Deconstructing the Digital Infrastructures Supporting Archaeological Knowledge

Jeremy Huggett

Abstract

The last 30 years have seen significant investments in the development of digital infrastructures to support archaeological practice. From field recording systems to national data archives, these have come to play an increasingly dominant role in the collection, management and access to data used in the creation of new archaeological knowledge. While archaeologists have paid a lot of attention to the technical creation of these infrastructures, much less is said about their wider political, cultural and social aspects. Despite this, more and more countries are building digital infrastructures to support cultural heritage management, the curation of archaeological data and to provide access for data reuse. A lack of critical reflection surrounding these infrastructures opens archaeologists, heritage organisations and their wider user communities to unforeseen outcomes, hidden socio-political and technical biases, and the promotion of conventions and processes which ultimately carry consequences for knowledge practice. The way that infrastructures become embedded in practice means that a critical understanding of their implementation and application – the opportunities they offer, the constraints they impose, and the perspectives they adopt – needs to become part of a wider debate surrounding their informed use.

Keywords: infrastructure, data, knowledge, interoperability, sustainability, standards, metadata, interface, failure

Archaeology, School of Humanities, University of Glasgow
jeremy.huggett@glasgow.ac.uk

This is an Open Access article distributed under the terms of the Creative Commons 4.0 International licence (CC BY 4.0) (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction

In recent years, archaeology has seen a rapid growth in the collection and use of born digital data alongside a growing dependence on digital cultural heritage data infrastructures. During this time, debate surrounding these infrastructures has been limited and narrowly focused, and many of the issues remain largely unchanged. In 1981, for example, Gaines (1981b:vii) wrote of a new focus on databank management in archaeology arising from the growth in interest in complex questions that required large and diverse bodies of data. At the same time, she noted problems associated with access, control, the variability of archaeological data, and issues associated with the selection and use of thesauri (Gaines 1981a:224). The intervening period has seen the development of a range of project- and organisation-based (for example Intrasis, ARK, FAIMS), national (for example ADS, tDAR, DANS, SweDigArch) and international data infrastructures (for example ARIADNEplus), using dramatically faster hardware, cheaper and more extensive storage, more complex software and more elaborate data structures. Nevertheless, archaeologists continue to wrestle with many of the same kinds of challenges.

There is an extensive body of work accompanying such developments, ranging from discussions of technical standards, requirements and tools through to perspectives on the creation and management of digital networked archives. Inevitably, this work is presented primarily by those involved in developing or implementing the systems, in part reflecting the relative novelty of the tools, but this raises the prospect of what can be characterised as an 'advocacy perspective' (Meyer & Schroeder 2015:183), an 'institutionalized discourse' (Mongili & Pellegrino 2014:xxiii) or a 'master narrative … [a] voice which speaks unconsciously from the presumed center of things' (Star 1999:384). As a result, it risks a tendency to technosolutionism (for example Paris et al. 2023:18), reinforcing a political status quo, prioritising particular values (for example Slota & Bowker, 2015:2; Gupta et al. 2023:78), and presenting a narrative bias (for example Pollock & Williams 2010:529). Certain infrastructures can also dominate the attention space by virtue of being early, and thereby inadvertently encourage certain approaches which are deemed useful while closing down others. For example, the Archaeology Data Service (ADS) is frequently used as an exemplar because of its early lead as the oldest archaeological digital repository in the world (Richards 2021), and so has exerted considerable influence on developments elsewhere.

Coupled with this advocacy perspective is the way in which debates focus on the components of these large-scale archaeological systems, rather than on the systems themselves. For example, there are (rightly) extensive

debates over terminologies, structures, organisation, policies, ethics and so on, but beyond the desirability of their creation, issues surrounding the nature of the infrastructures that these aspects contribute to are less commonly debated. As Wright and Richards (2018:S60) observe,

> … there is continued emphasis on technological and methodological innovations themselves rather than on the complex social factors that contribute to their success or failure and the connections they facilitate, but this has begun to change.

Discussions which extend beyond this technical/methodological focus include Wright and Richards (2018) on broader questions of stewardship and equity, Kansa (2022) on dependencies and sustainability, Optiz et al. (2021) on the support of transdisciplinarity, and the wider information studies perspective of Börjesson and Huvila (for example Börjesson & Huvila 2018; Huvila 2019a; Börjesson 2021). However, even in cases such as these, the tendency is to focus on specific areas or aspects, rather than examine the development and influence of the infrastructure as a whole. Consequently, there is a sense in which the study of these large-scale systems is approached in a bottom-up manner, looking in detail at their components and debating their utility, examining the parts rather than the whole (cf. Gupta & Devillers 2017:872). As a result, there is a gap in our approach to digital infrastructures: the focus on components and individual aspects means that an oversight of the nature of the whole infrastructure is lost within the detail. What are we building these infrastructures for? How do these infrastructures influence our practice? Are there alternative conceptions of archaeological digital infrastructures to those currently in use? And, how do the technical, political and ontological decisions made during the construction of these infrastructures influence the creation of archaeological knowledge? To begin to address questions such as these requires raising the gaze from specific tools, terminologies, structural models and so on, in order to take a broader perspective on the development of digital archaeological infrastructures.

What are infrastructures?

The term infrastructure itself is described as a '"plastic word" often used to signify any vital and widely shared human-constructed resource' (Thylstrup 2018:26). Larkin (2013:329) writes of the 'peculiar ontology' of infrastructures, in that they are both things and the relations between things, and this duality makes them 'conceptually unruly'. The most widely cited definition of infrastructure is that of Star and Ruhleder (1996:113;

see also Bowker & Star 1999:35) who characterise infrastructure as being embedded in other structures, transparent (in the sense of invisible through habituation, for example, becoming visible through breakdown), extending beyond a single event or place, learned as part of membership of a particular community of practice, incorporating standards and conventions, built on and constrained by an existing base (requiring backwards compatibility with prior works, for instance) and incrementally modified through negotiation and adjustment to other systems. From this, infrastructures emerge as:

> … complex, adaptive sociotechnical systems, made up of many interacting agents and components. Some of these are technological: buildings, devices, software and other artifacts. Others are social: organizations, standards, laws, budgets and political arrangements. Finally, some are human individuals who contribute to the infrastructure's development and maintenance or simply make use of it in their daily lives (Edwards 2019:356).

Such explicit attempts to define infrastructure are rare in the digital archaeology literature, although Huvila (2018:128) echoes Star and Ruhleder's characterisation, emphasising the interrelationships between people, technologies and archaeological practices. Otherwise, a variety of usages can be gleaned from the context in which the term is used. For example, Huvila (2019b:149) distinguishes between knowledge management systems which are primarily project- or organisation-based, infrastructures such as the ADS or the Digital Archaeology Record (tDAR), meta-infrastructures such as ARIADNEplus which integrates multiple infrastructures under one interface, and virtual research environments managing the research data lifecycle. Alternatively, infrastructure may be used in the sense of the software system itself or the technical requirements to operate it. For example, the Field Acquired Information Management Systems (FAIMS) developed for structured archaeological data collection is described as infrastructure (for example Sobotkova et al. 2016:338), whereas discussion of the Archaeological Recording Kit (ARK) refers to infrastructure as the network access required for the software to operate online (Dufton 2016:382). Likewise, the infrastructure of the Silchester virtual research environment (VERA) focuses on the broadband and WiFi network together with the hardware used to run the software (Dunn 2011:100–101). Beyond software systems or technical underpinnings, Niccolucci and Richards (2013:82) emphasise the human component of a research infrastructure, and Benardou et al. (2017:3) underline the importance of the research community in their description of infrastructure as 'scholarly ecosystem'. Closest perhaps to Huvila's (2018:128) use of the term, Kansa (2022:1412–1416) takes a broad view of infrastructure as systems necessary to support archaeologi-

cal information management and communication ranging across curation and communication infrastructures, software and data infrastructures, security infrastructures, social infrastructures and the dependencies associated with them.

This inconsistent reference to infrastructure is not unique to archaeology; more generally the term is frequently used in overlapping ways and in combination with others. Unpacking these reveals changing attitudes to infrastructures and the repositioning of infrastructural studies over a relatively short period of time. For example, *information infrastructures* became seen as key to research via the provision of a wide range of resources including centres, repositories, standards, visualisation tools and high performance computing (Bowker et al. 2010:98). Recognition of the problematic separation of data from information and knowledge led to information infrastructures becoming seen as *knowledge infrastructures*, consisting of:

> … the network of institutions, people, buildings, and information resources which enable us to turn observation and contemplation of the world into a standardized set of knowledge objects (Bowker 2017:391).

At the same time, information infrastructures became more narrowly defined, focusing on technical communication architectures, or to national or international policy frameworks rather than the systems themselves (Borgman 2015:33). More recently still, the rise of big data approaches and development of deep learning and neural networks has led to the notion of *thinking infrastructures*. These are distinguished from knowledge infrastructures by their more collaborative, distributed and decentralised nature, and their elimination of intermediaries. While knowledge infrastructures generally distinguish between knowledge producers and consumers, thinking infrastructures remove this clear separation (Bowker et al. 2019:9), and by extension, knowledge producers may become the systems themselves, as large language models applied in natural language processing increasingly risk being perceived (for example Bender et al. 2021).

Where archaeological enterprises sit in this infrastructural spectrum is open to debate, although most could be seen as information infrastructures focusing on the creation and management of resources, with more developed examples perceived as knowledge infrastructures supporting the construction of archaeological knowledge through access to large bodies of data. More realistically, however, the changing conception of infrastructures and the inconsistent application of the different interrelated terms means that infrastructures in archaeology contain a mixture of elements drawn from across these approaches without necessarily falling explicitly in one category or another.

Why focus on infrastructures?

The variability in definition and usage highlights the range of concepts that are embedded in the notion of infrastructure. Its imprecise use within digital archaeology tends to disguise this, despite the way that 'Infrastructure both enables and constrains what we can and cannot accomplish and how we go about our own work' (Kansa 2022:147; see also Huvila 2018:138). Few archaeological digital infrastructures are more than 20 years old. Most are considerably more recent, and many are in the early stages of development (for example SweDigArch, and see contributions to Jakobsson et al. 2021). None can be thought of as complete. Even older-established infrastructures remain works in progress: for example, the ADS may be recognised as an exemplar of best practice but after more than 25 years of effort it is possible to claim only that 'it is still far from standard, but the situation is improving' (Richards 2021). Similarly, tDAR has had mixed success in attracting depositors despite legal requirements for the long-term preservation of data (Witze 2019:42–43; see also Nicholson et al. 2023:64). Infrastructures are increasingly seen as essential – even transformative – for archaeological practice, required in order to undertake large-scale data analysis, integration and synthesis, and to enable archaeology to make a contribution to large transdisciplinary scientific research questions such as long-term social dynamics and climate change (for example Kintigh 2006:573; Buckland & Sjölander 2022:110). Viewed in this light, infrastructures do not simply curate and provide access to data but provide the means for developing the discipline in new directions (for example Kintigh et al. 2015:3; McManamon et al. 2017:240; Meghini et al. 2017:2), supporting novel analytical methods and knowledge creation and thereby increasing the reliance of practitioners upon their access and use (Börjesson 2021:1642). Investigation of the nature of these infrastructures before they are considered complete, wholly disappear into the background, and become taken for granted components of archaeological practice, is therefore crucial (for example Marttila & Botero 2017:103; Karasti & Blomberg 2018:237).

This is because infrastructures not only facilitate new methods and support new opportunities; they also limit practice and close down alternative approaches, as Kansa and Huvila have previously observed (above). For instance, most repository infrastructures organise data in particular ways to facilitate its discovery, requiring conformity with an institutionalised worldview. From the earliest studies of infrastructures (for example Star & Ruhleder 1996:113; Bowker & Star 1999:35) one of the characteristics identified with them is their transparent – as in invisible – nature: a 'good' infrastructure is one which disappears into its surroundings (Millerand & Baker 2020:10). Edwards (2019:358) identifies three forms of transpar-

ency: hiding or black-boxing the underlying technologies and techniques; habituation or invisibility through familiarity; and what he calls 'infrastructuration', whereby the infrastructure 'both shapes and relies upon the continual performances or rehearsals of agents'. Additionally, invisible work is involved in maintaining the systems that underpin the infrastructure which goes largely unrecognised (Borgman 2015:34). There are clearly ethical as well as practical reasons for addressing such invisibilities (for example Dennis 2020; Huggett 2021:424–429).

Infrastructures need therefore to become the centre of analysis (for example Heine & Meiske 2022:11), rather than simply treated as the means by which data are gathered and analysed. Examining infrastructures in this way encourages the study of the formation of knowledge and its contexts of creation, offsets the advocacy perspective, promotes the invisible labour entailed within infrastructures and incorporates a range of broader social and environmental issues (Heine & Meiske 2022:11–12). Ultimately it is important to recognise that infrastructures are situated culturally, socially, politically, technologically and spatially (for example Svensson 2015:338), which should make a critical appreciation of their design, purpose, development and implementation a necessary precursor to their use.

Building infrastructures

Infrastructures are best conceived as emergent phenomena rather than being carefully designed or directed from the outset:

> … its eventual ends and forms will not be fully contained in its beginnings, but rather subject to change through the intricacies of scaling, transfer, consolidation, etc. (Jackson et al. 2007).

Jackson et al. (2007) suggest that ecological metaphors (nurturing, growing and so forth) might be better associated with the development of infrastructures 'to capture the sense of an organic unfolding within an existing (and changing) environment' (Edwards et al. 2009:369). Such a representation fits with the image of the infrastructure as an 'ecology of people, practices, technologies, institutions, material objects, and relationships' (Borgman 2015:4), all of which are in flux with each other. Infrastructures are a process of enactment, always in-the-making (Parmiggiani 2017:208). It is an approach which encourages balance:

> … conceptualizing infrastructure as a process over time ensures that the technical and logistical sides of infrastructure are not privileged over, or seen as separate from, its social and political, or formal and aesthetic sides (Appel et al. 2018:17).

EMERGENCE

During its unfolding, an infrastructure frequently encounters unexpected changes imposed by unforeseen limitations and unanticipated demands placed upon it. For example, the Archaeology Data Service (ADS) was originally envisaged as a distributed system, recognising that archaeological information was held and maintained by a wide variety of institutions (for example Richards 1997:1058). Consequently, the ADS was conceived as a central brokering hub using metadata to link to the datasets held in museums and local archives and to the data in regional Historic Environment Records and the National Monuments Records held separately for England, Scotland, Wales and Northern Ireland. Only orphan datasets which had no alternative home would be held by the ADS itself. As part of this distributed focus the ADS supported organisations in acquiring online access: for example, a joint project between the Royal Commission on the Ancient and Historical Monuments of Scotland (RCAHMS, now Historic Environment Scotland) and the ADS saw the launch of CANMORE-Web for the National Monuments Record for Scotland in 1998 (ADS 1997:6, 1998:8; Richards 1997:1058). Relative to other disciplinary services at the time, it was always claimed that a unique aspect of the ADS was that its data was derived from the destruction of primary evidence, but it was this distributed emphasis that really set the ADS apart. However, in the early development years the stress on this distributed nature shifted for a variety of reasons, including the realities of the available technological infrastructure at the time, slow uptake of internet access by potential partners, and the need for demonstrable products to satisfy funders' requirements. With external links largely unfeasible or unreliable, the emphasis instead became the development of the central metadata index to resources and the licensing of copies of datasets to be held centrally by the ADS rather than accessed remotely (ADS 1998:8).

INTEROPERABILITY

Underlying this ambition for distributed access to data is the infrastructural concept of interoperability. Interoperability 'allows digitized cultural memory institutions to exchange and share documents, queries, and services' (Thylstrup 2018:67) and is seen as a key feature of data infrastructures. Interoperability enables the bringing together of multiple datasets while avoiding their treatment as a single body of evidence (for example Leonelli & Williamson 2023:7). This is distinct from the integration or aggregation of multiple datasets, which is a key feature of big data methodologies, for example, and which places considerable demands on the standardisation of data and is not feasible where the data are of radically different origins

(an excavation database versus a national monuments record, for instance) except at the most general level. Interoperability is a 'more responsible form of data linkage' (Leonelli & Williamson 2023:7), as it enables incompatible data to be connected without permanently changing their nature, although the level of standardisation necessary may still be of concern (Williamson & Leonelli 2023:105). Interoperability effectively creates networks of datasets and infrastructures which connect across a range of interfaces:

> … numerous systems, each with unique origins and goals … are made to interoperate by means of standards, socket layers, social practices, norms, and individual behaviors that smooth out the connections among them (Edwards et al. 2013:5).

Interoperability therefore does not solely operate at the data or the technical level – it also operates at the social level of infrastructures (cf. Thylstrup 2018:68).

For example, the distributed vision of the ADS did not disappear because of the early limitations encountered. The ADS launched HEIRPORT in 2002, a proof-of-concept portal for the Historic Environment using the Z39.50 communication protocol for the search and retrieval of data over TCP/IP networks to link the ADS with targets hosted by organisations across England, Scotland and Wales (Austin et al. 2002). Since the now-centralised ADS catalogue would always be out of sync with external data providers such as the regional and national monuments records, live searchable gateways to these resources would ensure that results returned remained current (Richards 2002:353). Similarly, the ADS ARENA (Archaeological Records of Europe Network Access) project used Z39.50 to simultaneously search data held by six partner organisations across Europe (Kenny & Richards 2005; Waller 2005). However, the Z39.50 protocol was not as reliable as it might be, and users were frequently faced with one or more unavailable targets, making cross-resource searching something of a lottery. This kind of direct cross-searching across different data targets has since been dropped, even though technological developments make it more feasible than before. Instead, ARENA's successors, ARIADNE and ARIADNEplus, employ a centralised metadata catalogue rather than using direct connections out to data providers, although the metadata does provide links to source data where available. Similarly, HEIRPORT was not further developed and the centralised ADS metadata catalogue only links to individual external records where accessible. In both cases, interoperability becomes the means by which the centralised metadata index is updated in the absence of direct cross-searching of local and regional resources. One advantage of this approach is that it only requires the high-level index terms to be translated and standardised across the various providers.

Such accommodations for interoperability illustrate the effects of unforeseen limitations and difficulties encountered at the interfaces between infrastructures. Less obviously, it also demonstrates the importance of social interoperability through personal networks – for instance, ARENA was a network of individuals brought together through personal contacts and common perspectives (Kenny & Richards 2005, sec. 3.1). Similarly, the original consortium members behind the creation of the ADS were a network of friends and colleagues spread across various universities. Such social aspects are rarely emphasised in accounts, and the significance of personal contacts and social engagement in infrastructural development is largely unrecorded. The ADS example also demonstrates that infrastructures do not develop in isolation but are more often grafted onto other existing infrastructures (for example Meyer & Schroeder 2015:183). For example, the ADS worked alongside a mixture of long-established digital and paper-based infrastructures including Historic Environment Records, National Monuments Records, with their own standards, recording systems, regulations, responsibilities, and funding lines, requiring complex and at times delicate manoeuvring amongst all parties to ensure social and political interoperability.

SUSTAINABILITY

Sustainability and resourcing over the long term is a key issue for infrastructures: indeed, from the outset at the ADS it was recognised that its short-term funding was in tension with its role as a data archive. This was one reason behind the proposed distributed model and the emphasis on only archiving orphan datasets: should funding cease, most data resources would remain unaffected by the closure of the ADS. Infrastructures are by nature fragile, and long-term preservation and maintenance of access entails cost and effort, with lack of investment leading to rapid degradation (for example Borgman et al. 2019:901; Millerand & Baker 2020:21). The invisibility of an infrastructure embedded into regular practice as a structuring force can give it an illusion of permanence, the risk only revealed when it breaks down and its functionality is, even temporarily, unavailable (for example Huggett 2022:271–272). Recognition of the indispensability of an embedded infrastructure is one way to establish a case for its continued sustainability.

The ADS has depended upon the successful negotiation of cycles of funding and defunding over more than 20 years. At any point, the sustainability of the ADS was under varying degrees of threat, and in recognition of this instability the ADS developed a legacy fund to support the winding down or transfer of the service should that prove to be necessary. The ADS was initially established in 1996 as part of a UK university-based e-infrastructure initiative, the Arts and Humanities Data Service (AHDS), funded by the

Joint Information Systems Committee (JISC) of the UK Higher Education Funding Councils (Richards 1997:1057) with additional funding subsequently provided by the Arts and Humanities Research Board (now Arts and Humanities Research Council, AHRC). However, the AHRC withdrew funding for the AHDS in 2008 following a review, making it unviable and forcing JISC to remove funding (see below). The ADS was unique amongst the AHDS service providers in being able to negotiate a period of transitional funding following the closure of the AHDS to allow it to evolve to a commercial funding model. Several factors were key to the negotiation of this transitional arrangement. The ADS had already developed a charging policy for depositors and so was able to demonstrate a potentially viable funding stream (Hardman & Richards 2013:76). The ADS was also unique amongst the AHDS services in having a close relationship with commercial and governmental organisations, core to its initial conception (see above). Furthermore, it was argued that archaeological data had a unique quality in that they were often the only surviving outcome of the destruction of primary evidence. In combination, a successful case was made to the AHRC for a period of transitional funding, the shortfall to be made up by growing commercial and research project income. This transition was not without its problems, but the fact that the ADS still exists is testament both to its ability to attract funding and to the level of community support, from the staff themselves through to data providers and project funders. Recent changes in policy mean that the ADS is again in receipt of core infrastructure funding from the AHRC (ADS 2021:3).

These ebbs and flows of funding and consequent concerns about sustainability had a range of infrastructural impacts, both on the technical and personal side of the ADS. For example, supporting the case for sustainability requires constant demonstration of value and in turn this demands not simply maintenance but also development, albeit in limited areas. Commercial imperatives can be seen in the prioritisation of tools such as ADS-easy to streamline the deposition of project archives, while the main ArchSearch interface to the collections remained largely unchanged during this period. Development work continued elsewhere, such as the implementation of the ARIADNE/ARIADNEplus portal, but only where project or commercial funding was specifically targeted. The arrival of significant new AHRC core funding in 2022 has enabled the public face of the ADS website to be redeveloped and a new version of ArchSearch will be launched shortly.

The experience of the ADS suggests that a national data infrastructure can be sustained using a blend of commercial and project funding, but it is far from ideal and takes a toll on the personnel and the profile of the organisation. The ADS is certainly not unique in this regard. In the USA, for instance, tDAR has similar funding challenges in juggling multiple grants

over time from the National Science Foundation (NSF) and the Andrew W. Mellon Foundation amongst others (Kintigh et al. 2018:32), including drawing upon public donation. As Wright and Richards (2018:S61) observe,

> Archaeologists are continually encouraged to find ways to make their work marketable within commercial frameworks and this is invariably part of any sustainability plan, but rarely produces significant revenue. Successful models for the long-term stewardship of archaeological data remain limited.

Buckland and Sjölander (2022:125) suggest that access to national funding is a necessity for these kinds of research infrastructure, particularly appropriate if those infrastructures are embedded within a national regulatory framework. However, while contributions to Jakobsson et al. (2021) show many of the countries represented have some kind of regulatory framework in place, infrastructural support for archiving, managing and making data available does not automatically follow. The requirements of funding bodies that data should be made open (for example Richards et al. 2021) might suggest that adequate secure infrastructural funding would follow, but again this is not necessarily the case, or not in large enough or regular enough quantities. It is therefore perhaps not surprising that in their discussion of the development of a national data integration infrastructure for the USA, Ortman and Altshul (2023:99) are uncertain about the most suitable organisational setting for such a service (private company, non-profit, university or other) given the uncertainties surrounding resourcing. The paradox common to all is how a long-term infrastructure can be securely established on the back of largely short-term, one- to five-year funding cycles, and the consequent challenges associated with the management of the inevitable periods of financial drought and uncertainties of insecure staffing.

FAILURE

Not all infrastructures are destined to succeed or to survive long-term in the face of resourcing challenges, technological change or competing infrastructures. Such issues are expected but unpredictable and associated with the unfolding nature of infrastructures which do not follow predictable, linear developmental paths (for instance, Karasti & Blomberg 2018:239). Indeed, to assume infrastructures are orderly, dependable and immune to social and technological change is to more or less guarantee failure (Edwards 2003:195).

As discussed above, the ADS faced potential closure in 2008 with the ending of core AHRC and JISC funding for the AHDS. The AHRC withdrew funding for three reasons (Collins 2012:166): it claimed that researchers had now gained the technical know-how to undertake their own data curation; it considered long-term storage and sustainability was best handled within universities rather than a centralised service; and, consequently, the AHDS

funding could be best used elsewhere. The AHRC also considered that relatively little use was made of the resources held by the AHDS, making the cost unjustified (Robey 2012:150). Despite vigorous objections that the AHRC had misread the situation and threatened fragile digital resources, the AHRC withdrawal paid little attention to suitable exit strategies, the ADS aside. One of the lessons drawn by Robey (2012:153) is the importance of visibility in relation to sustainability: digital resources need to be recognised and used as much as possible by their target communities in a form of network effect. This was clearly a strength of the ADS which supported its case to move to a transitional funding arrangement rather than have all core funding abruptly removed (see above).

While the ADS avoided failure amidst the collapse of the AHDS infrastructure, the Archaeology Data Archive Project (ADAP) was not so fortunate. The ADAP was established in 1994 within the Center for the Study of Architecture, a not-for-profit organisation in Bryn Mawr, USA (Eiteljorg II 1995:245). A key reason given for the discontinuation of the ADAP in 2002 was that it had been unsuccessful in attracting data. According to Eiteljorg (2011:262), it had failed to attract a single completed dataset other than one taken from an already-published CD-ROM: 'only a handful of scholars' had deposited files (Eiteljorg II 2002). With funding for a pilot project to support the archiving process not forthcoming, it was determined that the ADAP should cease operation and those files that were held were returned to their original depositors (Eiteljorg II 2002). Primarily the ADAP was unable to become self-sustaining as there was little evidence that data depositors were able or willing to pay for the service (Eiteljorg II 2001). Network effects apply again, since it seemed unlikely that a tipping point would be reached within a reasonable timespan whereby the resources became useful for analytical purposes and might therefore attract grant funding (Eiteljorg II 2002).

Arguably the ADAP never reached what might be called infrastructural status but there are lessons to be drawn from its experience. In comparison to the ADS, the ADAP clearly fell short in creating visibility, attracting resources and becoming embedded in its target community by clearly and unambiguously demonstrating its use-value. Although, in retrospect, Eiteljorg (2011:262) argued that it was 'unrealistic to expect that a more complex and costly approach offering no better rewards will succeed where a simpler and much less costly one failed', paradoxically a charge-free approach may have discouraged potential data providers unable to envision a long-term future for the archive. Likewise, the organisational setting (cf. Ortman & Altschul, 2023:99) may have been a factor: had the ADAP been hosted within a university or similar institution there might have been more confidence in its future outcome. Furthermore, by 2008 the ADS had already

become closely associated with larger pre-existing infrastructures across UK archaeology, including the national governmental organisations responsible for the management of cultural heritage. This provided the ADS with powerful and influential supporters whereas the ADAP had relatively few advocates to make the case on its behalf.

Infrastructures and social processes

Social and cultural aspects of infrastructure have not been a strong component of archaeological studies although these are frequently implicit in discussions of archaeological practices which make use of infrastructures (Huvila & Huggett 2018:93; see also Dallas 2015). More than technical constructs, infrastructures are not things but 'bundles of relationships' and,

> … whether in collaboration, deliberation, or conflict, individuals and communities come together around them and interact in ways that have lived implications (Carse 2016).

Ethnographic approaches lend themselves to a study of these relationships, but this has not been a significant feature of archaeological infrastructure studies to date (although see Huvila 2016; Börjesson & Huvila 2018; Börjesson 2021, for example). People contribute to the development and maintenance of the infrastructure while others will simply use it, and these relationships will change over time (Edwards 2019:356). Other groups of individuals may exert influence without closer engagement, their involvement limited to its approval and its funding, for instance. Still others may be part of a broader community who, while not users, nevertheless appreciate knowledge of (and perhaps critique) its existence. The social constellation associated with an infrastructure is therefore more extensive and membership more flexible than is commonly claimed.

VALUES

This broad collection of communities and individuals influence an infrastructure in crucial ways:

> All infrastructures embed social norms, relationships, and ways of thinking, acting, and working. As a corollary, when they change, authority, influence, and power are redistributed (Edwards et al. 2013:23).

Values embodied in the infrastructure may introduce certain biases or politics into the system (for example Slota & Bowker 2015:2), such as through the incorporation of a particular set of regulations or standards. Some values may also be in conflict with each other. For instance, Huvila (2016)

describes the standardisation approaches of early-established archaeological data archives as 'attempts to seize control and find footing' in other areas, which is problematic in terms of imposing restrictive requirements and yet a necessary component of a successful infrastructure which 'has to be established as a network of relationships between all parties' (Huvila, 2016). Similarly, Buckland and Sjölander (2022:115–116) point to the tension between researcher-friendly designs, which potentially allow more imaginative approaches to data analysis, and developer-friendly designs which are likely to entail efficient coding and documentation and hence are easier to maintain. There may also be ethical challenges: for instance, the low spatial resolution of the data used in the Digital Index of North American Archaeology (DINAA) places restrictions on reproducibility while being an important means to address colonial issues associated with disadvantaged descendant groups (Kansa 2022:143–144). Elsewhere, the Portable Antiquities Scheme (PAS) database for England and Wales restricts the availability of location data to bona fide researchers to avoid looting. This restriction is a condition of reporting but may also be a consequence of collectors seeking to withhold the knowledge in order to preserve exclusive access (for example Barford 2020:108; Brodie 2020:91). In both cases, influences external to the infrastructure impact on the way in which data are presented.

COMMUNITIES

Identifying the range of social groups associated with an infrastructure and their interactions is an important part of understanding that infrastructure: for example, there may be communities of researchers, 'intentional' communities (special interest or support groups), and communities of practice (Bowker et al. 2010:105). These are frequently grouped under the heading of 'users' as distinct from 'developers', although users may be further categorised in a variety of ways. In a study of the ADS prepared for JISC (Beagrie & Houghton 2013) users are divided into two categories: 'depositors' and 'users', collectively referred to as 'stakeholders' (Beagrie & Houghton 2013:6), a narrow perspective which reflects the study's limited focus on the value of the collections. As Huvila (2016) argues, such studies are focused on:

> … estimating the (positive) societal significance of the repositories rather than critically explicating how the repositories are linked to the everyday practices of the different groups that influence or are affected by the repositories.

Millerand and Baker (2010:141) characterise three kinds of user: the 'hands-on user', who is engaged with the definition and development of the system; the 'social actor', who generates, exchanges, and consumes information

from the system; and the 'sociopolitical actor', whose role and position is affected and impacted by the system. In the ADS, for instance, the various management and advisory committees could be characterised as hands-on users given their role in defining and overseeing the development of the ADS; the social actors are represented by the data depositors and data users; while representatives of the various national archaeological bodies who interacted with the ADS might represent the sociopolitical actors. Of course, individuals may move between roles or hold several roles simultaneously: in the ADS, people may be both hands-on users and social actors, committee members and data depositors/users, for example.

This remains a partial picture with regard to social roles associated with infrastructures. For instance, from their case study Millerand and Baker (2010:143) identify three further groups: 'informatics specialists', essentially the developers who build the tools and work on the metadata specifications; 'scientists', researchers who are users of the system and its datasets; and 'information managers', responsible for curating the data and implementing standards. Again, there is a degree of overlap with 'scientists' broadly equivalent to 'social actors', for instance, but these further characterisations usefully broaden the range of social roles beyond the generic user. In the ADS, for example, applications developers, web developers and system managers can be seen to constitute the informatics specialists, while archives officers and collections managers may be identified as the information managers. Unsurprisingly, this broadly maps onto the staffing structure that exists within the ADS. A key advantage of this further categorisation is that it draws attention to the staff operating the infrastructure who may otherwise be largely invisible in accounts.

Relationships and interactions are not the same for everyone and depend on how they experience the infrastructure: some may find it supports their work, others encounter obstruction (for example Star & Ruhleder 1996:112–113; Star 1999:380; Edwards et al. 2013:13; Koch 2018:70–71). The negative aspects of infrastructure are frequently underestimated in general, and open to debate in archaeology, although the degradation of locational data referred to above might be one example. They may also be evidenced in the form of opposition, resistance, workarounds and the subversion of processes (for example Edwards et al. 2013:13–14; Huggett 2021:422–423), although this remains a poorly-explored area in archaeology. Elsewhere, differences in financial resources between communities, organisations and nations may negatively affect the ability to create and employ infrastructures, leading to a bias in objectives, structural characteristics and perspectives that favour the UK, Europe, and North America, for instance (Slota & Bowker 2015:5; cf. contributions to Jakobsson et al. 2021). It may also raise questions of sustainability in terms of differen-

tial availability and levels of funding for projects and programmes, and potentially the ability to support deposit fees and even access charges, for instance. Identifying the range of infrastructural communities therefore goes well beyond simply the recognition of audiences.

Infrastructures and knowledge creation

Data-based infrastructures are seen as a new form of cultural memory institution (Thylstrup 2018:22); indeed, infrastructures may be largely defined as being all about data (Edwards et al. 2007:31). However, what is lacking is a critical enquiry into the effects of infrastructural arrangements of data, and how these affect access and use of data in the construction of archaeological knowledge (although see Hacıgüzeller et al. 2021). As data is incorporated within infrastructures it becomes institutionalised, and the infrastructure determines what data and associated information will be available in the future (for example Borgman 2015:15). Such institutionalisation is seen as benign, even beneficial, if the alternative is data abandonment and loss, but infrastructural effects are critical to the use of data for knowledge creation as well as its long-term curation. Infrastructures reflect the priorities of the institutions behind them and the socio-political contexts in which they operate (Fullilove & Alimari 2023:66). Infrastructures are built on certain expectations or requirements, standards and protocols, which configure the data and its accessibility, making a critical perspective central to understanding their role in knowledge creation (Harvey et al. 2017:16). Both data and data infrastructures remain always in-the-making, and decisions taken concerning their treatment will affect the options and opportunities available to those who come after (Hacıgüzeller et al. 2021:1710). A potential paradox therefore exists: data are fundamental to knowledge creation and reliant on infrastructures to make them findable, accessible, interpretable and (re)usable (i.e. FAIR [for example see Nicholson et al. 2023]), but at the same time those infrastructures may limit certain actions, practices and relations (Van Rossem & Pelizza 2022:3). Discourse becomes centred on the data as represented within the infrastructure (cf. Lucas 2012:244), and consequently knowledge may become so deeply engrained that the infrastructure becomes difficult to challenge, or for new forms of expression or new ways of knowing to be considered (Bowker 2018:209).

STANDARDISATION

Standards are core to infrastructures: they enable them to behave in predictable ways to provide universal access, interoperability with other infrastructures, and assure technical sustainability into the future. The ADS

archive, for instance, is built on a range of standards at different levels, from the high-level Open Archival Information System (OAIS) model defining the basic components and functionality of an archive and its preservation issues, through a range of data content-related standards to low-level standards specified for preferred file formats. Standards therefore operate at every level of the archive, and some degree of standardisation is difficult to argue against given it is a prerequisite for sharing or linking data from different sources, even if the resulting complex and ill-defined web of standards is rarely discussed (Huggett 2012:542–543).

Bowker (2018:217) identifies what he calls a 'quintessential tension' with standards: an opposition between a desire for universality and the need for change. The risk of universality is that standardisation may encourage misinterpretation, disguise doubtful data sources, and facilitate 'data arbitrage' where the availability of data trumps its quality (Edwards et al. 2013:7). Rather than change, the costs associated with standardisation means there may be considerable inertia (Edwards et al. 2013:9). Once standards are set, they tend to stick, which is more often taken as a sign of success than as apathy. More significant than questions of effort or cost, however, is the way in which standardising data can change the nature of those data and promotes certain forms of knowing. It also makes some kinds of data undocumentable, and hence invisible (Van Rossem & Pelizza 2022:2). Strict schemas are problematic for messy archaeological data (for example Löwenborg 2018:51), and the study by Hacıgüzeller et al. (2021) shows how attempts to create structured data risk smoothing out variability or omit aspects which may not be represented within the data model. A category of data which frequently defies categorisation is the implicit or tacit knowledge behind the original data (for example Huggett 2020:9–11), and documenting it requires effort for which there is little resource, even if the desire is there (for example Opitz et al. 2021).

METADATA

Metadata are data about data but also a standard of standards. They are key to facilitating interoperability between datasets and infrastructures (for example Meghini et al. 2017:5). They also structure the data presented to the user: it is the metadata catalogues that are searched, and the results retrieved are based on those metadata. The metadata may be created automatically – for example, through a process of text mining to extract metadata (for example Richards et al. 2011) – or manually on accession. In either case, metadata is created as a high-level summary which allows data with similar characteristics to be identified.

However, metadata are frequently perceived as benign: they are not data themselves but a higher order of information (Boellstorff 2013, sec. 3),

underlined by the common view of metadata as providing administrative information about the data (who created it, who owns it). This attitude implies a lesser significance than data, yet metadata is more than simply a finding/linking aid and is capable of being treated as if it were primary data. Metadata increasingly shifts mode to be used as data in its own right: for instance, providing basic summary data including information about site/artefact types, time periods, and location, and used in everything from distribution analyses to 'big data'-style studies. Consequently, metadata becomes the data rather than simply how the data is located or linked, and is therefore another layer of abstraction at a remove from the original, primary record. Effectively, therefore, metadata may travel between being metadata and being data – what is metadata to one may be data to another (for example Huggett 2020:3; Buckland et al. 2022:19).

Of course, data are always collected and abstracted using criteria according to a specific frame of reference and this affects its capacity for knowledge creation. Metadata is no different, but its role in infrastructural data retrieval and interoperability places it in a different relationship with archaeological knowledge creation since the ideologies, politics, and perspectives that define the metadata influence the data located and the connections made in the first place. Metadata therefore increasingly govern what can be found and what can be known (Börjesson et al. 2020:207–208). The structuring imposed by metadata carries the biases and worldviews of the infrastructures that created them, and profoundly impact the meaning that can be derived from the data (Canning et al. 2022:12).

INTERFACE

If metadata reveals and limits the data that can be presented, the interface through which most users will experience the digital infrastructure is equally capable of inclusion and exclusion (Hookway 2014:4). Knowing how an interface structures our relation to data is essential since it is designed to function more or less invisibly, but successful invisibility also tends to hide its affordances. Like the infrastructure itself, it allows certain behaviours and actions to occur (Drucker 2013, para. 31). The interface acts as 'cognitive scaffolding' (Dieter 2015:170), empowering the user, but at the same time is a 'device of capture' (Dieter 2015:173), determining pathways and reducing autonomy. The interface, like the infrastructure, is not an object as much as a 'dynamic, systematized relation' (Dieter 2022:5). Like standards and metadata, interfaces are also abstractions, sitting atop a complex system and exposing some of that system's logic while hiding others. The effect of this abstraction is to distance the user from the underlying system: at the same time as the interface facilitates discovery and provides access to data, its underlying design and implementation shape what is revealed or hidden.

For example, the ADS search interface adopts a 'point and browse' strategy rather than a Google-like 'type and hope' approach, enabling the million plus metadata records to be swiftly reduced to a small, relevant subset (Richards et al. 2011:35). However, the underlying search methodology is hidden from the user despite what otherwise seems to be a transparent interface: it invisibly employs a fuzzy search despite the appearance of using a specific, constrained classification, which may give rise to initial doubts about the validity of the output (Huggett 2022:272). The search interface does not allow control of the Boolean search criteria used, and so the infrastructure constrains user action.

Burdick (2015:31) has described a series of attributes and qualities that seek to address such infrastructural restrictions. For example, she proposes the opening of the interface black box to make the underlying operations visible, and ideally alterable. She also argues for making multiple world views available, with the interface configurable using different ontologies rather than the default infrastructural perspective. Ambitiously, she also proposes that the interface should be capable of viewing and manipulating data in an infinite number of ways. For instance, current interfaces are predicated on text for data retrieval, which may not be the most appropriate method. As Bugaje and Chowdhury (2018:258; see also Bugaje & Chowdhury 2017) suggest, data is not read so much as visualised, combined or manipulated, and an interface which reflects this would be more natural and certainly more flexible. Addressing these and other design aspirations would help to support a more sophisticated engagement between infrastructures and knowledge creation.

Conclusion

According to Star and Bowker (2006:231),

> Something that was once an object of development and design becomes sunk into infrastructure over time. Therefore a historical, archaeological approach to the development of an infrastructure … needs complementary sociological, regulatory and technical studies.

This paper seeks both to start the debate and to set the stage for such extensive studies in relation to archaeological infrastructures. In doing so, it has largely focused on large-scale data archive infrastructures, but many of the issues discussed are equally relevant at other scales of infrastructure, from data management and publication systems (for example Open Context, ARCHES), to field recording systems (for example FAIMS, ARK, Intrasis), down to the level of the database (for example Burns & Wark

2020), and the different scales are interwoven in complex ways. In all cases, there have been discussions surrounding these infrastructures, but they have been rather piecemeal and consequently lack a clear overview of the range of questions and concerns encountered. Most contributions are by those who might be described as advocates for the infrastructure, reporting on technical details of implementation and application, but only lightly touching upon aspects such as the infrastructural influence on practice, its positive and negative effects, successes and failures. It is crucial, therefore, that such debates engage those external to the immediate context of infrastructure development and implementation, to offset the influence of advocacy perspectives and technical determinism. Furthermore, while even long-standing archaeological infrastructures are still in-the-making, others may never be started, or are delayed, or abandoned, and these are the norm rather than the exception (Carse & Kneas 2019:9). Archaeology has seen dozens of digital infrastructure-related developments over the years, few of which become established in practice, as evidenced in the Computer Applications in Archaeology (CAA) conference proceedings, for example. What happened to them and why have they disappeared? Which factors determined success or failure?

Part of the attraction of infrastructures lies in their combinatorial possibilities: the way in which each digital object made possible via the infrastructure may be combined and recombined with others, to create new objects and novel innovations (Baiyere et al. 2023:8–9). The expansion of these infrastructures into the corners of archaeological practice makes it important to understand their emergence, their development, their environment, their relationships, their social and cultural elements, their implications for practice and their unanticipated outcomes, as well as their benefits. Given the ways in which infrastructures infiltrate and influence, empower and constrain archaeological practice and thought, it is crucial to develop critical and extensive overviews rather than more of the fragmentary approaches adopted to date. A broader and deeper understanding of archaeological infrastructures today will also ensure that lessons from the past and present will carry forward into future developments.

Declaration of interest

The author was one of the original consortium members that saw the creation of the Archaeology Data Service in 1996, and subsequently served as chair of the management and advisory committees at various points in the intervening years. He is currently a vice-chair of the ADS management committee.

References

ADS 1997. *ADS Annual Report 1996–1997, Archaeology Data Service Annual Reports*. Archaeology Data Service, doi:10.5284/1098084.

ADS 1998. *ADS Annual Report 1997–1998, Archaeology Data Service Annual Reports*. Archaeology Data Service, doi:10.5284/1098085.

ADS 2021. *ADS Annual Report 2020–2021, Archaeology Data Service Annual Reports*. Archaeology Data Service, doi:10.5284/1098108.

Appel, H., Anand, N. & Gupta, A. 2018. Temporality, Politics, and the Promise of Infrastructure. In: Anand, N., Gupta, A. & Appel, H. (eds), *The Promise of Infrastructure*, pp. 1–38. Durham: Duke University Press.

Austin, T., Pinto, F., Richards, J. & Ryan, N. 2002. Joined up Writing: An Internet Portal for Research into the Historic Environment. In: Burenhult, G. & Arvidsson, J. (eds), *Archaeological Informatics: Pushing the Envelope. CAA2001. Computer Applications and Quantitative Methods in Archaeology. Proceedings of the 29th Conference, Gotland, April 2001*, pp. 243–252. Oxford: Archaeopress.

Baiyere, A., Grover, V., Lyytinen, K.J., Woerner, S. & Gupta, A. 2023. Digital 'x' – Charting a Path for Digital-Themed Research. *Information Systems Research*, doi:10.1287/isre.2022.1186.

Barford, P. 2020. Some Aspects of the Collection-Driven Exploitation of the Archaeological Record in England and Wales. *Revista d'Arqueologia de Ponent*. Vol. 30 pp. 101–125, doi:10.21001/rap.2020.30.5.

Beagrie, N. & Houghton, J. 2013. *The Value and Impact of the Archaeology Data Service: A study and Methods for Enhancing Sustainability*. London: Joint Information Systems Committee (JISC), https://repository.jisc.ac.uk/5509/1/ADSReport_final.pdf [Accessed 17 October 2023].

Benardou, A., Champion, E., Dallas, C. & Hughes, L.M. 2017. Introduction: A Critique of Digital Practices and Research Infrastructures. In: Benardou, A., Champion, E., Dallas, C. & Hughes, L.M. (eds), *Cultural Heritage Infrastructures in Digital Humanities*, pp. 1–14. Abingdon, Oxon: Routledge.

Bender, E., Gebru, T., McMillan-Major, A. & Shmitchell, S. 2021. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21)*, pp. 610–623. New York, NY: Association for Computing Machinery, doi:10.1145/3442188.3445922.

Boellstorff, T. 2013. Making Big Data, in Theory. *First Monday*. Vol. 18(10), doi:10.5210/fm.v18i10.4869.

Borgman, C.L. 2015. *Big Data, Little Data, No Data: Scholarship in the Networked World*. Cambridge, Mass: MIT Press.

Borgman, C.L., Scharnhorst, A. & Golshan, M.S. 2019. Digital Data Archives as Knowledge Infrastructures: Mediating Data Sharing and Reuse. *Journal of the Association for Information Science and Technology*. Vol. 70(8) pp. 888–904, doi:10.1002/asi.24172.

Börjesson, L. 2021. Legacy in the Making: A Knowledge Infrastructural Perspective on Systems for Archeological Information Sharing. *Open Archaeology*. Vol. 7(1) pp. 1636–1647, doi:10.1515/opar-2020-0213.

Börjesson, L. & Huvila, I. 2018. Digital Archaeological Data for Future Knowledge-Making. In: Huvila, I. (ed), *Archaeology and Archaeological Information in the Digital Society*, pp. 14–36. Abingdon, Oxon: Routledge.

Börjesson, L., Sköld, O. & Huvila, I. 2020. Paradata in Documentation Standards and Recommendations for Digital Archaeological Visualisations. *Digital Culture & Society*. Vol. 6(2) pp. 191–220, doi:10.14361/dcs-2020-0210.

Bowker, G.C. 2017. How Knowledge Infrastructures Learn. In: Harvey, P., Jensen, C.B. & Morita, A. (eds), *Infrastructures and Social Complexity: A Companion*, pp. 391–403. New York: Routledge.

Bowker, G.C. 2018. Sustainable Knowledge Infrastructures. In: Anand, N., Gupta, A. & Appel, H. (eds), *The Promise of Infrastructure*, pp. 203–222. Durham: Duke University Press.

Bowker, G.C., Baker, K., Millerand, F. & Ribes, D. 2010. Toward Information Infrastructure Studies: Ways of Knowing in a Networked Environment. In: Hunsinger, J., Klastrup, L. & Allen, M. (eds), *International Handbook of Internet Research*, pp. 97–117. Dordrecht: Springer, doi:10.1007/978-1-4020-9789-8_5.

Bowker, G.C., Elyachar, J., […] Nucho, J.R. & Pollock, N. 2019. Introduction to Thinking Infrastructures. In: Kornberger, M., Bowker, G.C., […] Nucho, J.R. & Pollock, N. (eds), *Thinking Infrastructures*, pp. 1–13. Bingley, Yorks: Emerald Publishing.

Bowker, G.C. & Star, S.L. 1999. *Sorting Things Out: Classification and its Consequences*. Cambridge, Mass: MIT Press.

Brodie, N. 2020. What is this thing called the PAS? Metal-detecting entanglements in England and Wales. *Revista d'Arqueologia de Ponent*. Vol. 30 pp. 85–100, doi:10.21001/rap.2020.30.4.

Buckland, P., Sjölander, M., von Boer, J., Mähler, R. & Linderholm, J. 2022. The intricate details of using research databases and repositories for environmental archaeology data. *ArcheoLogica Data*. Vol. 2 pp. 15–29, doi:10.13131/UNIPI/2785-0668/A1CC-XT56.

Buckland, P. & Sjölander, M. 2022. Approaches to Research Data Infrastructure for Archaeological Science. In: Watrall, E. & Goldstein, L. (eds), *Digital Heritage and Archaeology in Practice: Data, Ethics, and Professionalism*, pp. 109–134. Gainesville: University Press of Florida.

Bugaje, M. & Chowdhury, G. 2017. Is Data Retrieval Different from Text Retrieval? An Exploratory Study. In: Choemprayong, S., Crestani, F. & Cunningham, S.J. (eds), *Digital Libraries: Data, Information, and Knowledge for Digital Lives*, pp. 97–103. Cham: Springer, doi:10.1007/978-3-319-70232-2_8.

Bugaje, M. & Chowdhury, G. 2018. Data Retrieval = Text Retrieval? In: Chowdhury, G., McLeod, J., Gillet, J. & Willett, P. (eds), *Transforming Digital Worlds*, pp. 253–262. Cham: Springer, doi:10.1007/978-3-319-78105-1_29.

Burdick, A. 2015. Meta! Meta! Meta! A Speculative Design Brief for the Digital Humanities. *Visible Language: The Journal of Visual Communication Research*. Vol. 49(3) pp. 13–33.

Burns, R. & Wark, G. 2020. Where's the Database in Digital Ethnography? Exploring Database Ethnography for Open Data Research. *Qualitative Research*. Vol. 20(5) pp. 598–616, doi:10.1177/1468794119885040.

Canning, E., Brown, S., Roger, S. & Martin, K. 2022. The Power to Structure: Making Meaning from Metadata Through Ontologies. *KULA: Knowledge Creation, Dissemination, and Preservation Studies*. Vol. 6(3) pp. 1–15, doi:10.18357/kula.169.

Carse, A. 2016. The Anthropology of the Built Environment: What Can Environmental Anthropology Learn from Infrastructure Studies (and Vice Versa)? *Engagement*, 17 May, https://aesengagement.wordpress.com/2016/05/17/the-anthropology-of-the-built-environment-what-can-environmental-anthropology-learn-from-infrastructure-studies-and-vice-versa/ [Accessed 20 September 2023].

Carse, A. & Kneas, D. 2019. Unbuilt and Unfinished. *Environment and Society*. Vol. 10(1) pp. 9–28, doi:10.3167/ares.2019.100102.

Collins, E. 2012. The National Data Centres. In: Pryor, G. (ed), *Managing Research Data*, pp. 151–172. London: Facet Publishing.

Dallas, C. 2015. Curating Archaeological Knowledge in the Digital Continuum: From Practice to Infrastructure. *Open Archaeology*. Vol. 1(1) pp. 176–207, doi:10.1515/opar-2015-0011.

Dennis, L.M. 2020. Digital Archaeological Ethics: Successes and Failures in Disciplinary Attention. *Journal of Computer Applications in Archaeology*. Vol. 3(1) pp. 210–218, doi:10.5334/jcaa.24.

Dieter, M. 2015. Dark Patterns: Interface Design, Augmentation and Crisis. In: Berry, D.M. & Dieter, M. (eds), *Postdigital Aesthetics: Art, Computation and Design*, pp. 163–178. London: Palgrave Macmillan, doi:10.1057/9781137437204.

Dieter, M. 2022. Interface Critique at Large. *Convergence: The International Journal of Research into New Media Technologies*, doi:10.1177/13548565221135833.

Drucker, J. 2013. Performative Materiality and Theoretical Approaches to Interface. *Digital Humanities Quarterly*. Vol. 7(1), http://www.digitalhumanities.org/dhq/vol/7/1/000143/000143.html.

Dufton, J.A. 2016. CSS for Success? Some Thoughts on Adapting the Browser-Based Archaeological Recording Kit (ARK) for Mobile Recording. In: Averett, E.W., Gordon, J.M. & Counts, D.B. (eds), *Mobilizing the Past for a Digital Future: The Potential of Digital Archaeology*, pp. 373–397. Grand Forks, ND: The Digital Press at the University of North Dakota.

Dunn, S. 2011. Poor Relatives or Favorite Uncles? Cyberinfrastructure and Web 2.0: A Critical Comparison for Archaeological Research. In: Kansa, E.C., Kansa, S.W. & Watrall, E. (eds), *Archaeology 2.0: New Approaches to Communication and Collaboration*, pp. 96–117. Los Angeles, CA: Cotsen Institute of Archaeology Press, doi:10.2307/j.ctvhhhfgw.

Edwards, P.N. 2003. Infrastructure and modernity: Force, Time, and Social Organization in the History of Sociotechnical Systems. In: Misa, T.J., Brey, P. & Feenberg, A. (eds), *Modernity and Technology*, pp. 185–225. Cambridge, Mass: MIT Press.

Edwards, P.N. 2019. Infrastructuration: On Habits, Norms and Routines as Elements of Infrastructure. In: Kornberger, M., Bowker, G.C., [...] Nucho, J.R. & Pollock, N. (eds), *Thinking Infrastructures*, pp. 355–366. Bingley, Yorks: Emerald Publishing.

Edwards, P.N., Bowker, G.C., Jackson, S.J. & Williams, R. 2009. Introduction: An Agenda for Infrastructure Studies. *Journal of the Association for Information Systems*. Vol. 10(5) pp. 364–374, doi:10.17705/1jais.00200.

Edwards, P.N., Jackson, S.J., Bowker, G.C. & Knobel, C.P. 2007. *Understanding Infrastructure: Dynamics, Tensions, and Design*. Ann Arbor, MI: University of Michigan, https://deepblue.lib.umich.edu/bitstream/handle/2027.42/49353/Understand?sequence=3.

Edwards, P.N., Jackson, S.J., [...] Burton, M. & Calvert, S. 2013. *Knowledge Infrastructures: Intellectual Frameworks and Research Challenges*. Ann Arbor, MI: University of Michigan, http://hdl.handle.net/2027.42/97552.

Eiteljorg II, H. 1995. The Archaeological Data Archive Project. In: Huggett, J. & Ryan, N. (eds), *CAA94. Computer Applications and Quantitative Methods in Archaeology 1994*, pp. 245–248. Oxford: Tempus Reparatum, https://proceedings.caaconference.org/paper/38_eiteljorg-ii_caa_1994/.

Eiteljorg II, H. 2001 Archaeological Data Archive Pilot Project. *CSA Newsletter*. Vol. 14(2), https://csanet.org/newsletter/fall01/nlf0101.html.

Eiteljorg II, H. 2002. The Archaeological Data Archive Project Ceases Operation. *CSA Newsletter*. Vol. 15(2), https://csanet.org/newsletter/fall02/nlf0201.html.

Eiteljorg II, H. 2011. What are our Critical Data-Preservation Needs? In: Kansa, E.C., Kansa, S.W. & Watrall, E. (eds), *Archaeology 2.0: New Approaches to Communication and Collaboration*, pp. 251–264. Los Angeles, CA: Cotsen Institute of Archaeology Press, doi:10.2307/j.ctvhhhfgw.

Fullilove, C. & Alimari, A. 2023. Baladi Seeds in the oPt: Populations as Objects of Preservation and Units of Analysis. In: Williamson, H.F. & Leonelli, S. (eds), *Towards Responsible Plant Data Linkage: Data Challenges for Agricultural Research and Development*, pp. 65–84. Cham: Springer, doi:10.1007/978-3-031-13276-6.

Gaines, S.W. 1981a. Computerized Data Banks in Archaeology: The European Situation. *Computers and the Humanities*. Vol. 15(4) pp. 223–226, doi:10.1007/BF02395373.

Gaines, S.W. 1981b. Preface. In: Gaines, S.W. (ed), *Data Bank Applications in Archaeology*, pp. vii–viii. Tucson, AZ: University of Arizona Press.

Gupta, N. & Devillers, R. 2017. Geographic Visualization in Archaeology. *Journal of Archaeological Method and Theory*. Vol. 24(3) pp. 852–885, doi:10.1007/s10816-016-9298-7.

Gupta, N., Martindale, A., Supernant, K. & Elvidge, M. 2023. The CARE Principles and the Reuse, Sharing, and Curation of Indigenous Data in Canadian Archaeology. *Advances in Archaeological Practice*. Vol. 11(1) pp. 76–89, doi:10.1017/aap.2022.33.

Hacıgüzeller, P., Taylor, J.S. & Perry, S. 2021. On the Emerging Supremacy of Structured Digital Data in Archaeology: A Preliminary Assessment of Information, Knowledge and Wisdom Left Behind. *Open Archaeology*. Vol. 7(1) pp. 1709–1730, doi:10.1515/opar-2020-0220.

Hardman, C. & Richards, J.D. 2013. Digital Curation at the Archeology Data Service. In: Winghart, S. (ed), *Archäologie und Informationssysteme: Vom Umgang mit archäologischen Fachdaten in Denkmalpflege und Forschung*, pp. 76–78. Hameln, DE: Niedersächsisches Landesamt für Denkmalpflege, http://www.landesarchaeologen.de/fileadmin/Dokumente/News/NI_Archaeologie_und_Informationssysteme_Arbeitsheft42.pdf#page=76.

Harvey, P., Jensen, C.B. & Morita, A. 2017. Introduction: Infrastructural Complications. In: Harvey, P., Jensen, C.B. & Morita, A. (eds), *Infrastructures and Social Complexity: A Companion*, pp. 1–22. London: Routledge.

Heine, E.-C. & Meiske, M. 2022. Scientific Bonanzas: Infrastructures as Places of Knowledge Production. In: Heine, E.-C. & Meiske, M. (eds), *Beyond the Lab and the Field: Infrastructures as Places of Knowledge Production Since the Late Nineteenth Century*, pp. 3–20. Pittsburgh: University of Pittsburgh Press.

Hookway, B. 2014. *Interface*. Cambridge, MA: MIT Press.

Huggett, J. 2012. Lost in Information? Ways of Knowing and Modes of Representation in E-Archaeology. *World Archaeology*. Vol. 44(4) pp. 538–552, doi:10.1080/00438243.2012.736274.

Huggett, J. 2020. Capturing the Silences in Digital Archaeological Knowledge. *Information*. Vol. 11(5) p. 278, doi:10.3390/info11050278.

Huggett, J. 2021. Algorithmic Agency and Autonomy in Archaeological Practice. *Open Archaeology*. Vol. 7(1) pp. 417–434, doi:10.1515/opar-2020-0136.

Huggett, J. 2022. Data Legacies, Epistemic Anxieties, and Digital Imaginaries in Archaeology. *Digital*. Vol. 2(2) pp. 267–295, doi:10.3390/digital2020016.

Huvila, I. 2016. 'If we just knew who should do it', or the Social Organization of the Archiving of Archaeology in Sweden. *Information Research*. Vol. 21(2), http://www.informationr.net/ir/21-2/paper713.html.

Huvila, I. 2018. Ecology of Archaeological Information Work. In: Huvila, I. (ed), *Archaeology and Archaeological Information in the Digital Society*, pp. 122–142. Abingdon, Oxon: Routledge.

Huvila, I. 2019a. Learning to Work between Information Infrastructures. *Information Research*. Vol. 24(2), http://www.informationr.net/ir/24-2/paper819.html.

Huvila, I. 2019b. Management of Archaeological Information and Knowledge in Digital Environment. In: Handzic, M. & Carlucci, D. (eds), *Knowledge Management, Arts, and Humanities*, pp. 147–169. Cham: Springer, doi:10.1007/978-3-030-10922-6_8.

Huvila, I. & Huggett, J. 2018. Archaeological Practices, Knowledge Work and Digitalisation. *Journal of Computer Applications in Archaeology*. Vol. 1(1) pp. 88–100, doi:10.5334/jcaa.6.

Jackson, S.J., Edwards, P.N., Bowker, G.C. & Knobel, C.P. 2007. Understanding Infrastructure: History, Heuristics and Cyberinfrastructure Policy. *First Monday*. Vol. 12(6), doi:10.5210/fm.v12i6.1904.

Jakobsson, U., Novák, D., Richards, J.D., Štular, B. & Wright, H. (eds). 2021. Digital Archiving in Archaeology: The State of the Art. *Internet Archaeology*. Vol. 58, https://intarch.ac.uk/journal/issue58/index.html.

Kansa, E.C. 2022. On Infrastructure, Accountability, and Governance in Digital Archaeology. In: Garstki, K. (ed), *Critical Archaeology in the Digital Age: Proceedings of the 12th IEMA Visiting Scholar's Conference*, pp. 141–152. Los Angeles, CA: Cotsen Institute of Archaeology Press.

Karasti, H. & Blomberg, J. 2018. Studying Infrastructuring Ethnographically. *Computer Supported Cooperative Work (CSCW)*. Vol. 27(2) pp. 233–265, doi:10.1007/s10606-017-9296-7.

Kenny, J. & Richards, J.D. 2005. Pathways to a Shared European Information Infrastructure for Cultural Heritage. *Internet Archaeology*. Vol. 18, doi:10.11141/ia.18.6.

Kintigh, K.W. 2006. The Promise and Challenge of Archaeological Data Integration. *American Antiquity*. Vol. 71(3) pp. 567–578.

Kintigh, K.W., Altschul, J.H., [...] Ludäscher, B. & Lynch, C.A. 2015. Cultural Dynamics, Deep Time, and Data: Planning Cyberinfrastructure Investments for Archaeology. *Advances in Archaeological Practice*. Vol. 3(1) pp. 1–15, doi:10.7183/2326-3768.3.1.1.

Kintigh, K.W., Spielmann, K.A., Brin, A., Candan, K.S., Clark, T.C. & Peeples, M. 2018. Data Integration in the Service of Synthetic Research. *Advances in Archaeological Practice*. Vol. 6(1) pp. 30–41, doi:10.1017/aap.2017.33.

Koch, G. 2018. The Ethnography of Infrastructures. In: Benardou, A., Champion, E., Dallas, C. & Hughes, L.M. (eds), *Cultural Heritage Infrastructures in Digital Humanities*, pp. 63–81. London: Routledge, doi:10.4324/9781315575278.

Larkin, B. 2013. The Politics and Poetics of Infrastructure. *Annual Review of Anthropology*. Vol. 42(1) pp. 327–343, doi:10.1146/annurev-anthro-092412-155522.

Leonelli, S. & Williamson, H.F. 2023. Introduction: Towards Responsible Plant Data Linkage. In: Williamson, H.F. & Leonelli, S. (eds), *Towards Responsible Plant Data Linkage: Data Challenges for Agricultural Research and Development*, pp. 1–24. Cham: Springer, doi:10.1007/978-3-031-13276-6.

Löwenborg, D. 2018. Knowledge Production with Data from Archaeological Excavations. In: Huvila, I. (ed), *Archaeology and Archaeological Information in the Digital Society*, pp. 38–53. Abingdon, Oxon: Routledge.

Lucas, G. 2012. *Understanding the Archaeological Record*. Cambridge: Cambridge University Press.

Marttila, S. & Botero, A. 2017. Infrastructuring for Cultural Commons. *Computer Supported Cooperative Work (CSCW)*. Vol. 26(1) pp. 97–133, doi:10.1007/s10606-017-9273-1.

McManamon, F.P., Kintigh, K.W., Ellison, L.A. & Brin, A. 2017. tDAR: A Cultural Heritage Archive for Twenty-First-Century Public Outreach, Research, and Resource Management. *Advances in Archaeological Practice*. Vol. 5(3) pp. 238–249, doi:10.1017/aap.2017.18.

Meghini, C., Scopigno, R., […] Binding, C. & Vlachidis, A. 2017. ARIADNE: A Research Infrastructure for Archaeology. *Journal on Computing and Cultural Heritage*. Vol. 10(3) art.18, doi:10.1145/3064527.

Meyer, E.T. & Schroeder, R. 2015. *Knowledge Machines: Digital Transformations of the Sciences and Humanities*. Cambridge, MA: MIT Press.

Millerand, F. & Baker, K.S. 2010. Who Are the Users? Who Are the Developers? Webs of Users and Developers in the Development Process of a Technical Standard. *Information Systems Journal*. Vol. 20(2) pp. 137–161, doi:10.1111/j.1365-2575.2009.00338.x.

Millerand, F. & Baker, K.S. 2020. Data Infrastructures in Ecology: An Infrastructure Studies Perspective. *Oxford Research Encyclopedia of Environmental Science*, doi:10.1093/acrefore/9780199389414.013.554.

Mongili, A. & Pellegrino, G. (eds). 2014. *Information Infrastructure(s): Boundaries, Ecologies, Multiplicity*. Newcastle upon Tyne: Cambridge Scholars Publishing.

Niccolucci, F. & Richards, J.D. 2013. ARIADNE: Advanced Research Infrastructures for Archaeological Dataset Networking in Europe. *International Journal of Humanities and Arts Computing*. Vol. 7(1–2) pp. 70–88, doi:10.3366/ijhac.2013.0082.

Nicholson, C., Kansa, S., Gupta, N. & Fernandez, R. 2023. Will It Ever Be FAIR? Making Archaeological Data Findable, Accessible, Interoperable, and Reusable. *Advances in Archaeological Practice*. Vol. 11(1) pp. 63–75, doi:10.1017/aap.2022.40.

Opitz, R., Strawhacker, C., […] Szabo, V. & Thompson, P. 2021. A Lockpick's Guide to dataARC: Designing Infrastructures and Building Communities to Enable Transdisciplinary Research. *Internet Archaeology*. Vol. 56, doi:10.11141/ia.56.15.

Ortman, S.G. & Altschul, J.H. 2023. What North American Archaeology Needs to Take Advantage of the Digital Data Revolution. *Advances in Archaeological Practice*. Vol. 11(1) pp. 90–103, doi:10.1017/aap.2022.42.

Paris, B.S., Cath, C. & West, S.M. 2023. Radical Infrastructure: Building Beyond the Failures of Past Imaginaries for Networked Communication. *New Media & Society*, doi:10.1177/14614448231152546.

Parmiggiani, E. 2017. This Is Not a Fish: On the Scale and Politics of Infrastructure Design Studies. *Computer Supported Cooperative Work (CSCW)*. Vol. 26(1) pp. 205–243, doi:10.1007/s10606-017-9266-0.

Pollock, N. & Williams, R. 2010. E-Infrastructures: How Do We Know and Understand Them? Strategic Ethnography and the Biography of Artefacts. *Computer Supported Cooperative Work (CSCW)*. Vol. 19(6) pp. 521–556, doi:10.1007/s10606-010-9129-4.

Richards, J.D. 1997. Preservation and Re-Use of Digital Data: The Role of the Archaeology Data Service. *Antiquity*. Vol. 71(274) pp. 1057–1059.

Richards, J.D. 2002. Digital Preservation and Access. *European Journal of Archaeology*. Vol. 5(3) pp. 343–366, doi:10.1179/eja.2002.5.3.343.

Richards, J.D. 2021. Archiving Archaeological Data in the United Kingdom. *Internet Archaeology*. Vol. 58, doi:10.11141/ia.58.21.

Richards, J.D., Jakobsson, U., Novák, D., Štular, B. & Wright, H. 2021. Digital Archiving in Archaeology: The State of the Art. Introduction. *Internet Archaeology*. Vol. 58, doi:10.11141/ia.58.23.

Richards, J.D., Jeffrey, S., Waller, S., Ciravegna, F., Chapman, S. & Zhang, Z. 2011. The Archaeology Data Service and the Archaeotools Project: Faceted Classification and Natural Language Processing. In: Kansa, E.C., Kansa, S.W. & Watrall, E. (eds), *Archaeology 2.0: New Approaches to Communication and Collaboration*, pp. 31–56. Los Angeles, CA: Cotsen Institute of Archaeology Press.

Robey, D. 2012. Improving Sustainability of Publicly Funded Digital Resources. In: Hughes, L.M. (ed), *Evaluating and Measuring the Value, Use and Impact of Digital Collections*, pp. 147–156. London: Facet Publishing.

Slota, S. & Bowker, G.C. 2015. On the Value of 'Useless Data': Infrastructures, Biodiversity, and Policy. In: *iConference 2015 Proceedings*. *iConference 2015*, Newport Beach, CA: iSchools, https://hdl.handle.net/2142/73663.

Sobotkova, A., Ross, S., Ballsun-Stanton, B., Fairburn, A., Thompson, J. & VanValkenburgh, P. 2016. Measure Twice, Cut Once: Cooperative Deployment of a Generalized, Archaeology-Specific Field Data Collection System. In: Averett, E.W., Gordon, J.M. & Counts, D.B. (eds), *Mobilizing the Past: The Potential of Digital Archaeology*, pp. 337–371. Grand Forks, ND: University of North Dakota, doi:10.31356/dpb008.

Star, S.L. 1999. The Ethnography of Infrastructure. *American Behavioral Scientist*. Vol. 43(3), pp. 377–391.

Star, S.L. & Bowker, G.C. 2006. How to Infrastructure. In: Lievrouw, L.A. & Livingstone, S.M. (eds), *Handbook of New Media: Social Shaping and Social Consequences of ICTs*. Updated Student Edition, pp. 230–245. London: Sage Publications.

Star, S.L. & Ruhleder, K. 1996. Steps Toward an Ecology of Infrastructure: Design and Access for Large Information Spaces. *Information Systems Research*. Vol. 7(1) pp. 111–134.

Svensson, P. 2015. The Humanistiscope: Exploring the Situatedness of Humanities Infrastructure. In: Svensson, P. & Goldberg, D. (eds), *Between Humanities and the Digital*, pp. 337–353. Cambridge, MA: MIT Press.

Thylstrup, N.B. 2018. *The Politics of Mass Digitization*. Cambridge, MA: MIT Press.

Van Rossem, W. & Pelizza, A. 2022. The Ontology Explorer: A Method to Make Visible Data Infrastructures for Population Management. *Big Data & Society*. Vol. 9(1), doi:10.1177/20539517221104087.

Waller, S. 2005. Future Connections: The Potential of Web service and Portal Technologies for the Historic Environment. *Internet Archaeology*. Vol. 18, doi:10.11141/ia.18.8.

Williamson, H.F. & Leonelli, S. (eds). 2023. *Towards Responsible Plant Data Linkage: Data Challenges for Agricultural Research and Development*. Cham: Springer, doi:10.1007/978-3-031-13276-6.

Witze, A. 2019. Disappearing Digital Data. *American Archaeology*. Vol. 23(1) pp. 40–45.

Wright, H. & Richards, J.D. 2018. Reflections on Collaborative Archaeology and Large-Scale Online Research Infrastructures. *Journal of Field Archaeology*. Vol. 43(sup1) pp. S60–S67, doi:10.1080/00934690.2018.1511960.