POSITION PAPER

Mapping Our Heritage: Towards a Sustainable Future for Digital Spatial Information and Technologies in European Archaeological Heritage Management

Peter McKeague*, Rein van 't Veer†, Isto Huvila‡, Anne Moreau§, Philip Verhagen‖, Loup Bernard¶, Anwen Cooper**, Chris Green†† and Niels van Manen‡‡

The enormous increase in digital spatial information has led archaeologists all over Europe to rely ever more on digital data to prepare and carry out archaeological research, both in academic and heritage management contexts. Spatial information is increasingly used to guide heritage management policies, from urban design to rural planning and tourism. Furthermore, spatial information is more and more employed to involve the general public, using digital technologies in museums and other places of archaeological interest, but also to involve amateur archaeologists in data collection programmes using crowdsourcing. With this proliferation of data and data use, issues of sustainability of digital data repositories, accessibility and reliability of data, standardization of data formats and management of property rights are currently widely debated inside and outside archaeology. However, they have not yet led to generally accepted practices of data management across or even within European countries. In this paper, we sketch the state-of-the-art of archaeological spatial data management, identify the major problem areas and reflect on potential improvements. We conclude that technical solutions are available, but will need a long-term transnational strategy in order fulfill the promise of open and sustainable spatial archaeological data for all user groups involved.

Introduction

The escalation of digital spatial information has led archaeologists all over Europe to increasingly rely on digital data to prepare and carry out archaeological research, both in academia and heritage management. Spatial information, collected in large quantities by archaeologists since the implementation of the Valletta Convention (Council of Europe 1992), is also progressively used to guide heritage management policies, from urban design to rural planning and tourism (e.g. Cuca et al. 2012; McKeague et al. 2012; Huvila 2017). Furthermore, spatial information is more and more used to involve the general public,

* Historic Environment Scotland, GB
† Geodan/Vrije Universiteit Amsterdam, NL
‡ Uppsala University, SE
§ INRAP/UMR6173 CITERES, FR
‖ Vrije Universiteit Amsterdam, NL
¶ UMR 7044 ArchiMedE/University of Strasbourg, FR
** University of Manchester, GB
†† University of Oxford, GB
‡‡ SpinLab/Vrije Universiteit Amsterdam, NL
Corresponding author: Philip Verhagen (j.w.h.p.verhagen@vu.nl)

using digital technologies in museums and other sites of archaeological interest (e.g. González-Tennant 2016; Seitsonen 2017), but also to involve amateur archaeologists in data collection programs, for example using crowdsourcing (e.g. Dhonju et al. 2017; Seitsonen 2017; CReAAH 2019; MOLA 2019; SCAPE 2019; Schweizerische Eidgenossenschaft 2019).

Since the quality of research results and heritage management decisions is highly dependent on the nature of the available data, issues of sustainability of digital data repositories, accessibility and reliability of data, standardization of data formats and management of property rights are currently widely debated. The lack of consistency in (spatial) data standards greatly inhibits the ability to develop sustainable solutions for managing, sharing and analysis of that data. Cross-regional or even supranational analysis of data sets is therefore in most cases a highly time-consuming or even impossible task.

Even when datasets are managed according to well-documented standards, the development of appropriate tools to access and present relevant (spatial) information to researchers, heritage managers and general public is still very much in the stage of exploration and focuses mainly on project-specific contexts that often have a short

lifespan. Software solutions typically have a short lifecycle as well and will cater for different user demands. The flexibility of interfaces is usually limited and does not support interoperability with other datasets. It can therefore be questioned whether these systems really provide the end users with the data and information that they would like to have.

In this position paper, we discuss the existing solutions and state-of-the-art approaches, and their effectiveness in providing the end users with relevant and up-to-date information, based on our experiences in developing data retrieval systems and spatial data infrastructures for various purposes. From there, we will reflect on how to develop more sustainable approaches and technologies for management and use of spatial data, particularly in research and archaeological heritage management, but also within the broader contexts of planning, design and public involvement. It is envisaged that data quality issues will be at the heart of future development work of spatial data infrastructures. However, the success of these infrastructures will also be guided by issues of Open Access and interoperability.

Historical overview
Development of national archaeological databases and their use
The origins of large (national) archaeological databases stem from the post-Second-World War shift in cultural heritage legislation and expansion of the archaeological enterprise in a number of developed countries (Willems 2000; Demoule 2012), leading to systematisation and standardization of archaeological surveys and the compilation of, at first, paper-based registries of archaeological sites (Norman & Sohlenius 2009; Niedziółka 2016). Although early computerised databases date back to the 1960s (Lock 2003), larger-scale deployment of digital information systems began to take place in the late 1980s and early 1990s (e.g. Willems 1997; Kuna 2002; Niedziółka 2016). The need to coordinate and manage archaeological fieldwork and effective decision-making required effective management of information. Maps and spatial data were especially important in this respect, as a core problem of archaeological heritage management and preservation is to collect and maintain accurate information on the location and extents of archaeological sites and monuments and investigated areas.

Even if the national sites and monuments registries were developed primarily as inventories (e.g. Sohlenius 2014) or administrative tools for heritage management (Niedziółka 2016), they also have proved to be useful catalogues of archaeological sites for researchers (Cooper & Green 2016; Niedziółka 2016). However, because of their aggregate nature, they tended to contain only a part of all documented information and their contents were seldom organised to be helpful for answering specific research questions pursued by individual researchers and projects (Meyer et al. 2007; Buckland & Eriksson 2014). In practice this has led to a proliferation of dedicated research project specific databases. Their heterogeneity and the lack

of centralised repositories or portals have made them difficult to access and use (Kintigh 2006; Löwenborg 2014).

Role of GIS and spatial database infrastructures
The large-scale development and use of spatial databases began only after the introduction of GIS software that was reasonably easy to use. Together with affordable personal computers this opened up the possibility to access and build databases and conduct spatial analysis for individual archaeologists and heritage administrators. The emergence of digital spatial data analysis has been one of the most influential factors affecting archaeological work during the past few decades, even when it has been subject to criticism (cf. Rajala 2004; Valenti & Nardini 2004; Conolly & Lake 2006). Apart from the evident usefulness of GIS in archaeological research (see e.g. Lock & Stančič 1995), its potential was also soon noted in archaeological heritage management (Limp 2000). Therefore, the level of standardisation of some spatial datasets is high, particularly in ones developed for regulatory purposes, whereas comparatively little effort has been made to standardize primary evidence from fieldwork and research. This has undoubtedly contributed to the fact that, currently, of all archaeological information, spatial data is probably best represented in digital repositories even if there are many obstacles to managing, preserving and making spatial information accessible.

A central ambition of national archaeological databases was also to make archaeological information more accessible. Many of the archaeological databases and repositories compiled during the 1990s were entirely, or partly, searchable on the Internet or had an ambition to open up for online access at a later date (e.g. Wise & Richards 1999; Kuna 2002). Since the first web-based interfaces, the comprehensiveness of accessible data has improved as well as the general understanding of the uses of, and users of, archaeological spatial data (e.g. McKeague et al. 2012). But the question of the best way to treat legacy data and old online sources remains open.

Creation of digital archives
The digitalisation of the production of archaeological information in general and spatial data in particular from the 1990s onwards meant that the old paper-based practices of archiving archaeological data were rapidly becoming obsolete (Huvila 2016). Even if a large number of digital archaeological databases were developed already during the 1990s with an aim of recording and making available information on archaeological sites to support both research and heritage management (e.g. Roorda & Wiemer 1992), they were generally designed for hosting only a subset of all documentation. As Wise and Richards (1999) underline, with few exceptions (e.g. Eiteljorg II 1995), there was a chronic lack of consideration of long-term preservation of information.

Obsolescence in both software and storage formats are major risks to the sustainability of digital data. In spite of the large-scale acquisition of spatial data, the paper report remained (and in many cases, still remains) the standard deliverable. The digital data either remains undiscoverable

with the data creator (Shaw et al. 2009) or is deleted, hence of no subsequent value beyond the project lifespan. Another problem was, and still is, the proliferation of project and site-specific databases with their own peculiar data structures, concepts and vocabularies (Oikarinen & Kortelainen 2013) that have proved to be difficult to harmonise on a meaningful level. This is a problem common to many disciplines (Bowker & Star 1999). In the UK, the need to interrogate data from multiple projects has been acknowledged and excavation and other fieldwork datasets are produced by commercial archaeology often creating data 'on a per-site basis structured according to differing schema and employing different vocabularies. Consequently, cross search, comparison or other reuse of the data in any meaningful way remains difficult. This hinders the reassessment of the original archaeological findings and reinterpretation in the light of evolving research questions' (Binding et al. 2015). The problem led to the development of the CIDOC-CRM EH extension and a range of semantic tools through the STAR and STELLAR projects (Binding et al. 2015). However, this semantic approach addressed data content rather than the spatial components (geometry) of project archives. Compared to other types of information, spatial data is relatively easy to work with but as a plethora of projects have shown (e.g. Green 2012; Löwenborg 2014), integrating spatial data from different sources and coordinate systems is an arduous undertaking. The increasing availability and use of digital spatial data have been, however, some of the cornerstones and central enablers of the integration of archaeological datasets as a whole (Huggett & Ross 2004).

Even when ensuring that data are accessible in the long term, they are mostly treated on a project by project basis with little or no standardisation across datasets: that is, the mechanisms to ensure the creation, exchange and use of spatial data do not exist for archaeological data within a Spatial Data Infrastructure (SDI), defined as a framework of technologies, policies, and institutional arrangements that together facilitate the creation, exchange, and use of geospatial data and related information resources across an information-sharing community (ESRI 2016).

A lot of significant work for the development of (non-spatial) data infrastructures was initiated in the context of research and development projects around the turn of the century. In terms of developing and maintaining national digital repositories for preserving and providing access to archaeological information, the Archaeology Data Service (ADS) in the UK, established in 1996 (Richards 2002), and the Dutch Data Archiving and Networked Services (DANS), established in 2005 (Gilissen & Hollander 2017), have been pioneers in the field. Even though both services provide excellent models for the preservation of digital data, there is no coordination of the spatial value of that data. Furthermore, the work towards establishing proper digital archives for keeping archaeological data has progressed slowly and many countries are still lacking comprehensive infrastructures for archiving archaeological information, including spatial data. Despite the symbiotic relationship between inventories and the archive, all too often the two

functions are performed by separate institutions. This artificial separation limits the spatial potential of (project) data to inform and enhance the inventory and for the inventory to signpost the archive through connected map layers.

State of the art
International guidelines and standards

The value of geospatial data is recognized internationally through the coordinating efforts of the Open Geospatial Consortium (OGC 2019a) in defining Open Data standards for the global geospatial community. Globally, the value of geospatial data is recognised through the United Nations Global geospatial Information Management (UN-GGIM) strategic framework (UN-GGIM 2018). Many international and national initiatives, including the INSPIRE Directive (European Commission 2019) that aim at realizing the potential of environmental data in Europe acknowledge the value of spatial data for a range of activities including decision making processes. Governments including The Netherlands, through Geonovum (Geonovum 2019) and the United Kingdom, through the Geospatial Commission (Gov.UK 2019) recognise the value of geospatial data to society and economy.

Approaches to the creation, collation and distribution of spatial data in archaeology across Europe are extremely fragmented. Although the Valletta Convention recognises the need for national inventories, to keep them up to date and to facilitate the national and international exchange of scientific information, implementation is left up to the national laws, institutions and approaches of individual signatory states.

Broadly the functions within each state may be defined as

· regulatory – the designation and management of the archaeological resource through formal (legislative) and informal (planning control and management agreements) processes;
· archives and collections – the long-term deposition and preservation of paper, physical and increasingly digital material; and
· investigation and research – undertaken by a range of organisations in the public, private and third sector (community, volunteer and crowdsourcing), each with their own priorities.

Public institutions are focused on delivering their key corporate objectives often with little opportunity for innovation whilst the growing commercial sector is predominantly driven by competitive processes focusing on project delivery rather than contributing to the bigger picture. Financial constraints, the need to demonstrate impact of *your* organisation coupled with institutional inertia in the face of the digital revolution reinforce data silo mentalities leaving the potential offered by digital datasets far from being realised. Moreover, Valletta could not foresee the revolution and opportunities offered by digital data so there are no incentives or drivers to work collaboratively towards developing an SDI for cultural heritage data.

Despite the obvious benefits of frameworks developed to contribute to interoperability such as the CIDOC CRM (Doerr 2003; ICOM CIDOC 2019a) or more pragmatic solutions required to collate and share data locally, there is no true coordination, leadership or mandate either at European or national level from within the profession to coordinate harmonising spatial data across organisations and jurisdictions.

Instead, the key drivers mandating the harmonisation and publication of datasets are external. In particular, the European Union INSPIRE Directive, transformed into the national legislation of member countries, requires that Protected Sites are published as WMS and WFS to agreed technical specifications, to inform Community environmental policies and policies or activities which may have an impact on the environment. As legislative frameworks and working practices differ in detail across Europe, approaches to the publication of 'protected sites' are uneven with some authorities only releasing formally designated data rather than full inventories.

INSPIRE only addresses authoritative datasets but archaeologists routinely create spatial data through a range of increasingly sophisticated digital fieldwork techniques. There is no requirement to collate spatial data from different archaeological projects into a single resource. For example, a map of the archaeological landscape may be created and published by applying consistent data standards to transcriptions of individual archaeological sites (**Figure 1**). This map may be used for internal management purposes, published online as part of a heritage portal (Canmore), and exposed through a metadata catalogue as WMS or WFS for others to access on their own GIS. Yet all too often data is created on a project by project basis following different conventions and formats, greatly restricting the ability to develop maps of the archaeological landscape.

Within Europeana (Europeana Foundation 2019), CARARE (Carare 2019) acts as an aggregator for archaeological and architectural heritage but the emphasis is firmly placed on the cultural value of objects within virtual collections rather than the spatial content. More recently
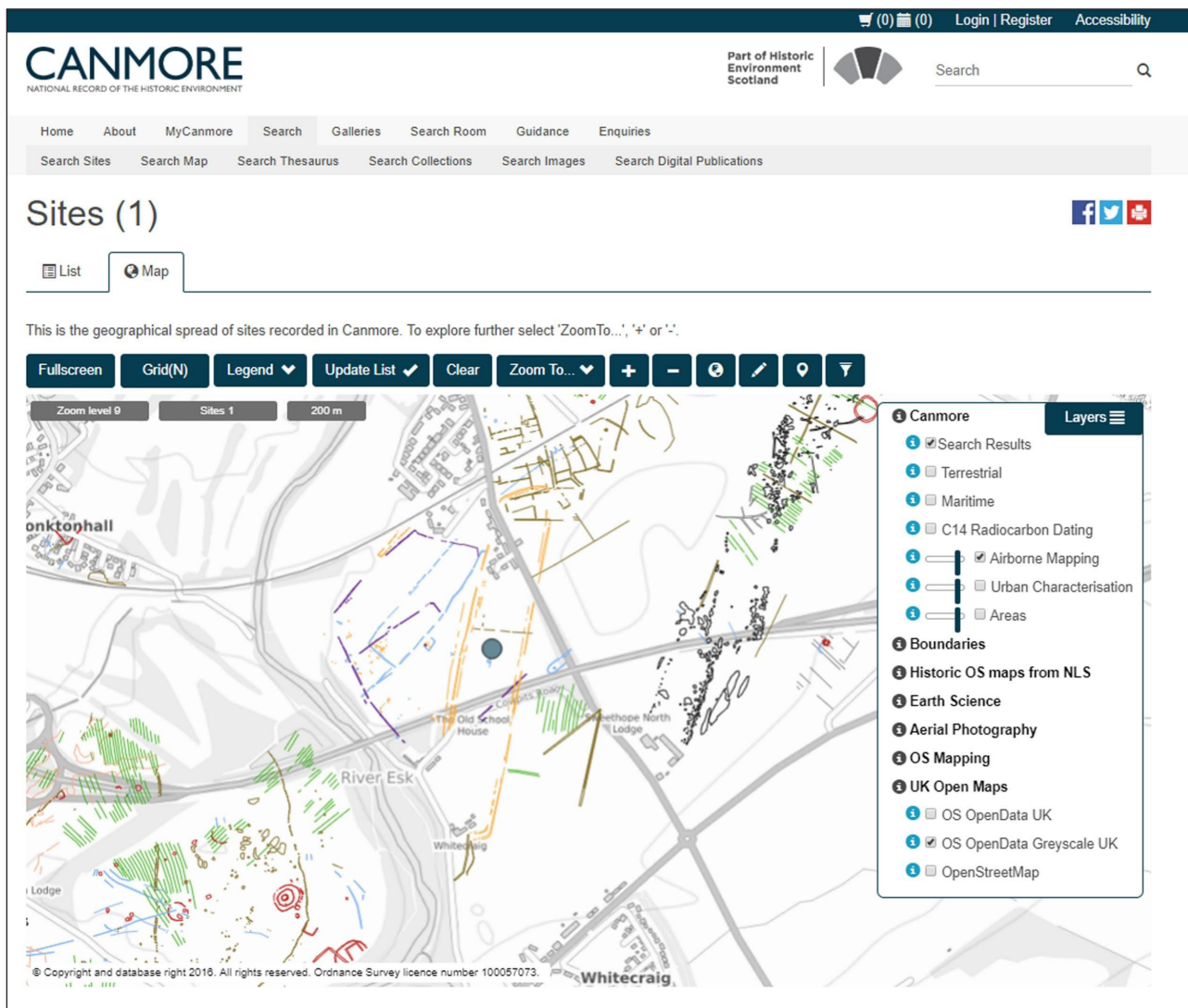
Figure 1: The archaeological landscape from Neolithic, through Roman to industrial at Inveresk, Scotland as revealed by cropmarks and published on Canmore: the online portal of the National Record of the Historic Environment for Scotland.
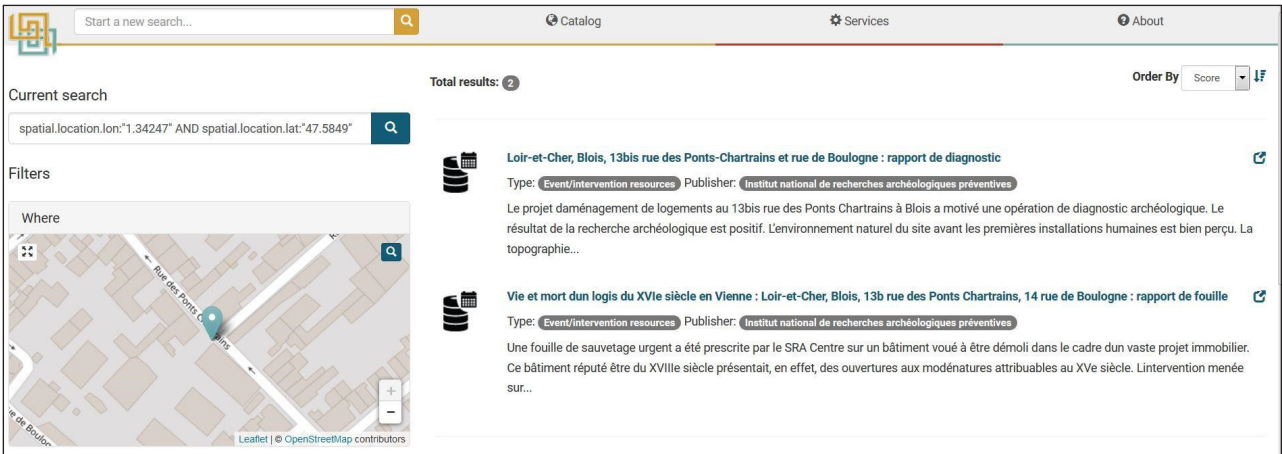
Figure 2: The ARIADNE portal displays the locational details of an archaeological intervention.

the ARIADNE Infrastructure project (ARIADNE 2019a) pooled existing archaeological research data infrastructures through new and powerful technologies to provide a European-wide interoperable dataset. Spatial data is limited to location based searching rather than rendering the spatial footprint of the asset. The portal displays the location of records in the system. **Figure 2** shows the location of an excavation undertaken by INRAP in Blois, France. Users can read a summary of the project on the portal (ARIADNE 2019b) or follow a hyperlink through to the contributing organisation's own resource (INRAP 2019). However, INRAP also maps the spatial footprint of excavations and the features revealed (**Figure 3**). Rich attribution allows

the user to explore the data. This data could be added to the portal as either a WMS or WFS. The limit of that kind of project is that only the INRAP data is available online. This excludes all data from the 20th century, which was available in the former PATRIARCHE system and will soon be accessible through the Pleade portal (PLEADE 2019) from the French Ministry of Culture. Even this platform does not share research and university data, despite the existence of European top-down tools like ARIADNE or Europeana and several bottom-up solutions, like ArkeoGIS (2019) or very local tools like those of the ChasseoLab (2019).

The two examples presented here (**Figures 1–3**) require the consistent application of data standards to combine

Figure 3: Spatial data is more than a place marker: much more spatial data was recorded by INRAP during the fieldwork, including both the project and trench extents as well as the locations of individual features.

data from multiple projects into organised datasets managed as part of core business applications and published online through web-GIS portals. For INSPIRE mandated datasets data is also available as WMS and WFS enabling other organisations, such as environmental consultancies, and researchers to work more efficiently by accessing information directly into their own systems. Although the ARIADNE Portal does not use dynamic WMS and WFS mandated by INSPIRE (in part the attribution is not rich enough), the portal could both signpost and consume those services to deliver truly spatial datasets within a dedicated heritage portal.

Archaeological spatial data tend not to conform to recognised ISO standards. The ISO Standard ISO 19115-2014 (ISO 2014) defines the schema for describing geographical information and associated services, including contents, spatial-temporal extents, data quality, access and rights to use (Shaw et al. 2009: 8). Metadata should be discoverable online so that a user can understand the nature, content and extent of the data. Additionally, there should be enough information to allow the user to explore and assess if the data is fit for purpose. For complex datasets such as remote sensing data, this should include technical metadata documenting the instrumentation and resolution of the capture scale. Finally, exploitation metadata is essential to allow the user to access transfer and apply the data in their own systems both within the heritage sector and for wider environmental benefits. These technical specifications need to be defined and documented, including mapping to the CIDOC-CRM and CRMgeo (Hiebel, Doerr & Eide 2017; ICOM CIDOC 2019b) standards.

Then there is the need to demonstrate the benefits of developing a standardised approach for spatial information in heritage to both archaeologists and those outside the profession. Project-led recording pays little attention to consistent, interoperable data so that it is time consuming to collate data from different projects – even those undertaken by the same organisation. Project driven, fragmented spatial data presents archaeology as amateurish, although this is not a problem unique to archaeology. It is time consuming to find and combine data and to the end user the lack of consistency in how information is displayed can appear unprofessional. Benefits to a consistent approach are both practical and creative, so long as space remains for local specificity in data. Users are able to work more efficiently through information access and consistent data has the potential to offer new insights.

Technologies used
Problem statement
The introduction of spatial technologies has created unprecedented opportunities for data collection, analysis and archiving. However, it could also be argued this has in some ways worsened rather than improved the interoperability of archaeological information. Hitherto, any person knowledgeable in archaeological terminology could access either handwritten or printed materials to inspect the results of archaeological research. The digitization process created new opportunities, but also imposed new hurdles on research data accessibility:

· The new technologies require a significant level of geospatial information (GI) literacy (see Nazari 2011; De Kleijn et al. 2014) that is at odds with the text-oriented skills of humanities scholarship.
· Even if properly curated, the differences in digital strategies used to describe the archaeological interpretations and methodology makes it very hard to combine information from different archaeological projects. This is especially the case for spatial archaeological information. Previously an archaeologist required no more than paper maps to (re-)interpret, now one has to deal with spatial data files in possibly differing formats, coordinate systems and information schemas. The visually interpretable world of archaeology has become much less intuitive. When a standardized format or system is used, differences in format and coordinate system can be easily overcome in a technical sense. Yet, differences in information schemas are not so easily alleviated if one does not have access to the exact semantics of the differences.
· Authorities have been lagging behind in adopting standards that alleviate these tensions in the (re-)use of digital archaeological data. For instance: Dutch archaeology created the SIKB0102 information exchange format (Van 't Veer 2012; Boasson & Visser 2017), but this standard has to date not been enforced by any organisation. So, currently most information models for archaeological data collections are hidden in the database applications where these data collections reside.
· Although GIS systems have contributed a lot to improve working with large spatial databases on a scale where paper maps become useless, the curation of archaeological field data still requires a lot of time and effort.

Seen from the perspective of knowledge and information systems, there have been important developments in the subdomains of knowledge systems known as knowledge representation and reasoning (or KRR) and artificial intelligence. These fields represent the state of the art when it comes to advanced knowledge representation and data analysis strategies and technology. The question therefore is: how can these fields contribute to the improved (re-)use of archaeological data, digital spatial data in particular?

Knowledge representation in archaeology
One of the largest successes in the field of knowledge representation has been in the development of the Semantic Web movement (Berners-Lee et al. 2001). The implementation of semantic web strategies is called Linked Data and its data framework RDF – the Resource Description Framework. Linked data is information broken down in so-called triples, each of which represents an entity-attribute-value statement (Antoniou & van Harmelen 2004: 24–25). Each triple is thus a combination of a subject (entity), a predicate (attribute) and an object (value). While this idea may seem abstract, its example application for archaeology is

very straightforward: the information elements making up a Harris matrix are exactly that: a subject (an archaeological feature) being related through a predicate (e.g. *cuts*, *cut by*, etc.) to another object (another archaeological feature).

In fact, archaeological field data, in the relationships between archaeological contexts and other archaeological contexts, finds and methodological constructs, function very much like a complicated graph structure that is at odds with the relational database structures that we tend to use to describe these entities. Instead, the RDF structure of describing these entities in terms of their relationships to other entities allows for easier deep graph traversal and network analysis. The use of Linked Data technologies, however, is by no means simple. The problem of GI literacy may be reduced in the future by the rise of generations of archaeologists that know no other world than the digital, but the technical expertise required to express archaeological data as Linked Data is considerable.

The second area where semantic web technologies can contribute is in describing the semantics, using web addresses. The prevalent way to embed the semantics of data in the data itself is through Linked Data. RDF is often used as a standard to describe the properties and semantics of data in a way that is much more descriptive than just an ambiguous label that is often used in databases, such as "date". It is easy to imagine confusion over the property "date": is this the dating of some find, the time it was collected in the field, when it was recorded in the field recording system, when it was entered in the database, altered in the database or when it was deposited in an archive? Best practices in RDF require you to model your data in more specific terms than is customary in most current (spatial) database systems, allowing users to look up the exact intended use of an instance just by following the links that describe the types and predicates of the data. Even if two data sets use a different vocabulary to describe the same types of entity, the classes used in these vocabularies can be (locally) equated to or subsumed, allowing the user to query across these data sets using one vocabulary.

The development of CIDOC-CRM extensions CRMgeo and CRMarchaeo (ICOM CIDOC 2019c; Nicolucci 2017) is highly relevant here. After its CRM-EH predecessor (CRM-EH 2019), it may be the first internationally drafted metastandard for describing archaeological data using RDF. In the long term, it could go a long way towards solving problem statement no. 2 above: the lack of interpretable data. Also, if developed further, it could serve to help with problem statement 3: the lack in adoption of shared archaeological data standards. However, it is still in a very early stage, and, as part of the CIDOC-CRM initiative, is not a domain ontology, but a top ontology, intended to be implemented in domain ontologies. As such, it covers archaeological concepts only on a highly abstract level, such as generic archaeological process units. CRMarchaeo differentiates between different archaeological process unit subtypes such as trenches and sections, which is a level of description required to express the meaning of actual data records. As a consequence, at present there is no overarching domain ontology derived from CRMarchaeo that allows easy use and implementation. Furthermore, it suffers from the same issue in exacerbating problem statement no. 1: semantics are complex to express in Linked Data: one not only needs a very high degree of understanding of how archaeological concepts are related, but also considerable Linked Data expertise to create a good semantic model.

Coordinate systems, place name gazetteers and ontologies provide tools that enable the cross-searching and analysis of spatial data but spatio-temporal data, often modelled locally or regionally, presents a particular challenge. The problem has been addressed through PeriodO (2019), cataloguing 'not global period concepts, but specific period definitions: authoritative assertions about the chronological and geographical coverage of period concepts, expressed using machine-readable coordinates (including start and end dates as well as geographic boundaries)' to 'facilitate the discovery of chronologically related data across heterogeneous digital resources' (Rabinowitz et al. 2016: 44–45).

The CIDOC-CRMgeo extension is an ontology extension modeling spatio-temporal aspects of heritage in general. It bridges the gap of cultural heritage modeling with geospatial and temporal properties, opening the possibility for integrating knowledge bases between heritage subdomains. CRMgeo interfaces with a geospatial extension called GeoSPARQL (OGC 2019b), that, in theory, facilitates spatial interoperability.

The GeoSPARQL standard is composed of a vocabulary part and a functional SPARQL extension. The GeoSPARQL vocabulary is highly useful in separating and describing spatial properties (geometries) from their objects (features), where in non-semantic data formats, these separate entities are often confused and concatenated. In GeoSPARQL, multiple geometries can describe the same object, at different functional levels: one can supply a centroid point coordinate for the use of labeling, and a complex multipolygon to describe its spatial extents. The extension links the GeoSPARQL vocabulary to CIDOC-CRM spatial and spatiotemporal concepts. At first sight this is of great value to incorporate cross-domain geospatial standards rather than to re-invent the wheel, but the practical value of both the CIDOC-CRM geospatial and archaeology extensions needs careful evaluation in future work.

The GeoSPARQL *functions* on the other hand offer little advantage over the range of functionality found in common spatial data infrastructures and GIS, except that it can be used in conjunction with Linked Data. For instance, common functions to calculate polygon areas, reproject to different coordinate systems or return the number of vertices in a geometry are lacking in the current version 1.0 (Perry & Herring 2012).

In considering the CRMarchao and CRMgeo extensions, there is still a large void between an abstract set of top concepts in CIDOC-CRM and their practical implementation as implementation documentation and manuals for the extensions appear to be unavailable.

Artificial Intelligence for archaeology
The field of artificial intelligence and, in particular, of machine learning has been diversifying and improving fast. Developed by large technology companies, the

available technologies tend to be highly abstract and hard to use but are rapidly gaining in usability and performance. Machine learning comprises a set of data analysis strategies that rely mostly on principles derived from statistics and linear algebra (Goodfellow et al. 2016).

Archaeological applications of machine learning are still rare and mainly confined to issues of automated classification (Van der Maaten et al. 2007; Hörr, Lindinger & Brunett 2014) and prediction (Oonk & Spijker 2015). In the ARIADNE project, an experimental automatic hypothesis generating setup was tried to see if novel hypothetical viewpoints could be formulated from patterns in the data itself (Wilcke et al. 2017). Other studies are underway, such as in the ArchAIDE project (ArchAIDE 2019), but publications in general on AI and archaeology are still few and far between.

However, it is likely that machine learning may have a tremendous impact on almost every data heavy analysis in archaeology, in support systems for

- classifying, tagging or captioning of digital field or artefact images
- automated extraction of feature geometries from (unmanned) aerial photography and LiDAR (Verschoof-van der Vaart & Lambers 2019)
- classification and description of archaeological features (Van't Veer et al. 2018)
- classifying, dating and matching of artefacts
- error detection in archaeological records
- predictive modelling of site locations and routes.

These goals are getting within reach using the current state of technology in machine learning. With regards to problem statement 4 – the time-consuming curation of archaeological information, and issues relating to the fragmentary nature of data (e.g. Löwenborg 2018) – the machine learning technologies are likely to be able to play a significant role but with so few practical results it is too early to tell.

Challenges
Sustainability
As highlighted earlier, efforts to guarantee the sustainable archiving of digital archaeological data have been partially successful at the local and national level, even when the spatial dimension of archaeological data has not been the primary focus. However, even well-established initiatives are still very vulnerable to changes in the political climate. Almost all organisations curating archaeological data do so on the basis of governmental funding, and changes in political priorities can easily lead to loss of datasets if curating these is not part of the legal obligations of administrative authorities at local, regional, national or transnational levels. The Dutch national digital archive DANS, for example, is currently financed by the Dutch government as a partnership between the Royal Dutch Academy of Sciences (KNAW) and the Netherlands Organisation for Scientific Research (NWO), as part of a larger effort to curate scientific research data. However, there is no guarantee that this agreement will continue in the future, and if so,

whether accessing data will remain free of charge, and if DANS will continue to receive enough funding to adapt its services to changing technological and research environments. And the Netherlands are ahead of many other countries in acknowledging that there is a problem that needs to be addressed at the national level. The funding for the German IANUS initiative (IANUS 2019), for example, was ended in 2017 without a follow-up strategy for implementing the project's results. Clearly, sustainability is not just a matter of developing exchange standards and creating repositories, but also of developing infrastructures that have a sound financial and legal basis.

Open Data and the FAIR Principles
External factors can act as a catalyst for change in archaeological data management. In publication the move to Open Access is gradually changing working practices. The growth of data papers heralds new working practices enabling reuse and critical reassessment of primary data. The vision for developing a sustainable future for spatial data is aligned on the Open Data Charter (2019) and the FAIR principles (Data FAIRport 2019). Releasing data under an Open Data licence ensures that the terms for reusing data are clearly defined. Open Data, therefore, represents an important development in developing collaborative mapping solutions.

Open Data are typically thought of as textual data, but the principles and definitions are equally applicable to spatial data. An additional challenge for some spatial data is third party intellectual property rights and copyright for data sourced against national mapping agency or other vendor products. Going forward, openness needs to be built into data creation to ensure ease of re-use.

Legacy data presents another challenge. Archive datasets, from paper records to most born digitally, were created without thought of transformation or re-use in a spatial data infrastructure. Traditional approaches persist with most spatial data consigned to an illustration in a project report, fossilising in print the spatial knowledge often acquired digitally in the field. Resolving copyright issues aside, much of this information cannot be easily or accurately geo-referenced and the sheer time and effort required to capture legacy data appears daunting. However, a pragmatic approach might be to capture the data 'as required' to support other project work or through training programmes.

Counterarguments to opening data are familiar. Publishing site locations raises the risk of looting through metal detecting and other activities. There will always be a hard core of determined looters, but risks can be mitigated through education and cooperation. At the same time promoting heritage amongst local communities often raises awareness in their surroundings empowering a sense of stewardship. A middle ground between those interests is often found by permitting access to data at various levels of detail for various user groups. The Portable Antiquities of the Netherlands project (PAN 2019), for example, provides searching and mapping of (metal) items at the municipality level, but will not provide exact coordinates unless the user has a login to the system. The

Portable Antiquities Scheme in England (British Museum & Amgueddfa Cymru 2019) will provide the spatial information to the nearest 1 km square.

Open Data elsewhere also challenges perceptions of data ownership. Archaeologists are often reluctant to share data, arguing that in so doing they are giving away their research. Yet their research is built upon data often gathered at public expense. Under Open Data elsewhere, early release of the data can help inform research and challenge conclusions. The rise of data papers, providing citations, helps demonstrate the changing thinking towards data in general but perhaps more is required to acknowledge the work of the original researcher.

Some agencies have developed business models towards accessing data where a limited amount of data is freely available with the rest behind a paywall. The impact of Open Data elsewhere within government agencies on free-mium data models is yet to be fully explored. Approaches to opening data have to be tailored to the specific conditions of each country, but archaeologists also need to recognise the relationship their data has with the rest of the geospatial world. Through INSPIRE some data has been released to enable and inform others to make decisions that affect the historic environment. INSPIRE raises the expectation that public data is findable and reusable. Development of web portals accessing geospatial web services, and complex modelling tools like ecosystem services expect to find and use data. If archaeologists do not participate in these systems, there is a very real danger that their data will simply be ignored.

Data quality management
Despite the fact that standards for data description and exchange are now well developed and increasingly being accepted within the archaeological community, there is still a lack of common concepts to be used for the evaluation of data quality. The most common quality aspects of archaeological data can be related to four main types of uncertainty: spatial precision, chronological certainty, interpretational accuracy and completeness of documentation (Verhagen et al. 2016). However, methods to define levels of certainty and to compare datasets with different accuracies have hardly been a subject of debate within the archaeological community. For example, the practice of 'fuzzy' dating of archaeological objects and assemblages (see e.g. Green 2011; Crema 2012) does not fully solve the problem of understanding chronological uncertainties, since it is inevitably based on a subjective assessment of the accuracy of dating, unless absolute dating methods are available. To rely on absolute chronology, however, seems essential, even when lots of dates are fuzzy. Different online gazetteers (e.g. PeriodO, Pleiades) and ontologies (e.g. OWL-Time) appear to offer good intermediate solutions.

From a practical point of view, data quality assessment should start from the definition of an ideal dataset and listing its characteristics. For example, an 'ideal' field survey dataset should be based on dGPS-based coordinates, should follow a field walking strategy that is according to current scientific standards, and employ a standardized and accepted system for recording and describing finds. Already, such a seemingly simple data quality description will be quite complex to define, let alone 'mapping' existing field survey datasets onto it and evaluating data quality on this basis.

Towards realising the value of spatial data in archaeology
Multiple agencies/actors create and use spatial data in archaeology, but the potential of that data remains unrealized despite initiatives in the wider geospatial community. This paper has outlined existing approaches/initiatives to delivering value from spatial data. The strengths and weaknesses of the different approaches are summarized in **Table 1**.

The two key approaches outlined in this paper address different aspects of realizing the spatial value from archaeological data. Development of consistent spatial data standards, following the INSPIRE model, requires effective data management through the definition and implementation of appropriate data standards and technical specifications across a range of archaeological spatial datasets. This model works well where there is a requirement to publish defined datasets, including spatial extents but does not address the richness and variety of the archaeological data behind the mapped content.

Semantic approaches, based on conceptual reference models such as the CIDOC-CRM and extensions, address interoperability by defining an integrating framework for manipulating data from multiple projects. These approaches underpin research projects, including Pelagios (Pelagios Commons 2019) and ARIADNE, that aggregate data from multiple sources. Although very powerful, the voluntary nature of participation and incompleteness of the data means that data cannot be relied on for decision making purposes. Current initiatives downplay the value of the mapping element to location – the simple X and Y coordinates rather than the full geometries of the archaeology not required for their projects.

Much more effort is required to realise the potential of (spatial) data created through fieldwork and research. There are no agreed standards or specifications for documenting fieldwork with organisations and even separate projects within organisations each defining their own technical specifications. Too often data from fieldwork is seen as a means to an end – the production of a report for a client – without consideration of the bigger picture. Addressing this issue requires a dialogue between those who create and those who curate data to define and share a common approach to maximizing the value from spatial data from fieldwork.

Future agenda
The themes approached in this paper show how much research time, expertise and imagination has already been invested at local, regional and international scales. It is now time to move towards larger scales through sharing knowledge and expertise within a structured framework, if possible by keeping the habits of the end users and the positive aspects of as many projects as possible.

Table 1: Suggested strengths and weaknesses of different approaches to realising the value of spatial data from archaeological datasets.

	Infrastructure	Conceptual Reference Model	Data reusers	Data creators
Initiatives	INSPIRE	CIDOC-CRM and extensions	Research projects	Archaeological practitioners
Key driver	Legal requirement	Standards Research	Research	Business requirements
Engagement	Compulsory for mandated public sector datasets	Research led	Research led	Core part of business delivery
Strengths	Part of a robust infrastructure for sharing environmental data across Europe	Accepted international standard	Well-funded innovative development phase	Business efficiency
	Implementing organisations capable of sustaining data delivery through core funding	Provides an integrating framework for diverse data structures	Project partners represent a coalition of the willing	Can define own specifications
	Data discovery and delivery provides (relatively) easy access to, and reuse of, data		Aggregates data from multiple sources	
Weaknesses	Restricted to 'Protected Sites' data	High technological threshold to implement	Spatial component generally restricted to locational data only	Fragmented approach – project focused – multiplicity of recording standards
	Does not address data created through archaeological fieldwork and research	Skills deficiency in wider archaeological community	Partners predominantly research focused and coalitions of the willing	Long term value of data beyond project lifespan not recognised
	Only applies to public sector data	Low engagement from public sector	Dependent on research funding for sustainability	
			Tailored Conceptual Reference Models	
			Value rests with the data aggregator	
			Incomplete datasets are not suitable for decision making	
			No engagement with data creators	
Sustainability	Legislative requirement for mandated datasets	Core standard maintained and developed by specialist community	Project life cycle	Reliant on an appropriate archive to host the data

Quantitative indicators of data quality

Large datasets inherently have their imperfections, due to the effects of combining various sources with different standards of data entry and data management practices (see e.g. Cooper & Green 2016). However, the fear of criticism of less than perfect data hinders some individuals and institutions from sharing their data. It is therefore important that tools for assessing data quality will not be used to shame and blame the original contributors but will act as an incentive to work towards improving data quality. After all, it is up to the future users of legacy datasets to decide whether the data is good enough for purpose, and if it is not, to develop strategies to enhance data quality. The good news is that good curation tools for digital data, such as graphs, are increasingly available, but we also need to find to ways to pay and credit the curators.

Open Data

Ideally, using the FAIR principles and sharing both raw data and expertise can only be positive for research. The challenge will be to manage for digital data curators adopting Open Data and the FAIR Data principles as part of their business model, without compromising the ability to attract funding in a competitive market. Acknowledgement of the value of data through FAIR Data principles is an important step in changing attitudes of funders. Moreover, Open and FAIR Data aligns with the Valletta Convention's requirements to exchange scientific data.

A complex issue is the difference on policies and strategies towards Open Data in various parts of the world. Even inside the EU, and despite transnational protocols, the approaches vary enormously, from moving towards completely Open Access for public data to setting barriers for use by asking handling fees or even restricting access because of real or perceived privacy, copyright, ownership and security issues. The difference between federal and central states is another issue here. This is not something that the archaeological community can solve on its own, but it will be necessary to signal through international archaeological organisations, such as EAA, WAC or EAC, that archaeologists support and are committed to data sharing. This is also necessary to increase support for our work among non-archaeologists, and to maintain transparency about what we are doing as practitioners with data that in the end are paid for with public money.

Sustainability

Many old databases are stored in obsolete (or soon-to-be outdated) software formats, and only archived on media like CDs, zip-disks or even floppy disks. It is time to take action now to make sure that old data sets will not be lost forever within a few years' time. In order to guarantee sustainability of legacy data it is first of all important to have a good overview of what is there, but what isn't shared yet. Data owners need to join forces to investigate the size of the problem, and to develop strategies for dealing with it.

The current shared data initiatives are the frontrunners, and we need to make sure that the practices developed there will be followed by data owners around the world in academia, government and private institutions. Publishing data papers and documenting good practices seem to the most effective way to achieve this, since top-down solutions tend to be very vulnerable to the financing models applied and are often not distributed more widely because of institutional barriers.

Involving the archaeological community

Involving the wider archaeological community in working with linked data and machine learning requires showing that the advantages far outweigh the investment in understanding and applying the technologies involved. This proof should come from a careful comparative evaluation of these technologies in controlled experiments on a wide range of archaeological information types and archaeological projects, in a study that not only shows that the semantic models can correctly describe information, but that there is also a distinct advantage to this when compared to a baseline of common digital methods of representing archaeological information. Training can be organized through international initiatives, and should involve collaboration with computer scientists, as undertaken during the ARIADNE project. An organization like Computer Applications & Quantitative Methods in Archaeology (CAA 2019) could potentially play an important role in this.

Involving data owners and curators

Existing approaches to data collection and curation are ill-suited to the digital age. Compartmentalised roles of data creator and curator and archivist hinder maximising the potential of born digital data in general and the need to collate spatial data from multiple sources is recognised. The bigger map of the archaeological landscape has been lost to the immediacy of project-led archaeology lacking consistent data standards. Harnessing the potential of spatial data requires a collaborative approach through developing networks across the profession to help exploit the full potential of digital databases so that the data can be shared with other policy domains and linked to societal challenges – all sub-themes within the Amersfoort Agenda (Europae Archaeologiae Consilium 2015: 15–23). Doing so requires redesigning existing approaches to make them work more effectively. Digital data demands that it is collected once, combined seamlessly from different providers, maintained and published at the level where this can be done most effectively so that the data is findable, accessible, interoperable and re-usable.

The need for governance

The systemic failure to move towards a landscape approach for archaeological data was highlighted in a 2015 report by the Horizon 2020 Expert Group on Cultural Heritage (European Commission Directorate-General for Research and Innovation 2015) which found that spatial data about the historic environment should be at the heart of good decision making but is noticeable by its absence. The report identified the need to shift from 'an object-oriented approach towards a spatial approach in heritage planning' and to 'consider cultural landscapes early as part of land use and spatial planning processes' to get cultural heritage data to work for'.

Successful delivery of such a vision requires breaking down data silos to develop a sustainable infrastructure collating, maintaining and distributing data effectively. Authoritative datasets, such as designations, tend to follow well-defined schemas within national jurisdictions but are not easily interoperable across boundaries. Primary data collection from fieldwork is generally not standardised and requires harmonising across contributors. This raises the question of whose role is it to pull together these data which may eventually be deposited in a monument record or an archive. Organisational structures of archives adopt a hierarchical approach to tidying specific projects into folders within specific collections and are not open to re-engineering and amalgamating data into datasets let alone publishing that data online through a range of web services.

Delivery requires co-ordination through a thematic Spatial Data Infrastructure setting out the framework, policies, standards etc. to deliver the value from the wealth of spatial data in archaeology (McKeague et al. 2017). Yet that coordination across organisations and territories is conspicuous by its absence.

Conclusion

As this paper shows, the potential value of spatial data in heritage is currently not being realised. Beyond the fundamental issue of ensuring the long-term preservation of digital data in general, there is a lack of recognition of the value and potential of spatial data held in reports and datasets, notwithstand-

ing the obvious benefit of standardising and sharing spatial data for research and to inform environmental policies and activities that may impact the cultural heritage. The technical solutions exist from the Discovery Metadata, Web Map Services and Web Feature Services mandated by the INSPIRE Directive or through Linked Data approaches. However, solutions require capital investment in delivering change and a commitment to maintain services well into the future. With Linked Data, technical barriers constrain their implementation outside research institutes.

The lack of coordination in creating a sustainable future for spatial data – actual or legacy – is a major challenge for archaeology information management. The data-sharing envisaged by the Valletta convention remains limited mostly due to fear of looting, intellectual property issues and client confidentiality. Contrast the lack of coordination in archaeology with the approach adopted by geological sciences where the need for, and value of, a standardized approach to spatial data was recognized and addressed from the early 2000s (Jackson 2007). Now over ten years old, the One Geology portal (One Geology 2019) provides access to standardized spatial data from 113 countries. Multi-national data sharing initiatives, including Pelagios, ARIADNE and ArkeoGIS, are very much the exception in archaeology.

The INSPIRE Directive should form a blueprint for developing an SDI for cultural heritage datasets to demonstrate the need for and potential of a coordinated approach not only for archaeologists but the wider society. This will inform the evidence base to develop a manifesto for implementing a spatial data strategy for archaeological datasets. One channel for preparing a manifesto and strategy would be establishing a working group on spatial data under the auspices of the EAC, ideally jointly with the EAA and Computer Applications and Quantitative Methods in Archaeology.

Acknowledgements

Competing Interests

The authors have no competing interests to declare.

References

Antoniou, G and **van Harmelen, F.** 2004. *A semantic web primer*. Cambridge, MA: MIT press.

ArchAIDE. 2019. Archeo revolution is coming. Available at http://www.archaide.eu/ [Last accessed 20 April 2019].

ARIADNE. 2019a. ARIADNE Portal. Available at http://www.ariadne-infrastructure.eu/Portal [Last accessed 20 April 2019].

ARIADNE. 2019b. Loir-et-Cher, Blois, 13bis rue des Ponts-Chartrains and rue de Boulogne: diagnostic report. Available at http://ariadne2.isti.cnr.it/index.php/page/25109297 [Last accessed 20 April 2019].

ArkeoGIS. 2019. ArkeoGIS: sharing & quering of spacialized data about the past. Available at http://arkeogis.org/en [Last accessed 20 April 2019].

ARKWORK. 2019. About Arkwork. Available at https://www.arkwork.eu/ [Last accessed 20 April 2019].

Berners-Lee, T, Hendler, J and **Lassila, O.** 2001. The Semantic Web. *Scientific American* 284(5): 34–43. DOI: https://doi.org/10.1038/scientificamerican0501-34

Binding, C, Charno, M, Jeffrey, S, May, K and **Tudhope, D.** 2015. Template Based Semantic Integration: From Legacy Archaeological Datasets to Linked Data. *International Journal on Semantic Web and Information Systems* 11(1): 1–29. DOI: https://doi.org/10.4018/IJSWIS.2015010101

Boasson, W and **Visser, RM.** 2017. SIKB0102: Synchronizing Excavation Data for Preservation and Re-Use. *Studies in Digital Heritage* 1(2): 206–224. DOI: https://doi.org/10.14434/sdh.v1i2.23262

Bowker, GC and **Star, SL.** 1999. *Sorting Things out: Classification and Its Consequences*. Cambridge, MA: MIT Press.

British Museum and **Amgueddfa Cymru − National Museum Wales.** 2019. Portable Antiquities Scheme. Available at https://finds.org.uk/ [Last accessed 20 April 2019].

Buckland, PI and **Eriksson, EJ.** 2014. Strategic Environmental Archaeology Database (SEAD). In: Smith, C (Ed) *Encyclopedia of Global Archaeology*. New York: Springer, pp. 7076–7085. DOI: https://doi.org/10.1007/978-1-4419-0465-2_833

CAA. 2019. About CAA. Available at https://caa-international.org/about/ [Last accessed 20 April 2019].

Canmore. 2019. Canmore National Record of the Historic Environment Available at https://canmore.org.uk/ [Last accessed 17 May 2019].

CARARE. 2019. CARARE association. Available at http://www.carare.eu/about/ [Last accessed 20 April 2019].

ChasseoLab. 2019. ChasseoLab. Available at http://chasseolab.huma-num.fr/ [Last accessed 20 April 2019].

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

Cooper, A and **Green, C.** 2016. Embracing the Complexities of 'Big Data' in Archaeology: The Case of the English Landscape and Identities Project. *Journal of Archaeological Method and Theory* 23(1): 271–304. DOI: https://doi.org/10.1007/s10816-015-9240-4

Council of Europe. 1992. *European Convention on the Protection of the Archaeological Heritage (Revised). European Treaty Series 143*. Available at: https://www.

coe.int/en/web/conventions/full-list/-/conventions/treaty/143 [Last accessed 31 March 2019].

CReAAH. 2019. Archaeology, Littoral and Terrestrial Heating ALERT. Available at https://alert-archeo.org/ [Last accessed 20 April 2019].

Crema, ER. 2012. Modelling Temporal Uncertainty in Archaeological Analysis. *Journal of Archaeological Method and Theory* 19(3): 440–61. DOI: https://doi.org/10.1007/s10816-011-9122-3.

CRM-EH. 2019. CRM-EH – English Heritage extensions to CIDOC CRM for the archaeological domain. Available at http://purl.org/crmeh [Last accessed 20 April 2019].

Cuca, B, Oreni, D and **Brumana, R.** 2012. Digital Cartographic Heritage in Service to the Society: Landscape Analysis for Informed Decision Making. In *18th International Conference on Virtual Systems and Multimedia, VSMM 2012, Milan, Italy, September 2-5, 2012*, pp. 499–506. IEEE. DOI: https://doi.org/10.1109/VSMM.2012.6365964

Data FAIRport. 2019. Find Access Interoperate Re-use Data. Available at http://www.datafairport.org/ [Last accessed 20 April 2019].

Demoule, J-P. 2012. Rescue Archaeology: A European View. *Annual Review of Anthropology* 41: 611–626. DOI: https://doi.org/10.1146/annurev-anthro-092611-145854

Dhonju, HK, Xiao, W, Shakya, B, Mills, JP and **Sarhosis, V.** 2017. Documentation of Heritage Structures through Geo-Crowdsourcing and Web Mapping. In *ISPRS Geospatial Week 2017, 18–22 September 2017, Wuhan, China.* The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-2/W7, pp. 17–21. ISPRS. DOI: https://doi.org/10.5194/isprs-archives-XLII-2-W7-17-20

Directorate-General for Research and Innovation. 2015. *Getting Cultural Heritage to Work for Europe: Report of the Horizon 2020 Expert Group on Cultural Heritage.* Luxembourg: Publications Office of the European Union. Available at http://openarchive.icomos.org/1693/ [Last accessed 31 March 2019].

Doerr, M. 2003. The CIDOC Conceptual Reference Module: An Ontological Approach to Semantic Interoperability of Metadata. *AI Magazine* 24(3): 75–92. DOI: https://doi.org/10.1609/aimag.v24i3.1720

EAA. 2019. 271. Mapping our heritage. towards a sustainable future for digital spatial information and technologies in archaeological heritage management. Available at https://www.e-a-a.org/EAA/Conferences/Filmed_conference_sessions/Maastricht_2017/Session_271/EAA/Navigation_conferences/2017_videos/Session_271.aspx [Last accessed 20 April 2019].

Eiteljorg, H, II. 1995. The Archaeological Data Archive Project. In Huggett, J and Ryan, N (Eds.) *CAA94. Computer Applications and Quantitative Methods in Archaeology 1994.* BAR International Series 600. Oxford: Tempus Reparatum, pp. 245–248.

ESRI. 2016. Spatial Data Infrastructure: A Collaborative Network. Redlands (CA): ESRI. Available at: http://www.esri.com/library/brochures/pdfs/spatial-data-infrastructure.pdf [Last accessed 31 March 2019].

Europae Archaeologia Consilium. 2015. Amersfoort Agenda – Setting the Agenda for the Future of Archaeological Heritage Management in Europe. In: Schut, PAC, Scharff, D and de Wit, L (Eds.) *Setting the Agenda: Giving New Meaning to the European Archaeological Heritage. Proceedings of the International Conference Amersfoort, The Netherlands, 20–22 March 2014.* EAC Occasional Paper 10. Budapest: Archaeolingua, pp. 15–23.

Europeana Foundation. 2019. Europeana Collections. Available at https://www.europeana.eu [Last accessed 20 April 2019].

European Commission. 2019. INSPIRE Knowledge Base. Infrastructure for spatial information in Europe. Available at https://inspire.ec.europa.eu/ [Last accessed 20 April 2019].

Geonovum. 2019. De verbindende kracht van standaardisatie Vindbare, toegankelijke, uitwisselbare geo-informatie. Available at https://www.geonovum.nl/ [Last accessed 20 April 2019].

Gilissen, V and **Hollander, H.** 2017. Archiving the Past While Keeping up with the Times. *Studies in Digital Heritage* 1(2): 194–205. DOI: https://doi.org/10.14434/sdh.v1i2.23238

González-Tennant, E. 2016. Recent Directions and Future Developments in Geographic Information Systems for Historical Archaeology. *Historical Archaeology* 50(3): 24–49. DOI: https://doi.org/10.1007/BF03377332

Goodfellow, I, Bengio, Y and **Courville, A.** 2016. *Deep Learning.* Cambridge, MA: MIT Press.

Gov.UK. 2019. Geospatial Commission. Available at https://www.gov.uk/government/organisations/geospatial-commission [Last accessed 20 April 2019].

Green, C. 2011. *Winding Dali's Clock: The Construction of a Fuzzy Temporal-GIS for Archaeology.* BAR International Series 2234. Oxford: Archaeopress.

Green, C. 2012. Archaeology in Broad Strokes: Collating Data for England from 1500 BC to AD 1086. In Earl, G, Sly, T, Chrysanthi, A, Murrieta-Flores, P, Papadopoulos, C, Romanowska, I and Wheatley, D (Eds.) *Archaeology in the Digital Era: Papers from the 40th Annual Conference of Computer Applications and Quantitative Methods in Archaeology (CAA), Southampton, 26–29 March 2012.* Amsterdam: Amsterdam University Press, pp. 307–312. DOI: https://doi.org/10.1515/9789048519590-033

Hiebel, G, Doerr, M and **Eide, Ø.** 2017. CRMgeo: A Spatiotemporal Extension of CIDOC-CRM. *International Journal on Digital Libraries* 18(4): 271–279. DOI: https://doi.org/10.1007/s00799-016-0192-4

Hörr, C, Lindinger, E and **Brunnett, G.** 2014. Machine Learning Based Typology Development in Archaeology. *Journal on Computing and Cultural Heritage* 7(1): 2: 1–2: 23. DOI: https://doi.org/10.1145/2533988

Huggett, J and **Ross, S.** 2004. Archaeological Informatics: Beyond Technology. *Internet Archaeology* 15. DOI: https://doi.org/10.11141/ia.15.13

Huvila, I. 2016. Awkwardness of Becoming a Boundary Object: Mangle and Materialities of Reports,

Documentation Data and the Archaeological Work. *The Information Society* 32(4): 280–297. DOI: https://doi.org/10.1080/01972243.2016.1177763

Huvila, I. 2017. Land Developers and Archaeological Information. *Open Information Science* 1(1):71–90. DOI: https://doi.org/10.1515/opis-2017-0006

IANUS. 2019. IANUS – Forschungsdatenzentrum Archäologie & Altertumswissenschaften. Available at https://www.ianus-fdz.de/ [Last accessed 20 April 2019].

ICOM CIDOC. 2019a. CIDOC CRM Conceptual Reference Model. Available at http://www.cidoc-crm.org/ [Last accessed 20 April 2019].

ICOM CIDOC. 2019b. CRMgeo Spatiotemporal model. Available at http://www.cidoc-crm.org/crmgeo/ [Last accessed 20 April 2019].

ICOM CIDOC. 2019c. CRMarchaeo Excavation model. Available at http://www.cidoc-crm.org/crmarchaeo/ [Last accessed 20 April 2019].

INRAP. 2019. Dolia. Catalog of documentary collections and digital library of Inrap. Available at http://dolia.inrap.fr/flora/ark:/12345/0131064 [Last accessed 20 April 2019].

Jackson, I. 2007. OneGeology – Making geological map data for the earth accessible. *Episodes* 30(1): 60–61.

Kintigh, K. 2006. The Promise and Challenge of Archaeological Data Integration. *American Antiquity* 71(3): 567–78. DOI: https://doi.org/10.1017/S0002731600039810

Kleijn, M, de van Manen, N, Kolen, J and **Scholten, H.** 2014. Towards a User-Centric SDI Framework for Historical and Heritage European Landscape Research. *IJSDIR* 9:1–35. DOI: https://doi.org/10.2902/1725-0463.2014.09.art1

Kuna, M. 2002. The Archaeological Record of Bohemia: An Attempt at an Analytical Information System. In: García Sanjuán, L and Wheatley, DW (Eds.) *Mapping the Future of the Past: Managing the Spatial Dimension of the European Archaeological Resource*. Sevilla: Universidad De Sevilla, pp. 45–52.

Limp, WF. 2000. Geographic Information Systems in Historic Preservation. In: Williamson, RA and Nickens, PR (Eds.) *Science and Technology in Historic Preservation*. Boston: Springer, pp. 231–247. DOI: https://doi.org/10.1007/978-1-4615-4145-5_10.

Lock, G. 2003. *Using Computers in Archaeology: Towards Virtual Pasts*. London: Routledge. DOI: https://doi.org/10.4324/9780203451076

Lock, G and **Stančič, Z.** (Eds.) 1995. GIS and Archaeology: A European Perspective. London: Taylor and Francis.

Löwenborg, D. 2014. Recuperating GIS Data from Excavations: On the Use, or Lack of Use, of Digital Archaeological Information. In: Huvila, I (Ed.) *Perspectives to Archaeological Information in the Digital Society*. Uppsala: Department of ALM, Uppsala University, pp. 11–23.

Löwenborg, D. 2018. Knowledge production with data from archaeological excavations. In: Huvila, I (Ed.) *Archaeology and Archaeological Information in the Digital Society*. London: Routledge, pp. 37–52. DOI: https://doi.org/10.4324/9781315225272

Maaten, L, van der Boon, B, Lange, G, Paijmans, H and **Postma, E.** 2007. Computer Vision and Machine Learning for Archaeology. In: Clark, JT and Hagemeister, EM (Eds.) *Digital Discovery. Exploring New Frontiers in Human Heritage. CAA2006. Computer Applications and Quantitative Methods in Archaeology. Proceedings of the 34th Conference, Fargo, United States, April 2006*. Budapest: Archaeolingua, CD-ROM 476-482.

McKeague, P, Corns, A and **Posluschny, A.** 2017. Why the Historic Environment Needs a Spatial Data Infrastructure. *Internet Archaeology* 43(7). DOI: https://doi.org/10.11141/ia.43.7

McKeague, P, Corns, A and **Shaw, R.** 2012. Developing a Spatial Data Infrastructure for Archaeological and Built Heritage. *International Journal of Spatial Data Infrastructures Research* 7: 38–65.

Meyer, É, Grussenmeyer, P, Perrin, J-P, Durand, A and **Drap, P.** 2007. A Web Information System for the Management and the Dissemination of Cultural Heritage Data. *Journal of Cultural Heritage* 8(4): 396–411. DOI: https://doi.org/10.1016/j.culher.2007.07.003.

MOLA. 2019. Citizan. Available at https://citizan.org.uk/ [Last accessed 20 April 2019].

Nazari, M. 2011. A Contextual Model of Information Literacy. *Journal of Information Science* 37(4): 345–359. DOI:https://doi.org/10.1177/0165551511403544

Nicolucci, F. 2017. Documenting Archaeological Science with CIDOC CRM. *International Journal on Digital Libraries* 18(3): 223–231. DOI: https://doi.org/10.1007/s00799-016-0199-x

Niedziółka, K. 2016. The Prospect of Digitization of Polish Archaeological Record on an Example of Materials from the Turn of Bronze and Iron Age from the Area of Pomeranian Voivodeship (Northern Poland). *Sprawozdania Archeologiczne* 68: 121–44. DOI: https://doi.org/10.23858/SA68.2016.007

Norman, P and **Sohlenius, R.** 2009. ASIS – More than a Register of Ancient Monuments. In: Schut, PAC (Ed.) *Listing Archaeological Sites, Protecting the Historical Landscape*. EAC Occasional Paper no 3. Brussels: EAC, pp. 83–86.

OGC. 2019a. Welcome to The Open Geospatial Consortium. Available at http://www.opengeospatial.org/ [Last accessed 20 April 2019].

OGC. 2019b. GeoSPARQL – A Geographic Query Language for RDF Data. Available at http://www.opengeospatial.org/standards/geosparql [Last accessed 20 April 2019].

Oikarinen, T and **Kortelainen, T.** 2013. Challenges of Diversity, Consistency, and Globality in Indexing of Local Archeological Artifacts. *Knowledge Organization* 40(2): 123–135. DOI: https://doi.org/10.5771/0943-7444-2013-2-123

One Geology. 2019. OneGeology, what's it all about? Available at http://www.onegeology.org/home.html [Last accessed 20 April 2019].

Oonk, S and **Spijker, J.** 2015. A Supervised Machine-Learning Approach towards Geochemical Predictive Modelling in Archaeology. *Journal of Archaeological Science* 59: 80–88. DOI: https://doi.org/10.1016/j.jas.2015.04.002

Open Data Charter. 2019. Who we are. Available at https://opendatacharter.net/ [Last accessed 20 April 2019].

PAN. 2019. Portable Antiquities of the Netherlands project. Available at https://www.portable-antiquities.nl/ [Last accessed 20 April 2019].

Pelagios commons. 2019. Pelagios commons. Available at http://commons.pelagios.org/ [Last accessed 20 April 2019].

PeriodO. A gazetteer of period definitions for linking and visualizing data. Available at http://perio.do/en/ [Last accessed 20 April 2019].

Perry, M and **Herring, J.** 2012. *OGC GeoSPARQL − A Geographic Query Language for RDF Data.* OGC 11-052r4. Open Geospatial Consortium. Available at: http://www.opengis.net/doc/IS/geosparql/1.0 [Last accessed 31 March 2019].

PLEADE. 2019. Pleade a complete, free and Open Source solution for heritage valorization. Available at http://pleade.com/ [Last accessed 20 April 2019].

Rabinowitz, A, Shaw, R, Buchanan, S, Golden, P and **Kansa, E.** 2016. Making Sense of the Ways We Make Sense of the Past: The PeriodO Project. *Bulletin of the Institute of Classical Studies* 59(2): 42–55. OOI: 10.1111/j.2041-5370.2016.12037.x

Rajala, U. 2004. Sense and Sensibility – Reflections on the Epistemology and Ontology of GIS Studies. *Internet Archaeology* 16. DOI: https://doi.org/10.11141/ia.16.2

Richards, JD. 2002. Digital Preservation and Access. *European Journal of Archaeology* 5(3): 343–366. DOI: https://doi.org/10.1177/146195702761692347

Roorda, IM and **Wiemer, R.** 1991. Towards a New Archaeological Information System in The Netherlands. In: Lock, G and Moffett, J (Eds.) *CAA91. Computer Applications and Quantitative Methods in Archaeology 1991.* BAR International Series S577. Oxford: Tempus Reparatum, pp. 85–88.

Scape Trust. 2019. Scotland's Coastal Heritage at Risk. Available at http://scharp.co.uk/ [Last accessed 20 April 2019].

Schweizerische Eidgenossenschaft. 2019. Patrimoine pour tous / Kulturerbe für alle / Patrimonio per tutti. Available at https://kulturerbefueralle.ch/ [Last accessed 20 April 2019].

Seitsonen, O. 2017. Crowdsourcing Cultural Heritage: Public Participation and Conflict Legacy in Finland. *Journal of Community Archaeology & Heritage* 4(2): 115–30. DOI: https://doi.org/10.1080/20518196.2016.1252129

Shaw, R, Corns, A and **McAuley, J.** 2009. Archiving Archaeological Spatial Data: Standards and Metadata. In: *Online Proceedings Computer Applications to Archaeology 2009 Williamsburg, Virginia, USA. March 22–26, 2009*, Williamsburg (VA), pp. 1–15. Available at http://archive.caaconference.org/2009/articles/Shaw_Contribution187_c%20(1).pdf [Last accessed 31 March 2019].

Sohlenius, R. 2014. *Förstudie FMIS-Processen [Prestudy of FMIS Process].* Stockholm: RAÄ.

UN-GGIM. 2018. *Integrated Geospatial Information Framework. A Strategic Guide to Develop and Strengthen National Geospatial Information Management. Part 1: Overarching Strategic Framework.* Available at http://ggim.un.org/meetings/GGIM-committee/8th-Session/documents/Part%201-IGIF-Overarching-Strategic-Framework-24July2018.pdf [Last accessed 31 March 2019].

Valenti, M and **Nardini, A.** 2004. Modello Dei Dati e Trattamento Del Dato Sul GIS Di Scavo. *Archeologia e Calcolatori* 15: 341–358. Available at http://www.archcalc.cnr.it/indice/PDF15/20_Valenti.pdf [Last accessed 31 March 2019].

van't Veer, R, Bloem, P and **Folmer, E.** 2018. Deep Learning for Classification Tasks on Geospatial Vector Polygons. arXiv: 1806.03857. Available at https://arxiv.org/abs/1806.03857 [Last accessed 31 March 2019].

Verhagen, P, Vossen, I, Groenhuijzen, MR and **Joyce, J.** 2016. Now You See Them, Now You Don't: Defining and Using a Flexible Chronology of Sites for Spatial Analysis of Roman Settlement in the Dutch River Area. *Journal of Archaeological Science: Reports* 10: 309–21. DOI: http://dx.doi.org/10.1016/j.jasrep.2016.10.006

Verschoof-van der Vaart, WB and **Lambers, K.** 2019. Learning to Look at LiDAR: The Use of R-CNN in the Automated Detection of Archaeological Objects in LiDAR Data from the Netherlands. *Journal of Computer Applications in Archaeology* 2(1): 31–40. DOI: http://doi.org/10.5334/jcaa.32

Wilcke, WX, de Boer, V and **van Harmelen, FAH.** 2017. User-Driven Pattern Mining on Knowledge Graphs: An Archaeological Case Study. In: Duivesteijn, W, Pechenizkiy, M, Fletcher, GHL, Menkovski, V, Postma, EJ, Vanschoren, J and van der Putten, P (Eds.) *Benelearn 2017: Proceedings of the Twenty-Sixth Benelux Conference on Machine Learning, Technische Universiteit Eindhoven, 9–10 June 2017.* Eindhoven: Technische Universiteit Eindhoven, pp. 137–39. Available at https://pure.tue.nl/ws/files/72619856/benelearn_2017.pdf [Last accessed 31 March 2019].

Willems, WJH. 1997. Archaeological Heritage Management in the Netherlands: Past, Present and Future. In: Willems, WJH, Kars, H and Hallewas, DP (Eds.) *Archaeological Heritage Management in the Netherlands − Fifty Years State Service for Archaeological Investigations.* Amersfoort: ROB, pp. 3–34.

Willems, WJH. 2000. The Management of the Archaeological Heritage in the Netherlands. *Archeologia Polona* 38:153–168.

Wise, A and **Richards, J.** 1999. Digital Preservation in Archaeology. In: Dingwall, L, Exon, S, Gaffney, V, Laflin, S and van Leusen, M (Eds.) *Archaeology in the Age of the Internet. CAA97. Computer Applications and Quantitative Methods in Archaeology. Proceedings of the 25th Anniversary Conference, University of Birmingham, April 1997.* BAR International Series 750. Oxford: Tempus Reparatum, pp. 137–144.