conditions are often not apparent, or reuse may be constrained by funding conditions. Open licences must be promoted within the profession to ensure that spatial data, essential for project discovery, is freely accessible.

5.4.4. Findability—The Need for a Spatial Approach

Considerable investment has been made in developing digital infrastructures for object (Europeana \cite{80}) and project (ARIADNE \cite{47}) data. To date, these portals remain object- or site-focused, whereas engaging with the wider geospatial community requires a landscape approach to aid and signpost available View and Download Services to the full range of archaeological spatial datasets, and relevant related resources such as historical mapping and LiDAR datasets. One option would be development of the ARIADNE service page to collate and publish spatial datasets and portals \cite{129}. The wealth of GIS technologies in cultural heritage has long been recognised \cite{130} if not realised. An ongoing review of services \cite{131} demonstrates the considerable potential for coordinating a formal metadata catalogue as part of a wider archaeological digital infrastructure.
6. Conclusions

Spatial data from archaeological fieldwork has considerable untapped potential for research and stewardship of cultural heritage data, yet we continue to work inefficiently through consigning that value to the project report and archive. The challenge of developing a spatial approach to our data is not unique to archaeology. Other disciplines, including geology and the marine sector, have successfully addressed the challenge through collaborative approaches supported by a common infrastructure and vision to deliver efficiencies and benefits to the wider community. Cost–benefit analysis on spatial data and archaeological data generally indicates a two- to 16-fold return on investment. We must do the same; bridging the divide between data creators and data curators to realise the potential of our spatial data. Many of the building blocks that would help archaeologists move from established processes to the maturing geospatial approaches for societal benefit already exist. INSPIRE provides the template and foundations for sharing a range of datasets including Protected Sites data. SEADDA [27] is addressing accessibility and reuse of digital data in archives. Through ARIADNE [47,81], there is an emerging research data infrastructure, and more and more vocabularies are available in SKOS formats. Many countries are developing digital processes and incorporating data standards to manage the data flow from field to record. With standardised reporting, it is a short step to routinely collate and publish spatial data from primary archaeological research. It is a case of bringing together these building blocks but rather than developing new solutions we should build on the emerging digital infrastructure that is ARIADNE. It is time to realise that potential so that: Spatial data from archaeological research is shared openly, enhances the study and stewardship of the past, and positively contributes to the broader geospatial environment.

Author Contributions: Conceptualisation, P.M.; writing—original draft preparation, P.M., A.C., Å.L., A.M., A.P., K.V.D., T.E. writing—review and editing, P.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was financed through an Arts and Humanities Research workshop funding award from the Royal Society of Edinburgh, Scotland, United Kingdom.

Conflicts of Interest: The authors declare no conflict of interest.

References
4. Panek, J.; Netek, R. Collaborative Mapping and Digital Participation: A Tool for Local Empowerment in Developing Countries. *Information* 2019, 10, 255. [CrossRef]
9. Wright, H.; Richards, J.D. Reflections on Collaborative Archaeology and Large-Scale Online Research Infrastructures. J. Field Archaeol. 2018, 43 (suppl. 1), 60–67. [CrossRef]
27. SEADDA. About SEADDA. Available online: https://www.seadda.eu/?page_id=1274 (accessed on 8 April 2020).


31. OASIS. OASIS. Available online: https://oasis.ac.uk/pages/wiki/Main (accessed on 15 February 2020).


36. Larsson, Å.; Smith, M.; Sohlenius, R.; Klafver, T. Digitising the Archaeological Process at the Swedish National Heritage Board: Producing, managing and sharing archaeological information. Internet Archaeol. 2017, 43. [CrossRef]


44. MASA. PACTOLS Thesaurus for Archaeology. Available online: https://masa.hypotheses.org/pactols (accessed on 15 February 2020).


56. McKeague, P.; Corns, A.; Posluschny, A. Why the Historic Environment needs a Spatial Data Infrastructure. *Internet Archaeol.* **2017**, *43*, [CrossRef]
72. Meylemans, E.; De Bie, M.; Petermans, T.; Van Vaekkenborgh, J.; Van Daele, K. The development and application of very high resolution lidar derived maps for the mapping and management of archaeological features and landscapes. (Unpublished; manuscript in preparation).
94. INTRASIS. INTRASIS. Available online: https://www.intrasism.com/ (accessed on 17 February 2020).
Information 2020, 11, 222


101. Richards, J.D.; Niccolucci, F. ARIADNE and ARIADNEplus. In The ARIADNE Impact; Richards, J.D., Niccolucci, F., Eds.; Archaeolingua: Budapest, Hungary, 2019; pp. 7–26. [CrossRef]


111. Huvila, I. Land developers and archaeological information. Open Inf. Sci. 2017, 1, 71–90. [CrossRef]


