The British Academy

Journal of the British Academy, 10: 93–120
https://doi.org/10.5871/jba/010.093
Posted 11 October 2022

Facing the challenge of digital sustainability as humanities researchers

Joanna Tucker

Abstract

Humanities researchers are increasingly united in their concerns about the long-term sustainability of digital resources. Much of our work is now reliant upon the use of resources such as databases, online research tools and digital editions. Libraries and archives are undertaking programmes of 'mass digitisation', making our primary sources available to view as digital images on the web. This article presents a view of the current landscape as well as thoughts for the future of digital research for scholars in the humanities. It considers four aspects of sustainability (technological, financial, environmental and 'human'), and offers a new working definition for digital sustainability in this context. This centres on sustainability not just as a technical concern but as a multifaceted activity within which humanities researchers can play a crucial role.

Keywords

Sustainability, Digital Humanities, environmental sustainability, digitisation, digital editions, databases.

Note on the Author

Dr Joanna Tucker is a Lecturer in History at the University of Glasgow. Her research has focused on approaches to medieval charters and cartulary manuscripts, especially developing a new methodology for analysing scribes and their work in her recent book *Reading and Shaping Medieval Cartularies: Multi-Scribe Manuscripts and their Patterns of Growth* (2020). She has also explored Digital Humanities approaches to manuscripts and texts, including as co-director of http://www.modelsofauthority.ac.uk and as a Co-Investigator on the research network *A Digital Framework for the Medieval Gaelic World* (2020–21).
https://orcid.org/0000-0001-7464-9492 | Joanna.Tucker@glasgow.ac.uk

There is no doubt that 'Digital Humanities' (broadly defined) has had a profound impact on the horizons of scholarly research. At least half of the British Academy's research projects have involved the production of digital resources: these relate to every Humanities Section and most of the Social Sciences Sections.[1] The way that scholars understand Digital Humanities has often been formed by our own individualised experiences with specific digital resources such as these, designed for our particular fields of research. There is now, however, a pressing issue that increasingly concerns all of our work: the question of 'digital sustainability'. This obliges us to step back from our individual experiences and collectively consider the longer-term future for our digital research methods and outputs.

It is probably fair to say that, for many, the issue of 'digital sustainability' is synonymous with anxiety and unease. The term might conjure up thoughts of the inevitable deterioration of websites and digital tools, particularly those which are outcomes of short-term funded research projects. As something inherently technological, digital sustainability is surely far beyond the control of most humanities scholars. Nevertheless, it is progressively becoming a shared concern for researchers in a range of humanities fields, given our development of – and now often reliance upon – digital resources for our research as well as for teaching and public engagement.

Researchers can be comforted by the fact that many people outside of the academy are also concerned about digital sustainability. In 2015, Google's Vint Cerf warned of the 21st century becoming a 'digital Dark Age' in the future if we do not take steps to prevent the obsolescence of our hardware and software.[2] For those of us who work on the Middle Ages, it is notable that Cerf described his conceptual solution as 'digital vellum', given vellum's proven longevity. This serves as a humbling reminder of the relatively robust and sustainable information technology inherited from the medieval period in the form of parchment manuscripts, and also other material culture.

The question of sustainability is also not an especially new one in Digital Humanities (DH). It does, however, seem to be all the more pressing now. It should be recognised that this is largely due to the success of so many DH projects and DH research generally, especially in widening out to non-DH specialists who have found that digital resources are now embedded in how we research, how we teach, how we access materials, and even how we expect to view and interact with our sources. It might be assumed that 2020's sudden 'online pivot' as a result of Covid-19 restrictions has accelerated this, embedding our reliance on digital ways of working and communicating. The longer-term success of DH is also a reflection of the changing nature of humanities disciplines themselves, especially the growth of cross-disciplinary collaborations between humanities researchers and software engineers.

There is a growing need for designers, developers, users and hosts of digital research resources to think collectively about sustainability. Long-term sustainability has been a concern among DH scholars for as long as the field itself has existed. As early as 2004, the Blackwell *Companion to Digital Humanities* had a section devoted to 'Production, dissemination, archiving' which included chapters on 'Designing sustainable projects and publications' and 'Preservation'.[3] More recently, those active in DH projects have begun to address sustainability explicitly, particularly in articles in DH journals.[4] Individual DH research projects also engage with the topic and sometimes publish

[1] British Academy, 'Academy Research Projects'. https://www.thebritishacademy.ac.uk/programmes/academy-research-projects/ [accessed 3 October 2021].
[2] Ghosh (2015).
[3] Pitti (2004) and Smith (2004).
[4] For example, Edmond & Morselli (2020); Barats *et al.* (2020); and Drucker (2021).

their approach.[5] Another context for work on sustainability is libraries and archives where there is a more longstanding literature, given the duty of these institutions to provide continuous access to their holdings, whether print or digital.[6] Funders have also been engaging with this issue: for example, in 2006 the Arts and Humanities Research Council's ICT Methods Network published a report on digital sustainability, surveying the current hurdles and some possible steps for overcoming them.[7]

It is less common for humanities researchers to be leading the conversation on digital sustainability. A recent exception was a symposium hosted by the University of Glasgow in June 2021 as part of a series in the AHRC- and IRC-funded network *A Digital Framework for the Medieval Gaelic World* (August 2020 – September 2021). This brought together scholars of history, language, literature and archaeology as well as archivists, librarians, digital scientists and DH specialists from across Britain and Ireland to consider the current state of the field. The discussions from this symposium led to a report and series of recommendations.[8] The present article develops the issues arising from this symposium more widely, drawing on examples from the medieval Gaelic world in order to reach a new understanding of digital sustainability from a humanities perspective, rather than solely from a collections or DH point of view. It also looks at the topic in relation to the whole field, rather than reflections from a single research project. The conversation is relevant for all humanities researchers since many digital resources relating to the medieval Gaelic world are of a kind that is replicated across the humanities, such as searchable dictionaries and catalogues, collections of images, critical editions, or more bespoke digital research tools such as 'factoid prosopographies'.[9] Because of this widespread range of resources, there is now a particular urgency to the general issue of long-term access in the digital sphere.

Throughout the course of the network, it became clear that a fundamental obstacle can be our own understanding of what 'sustainability of digital resources' actually means. It is worth considering, therefore, whether we, in the humanities, need to develop a new approach to what sustainability means in practice. This article attempts to address this. In doing so, it seeks to raise awareness of, and offer reassurances about, sustainability for scholars in the humanities who work with digital resources. It will address four aspects in turn: technological, financial, environmental and 'human' sustainability.[10] Though it makes no claim to be comprehensive in any of these areas, it will bring together some collective insights about the present and future of digital scholarship. It will also suggest a new working understanding of digital sustainability which centres on seeing this not just as a technical concern but as a wide range of activities within which humanities researchers can play a crucial role.

[5] Blaney & Winters (2010) is an example of a discussion of sustainability in relation to a particular institutional project (the British History Online Digital Library) led by the Institute for Historical Research (IHR).
[6] For example, Bradley (2007); Chowdhury (2014); Varela (2016); and Eschenfelder *et al.* (2019).
[7] AHRC ICT Methods Network (2006).
[8] Stifter *et al.* (2022).
[9] Examples of such resources which have been funded by the British Academy can be found here: https://www.thebritishacademy.ac.uk/programmes/academy-research-projects/ [accessed 3 October 2021].
[10] It is common to break down sustainability into essential areas such as these, although writers tend to differ on how they divide it up. For example, sustainability is divided into economic, social and environmental in Chowdhury (2014). The AHRC's ICT Methods Network report (2006) distinguishes between a scholarly digital resource's academic sustainability (i.e., the content) and its technical sustainability (i.e., the platform).

Case studies of digital sustainability in the medieval Gaelic world: past, present and future

DH research projects have resulted in a wide range of tools, applications, methodologies and resources, including databases of people or images, digital editions, interactive reference works, digital archives, online dictionaries and glossaries, and much else besides. Before unpacking the concept of sustainability, it will be useful to begin by surveying some concrete examples of digital research resources. Illustrations will be drawn from the medieval Gaelic world – past, present and future – with a particular eye on their digital sustainability prospects. This will allow us to take stock of the current situation and consider what sustainability can look like 'in practice'.

First, let us look at three examples of well established digital projects and how they have managed their own sustainability. The *People of Medieval Scotland* (PoMS)[11] was the result of a series of funded projects beginning in 2007. Such continuity of funding is rare, and is not itself a long-term solution to sustainability. PoMS has recently entered a new phase with its data converted by John Bradley into a form of 'linked data' known as 'Resource Description Framework' (RDF) triples.[12] This represents a more open, more interoperable, and therefore more sustainable form of accessible data than PoMS's current database. A second example of success is the *electronic Dictionary of the Irish Language* (eDIL),[13] another longstanding digital project, beginning in 2003. Part of its success lies in the fact that the dictionary is based on sustainable technologies like TEI XML. The website is currently hosted on a server at Queen's University Belfast on an ongoing basis, while the data itself is stored in GitHub. A third example of a successful, large-scale digital resource is the *Corpus of Electronic Texts* (CELT),[14] begun in 1997 and now hosted by University College Cork. CELT's ongoing existence is arguably not just a result of its hosting and technical standards, significant though these factors are. The impetus for CELT's sustainability has also come down to its fundamental role in making texts available and accessible, allowing users to download and re-use them. In other words, the function of the resource itself and its central place in research and teaching relating to the medieval Gaelic world has contributed to its long-term sustainability. The journey has not been smooth for any of these projects, and none of their paths have been predictable. All have, however, achieved a level of sustainability which should hopefully save them and their contents from the 'digital graveyard' for some time to come.

For other older projects the path to sustainability is often less clear, not least because there can be significant challenges for updating. This has been the case, for example, for *Monasticon Hibernicum*,[15] and for *Thesaurus Linguae Hibernicae* (TLH)[16] which has only recently been brought back online after some risks to its integrity and operability were tackled. For such projects, a range of steps might need to be taken to ensure security and usability and to generally bring the resource up to standard: for example, uploading data to a secure storage space (such as GitHub); enhancing the website's basic look and optimising the interface for mobile devices; allowing for records and search results to be saved or exported (as a CSV file, for example); or updating

[11] *People of Medieval Scotland 1093-1371*. https://www.poms.ac.uk
[12] *People of Medieval Scotland: RDF Server.* https://www.poms.ac.uk/rdf
[13] *eDIL – Electronic Dictionary of the Irish Language.* https://dil.ie
[14] *CELT: Corpus of Electronic Texts*. https://celt.ucc.ie
[15] *Early Christian Ecclesiastical Settlement in Ireland 5th to 12th Centuries: Database of the Monasticon Hibernicum Project*. https://monasticon.celt.dias.ie
[16] *Thesaurus Linguae Hibernicae*. https://www.ucd.ie/tlh/index.html

copyright restrictions. Post-factum updates such as these can be difficult to manage and fund. Where they are possible, however, they do have the advantage of being able to draw on real user experiences, responding to the resource's actual use beyond the original project itself.

For new projects, planning for sustainability is now much more of a concern in the development process. Recent digital projects in the field of Gaelic studies include Pádraic Moran's *Gloss Corpus*[17] and his recently launched *Manuscripts with Irish Associations* (MIrA).[18] Another example is *Chronologicon Hibernicum* (ChronHib), which resulted in the creation of the lexicographic database *Corpus PalaeoHibernicum* (CorPH).[19] While there are no straightforward solutions, the creators in each case have been mindful of sustainability concerns to a greater extent than was likely in earlier projects, especially in the planning stages.

Many more examples could be highlighted here. In each instance, 'sustainability' means something slightly different in practice and is certainly not something that can be easily predicted. What, then, should we take to be the overarching definition of 'digital sustainability'?

What is digital sustainability?

In facing the concerns about digital sustainability, it is worth first unpacking the concept. A few dimensions can be highlighted. First, digital sustainability is, of course, inherently about technological decisions. This can include the choice of software and infrastructure, the data models and methods employed, the interface and its presentation on different digital devices, the long-term access and storage of the data, as well as ongoing security risks. Such issues are matched in their significance by financial sustainability, especially the challenge of how to continuously access funding for updates for the maintenance of infrastructure, fixing bugs and responding to other emerging needs. Increasingly, there are also alarming concerns about environmental sustainability, particularly because of the continuous energy required for digital infrastructures. This can often feel remote for anyone planning a DH research project but is nevertheless a very real concern. A final aspect is 'human' sustainability – investment in the people that create, and also those who use, the resource. It is, after all, people who enable the growth of any resource in response to its use over time and support its community of users in an ongoing way. It is also people who promote the resource and train new potential users, whether students or the public or scholars in other fields. All of this is central to the continuous support of digital resources and is therefore key for achieving real sustainability. To put it crudely, sustainability is not something abstract which simply 'happens'. Sustainability requires human resources which also need to be sustained and nurtured. Sustainability is therefore best regarded as a range of activities with people at its heart, not simply a 'state' that can be achieved by any single digital resource.

Many of those who have been involved in DH projects will have direct experience of planning for sustainability. A particularly significant example is the Collaborative European Digital Archival Research Infrastructure (CENDARI),[20] where sustainability is seen 'as a process, rather than a state', with 'transformation and reuse' as the project's end goal, building on the 'know-how and

[17] *Gloss Corpus*. http://www.glossing.org/glosscorpus/
[18] *Manuscripts with Irish Associations (MIrA)*. http://www.mira.ie
[19] *Chronologicon Hibernicum (ChronHib)*. https://chronhib.maynoothuniversity.ie/chronhibWebsite/home
[20] *Collaborative European Digital Archival Infrastructure (CENDARI)*. http://www.cendari.eu/

expertise … and the community of users' gained during the project's development.[21] For many others, however, it is all too easy to despair at sustainability as a seemingly technological problem that can only be overcome with large pots of money. Perhaps it is time, therefore, that we collectively recognise sustainability as representing something that includes the activity of ongoing use and promotion of digital resources by individuals and communities and is therefore much wider than just technological infrastructure. To fully understand all of the dimensions that can be at play, and how humanities scholars might engage with them, let us look at each of the four areas outlined above in turn.

Technological aspects of digital sustainability

Discussions of digital sustainability are often framed in terms of technology. While this is entirely natural, it can mean that humanities researchers become quickly isolated from the conversation and remain uninformed about the crucial issues. It is important, however, to become familiar with the emerging areas of concern as well as opportunities for the future, some of which will be discussed here.[22]

DH scholarship now routinely talks of research materials as 'data'. For some researchers, this is a well established mentality, part of an acquired 'digital literacy'. One context in which this has become embedded is in the funding application's requirement for a 'data management plan' for all research projects. For others, however, this represents a relatively unfamiliar way of thinking about the work that we produce or its perceived uses and value.[23] It can therefore present an initial hurdle for those seeking to engage with wider discussions of digital sustainability. Nevertheless, there are contexts where this language can be helpful. The early stages of planning a digital research project, for example, ought to begin with a frank conversation between the humanities researchers and the digital development team about the nature of the project's resulting 'data', and how this will be stored and accessed beyond the project's funded lifetime.[24] Rather than limiting the project's aspirations, this can in fact be an opportunity to refine what really lies at the core of the research, and therefore what needs to last beyond the project itself, all of which are humanities-led questions. Such discussions will typically result in a detailed data management plan.

From a sustainability perspective, the gold standard, we are often now told, is data that will be 'interoperable', able to be reused in future research projects. For this to happen, it should be 'machine readable' but also comprehensible to humans. While discussions of interoperability might seem to be highly technical, it is becoming clear that for this prospect to be realised the data's 'quality' is just as important as its format. This includes how the data has been structured and 'marked up' (or interpreted), all of which are decisions that rest with humanities scholars. It has also been generally acknowledged that descriptive metadata and documentation are also fundamental to 'good quality data', to help future readers/users to understand the context in which that

[21] Edmond & Morselli (2020: 1024).
[22] A useful place to get a flavour of the kinds of technical issues posed by sustainability is to look at the work of *The Endings Project* which aims to help projects plan for sustainability from the outset: *The Endings Project. Building Sustainable Digital Humanities Projects*. https://endings.uvic.ca/index.html [accessed 2 July 2021]. For a brief project report, see Carlin (2018).
[23] Posner (2015).
[24] Smithies *et al.* (2019: paragraph 26).

data was initially produced or collected.[25] This is one of the reasons that humanities researchers must work closely alongside those developers producing a web resource in order to ensure that the information is well 'contextualised'.[26] Quality control is therefore a crucial part of making our data more sustainable in the long term.

The future 'reuse' of data is certainly an emerging theme in some areas of DH. James Cummings, for example, has recently said of digital scholarly editions that:

> … the most interesting repurposing of any digital edition will likely not be done by the original creators. In other words, assuming the survival of a well-documented edition's data into the distant future, the edition (or the data) is much more likely to be repurposed as technology develops and exploited in ways which we could not have predicted. And yet, current reuse of digital edition data by others is very rare.[27]

While this basic idea of reuse is well established in, for example, an archives and library setting – where reuse, and the enhancement of metadata to allow for reuse, is at the core of what these institutions do – for humanities scholars the production of reusable data could represent a significant shift in mindset.[28] It moves away from the idea of research data as unique and personal to an individual (which is traditionally what researchers are rewarded for), and instead moves the discipline towards a more explicitly open and collaborative scholarly method, where the fruits of research are not just packaged into articles or books. While there is nothing new about collaborative scholarly research, it is arguable whether we have truly developed a language of the 'authorship' of research 'data', and whether it is viewed on an equal par with traditional outputs.[29]

There is certainly further to go in terms of truly interoperable DH research tools. Websites may include links to other websites, but this is perhaps a light-touch version of what is envisioned by 'interoperability', where an existing dataset can truly be harnessed to further the research questions of a new project.[30] It might be assumed that this is because of technical limitations – too many digital resources have been written in a bespoke way that does not allow for the kind of interoperability that might be possible had they all been coded in the same language or using the same framework. The question should also be asked, however, from a humanities perspective. How crucial is it for our own individual research questions that we are able to harness the data from

[25] Barats *et al.* (2020: paragraphs 19 and 28).

[26] A good example of 'contextualised' data relating to medieval manuscripts research is Bill Endres' publication of digital images of the St Chad Gospels and Lichfield Cathedral's Wycliffe New Testament (with the permission of the Chapter of Lichfield Cathedral under a Creative Commons non-commercial license), which are now freely available via the Open Science Framework website: 'St Chad Gospels (8c) – CC images (University of Oklahama)'. https://osf.io/e9dyt/ [accessed 4 October 2021].

[27] Cummings (2019: 183).

[28] See, for example, the comments in Kálmán *et al.* (2019: 114): 'In the arts and humanities in particular, however, data sharing and reuse among researchers is not a commonly established practice. … Even if there were an increased awareness of the need and benefit of sharing resources within the disciplines of the arts and humanities, much needs to be done to make it an integral part of an everyday research practice.'

[29] Guidelines for describing the role of individuals in relation to research were published by the Consortia Advancing Standards in Research Administration Information (CASRAI). 'CRediT' (Contributor Roles Taxonomy: https://credit.niso.org/) includes as one of its roles 'data curation'. This framework is not, however, routinely used to describe roles in humanities research projects.

[30] For this point in relation to digital editions, see the comment in Turska *et al.* (2016: paragraph 1): 'The number of publicly accessible digital editions is constantly growing, but only a relatively small percentage of them make their encoded source files openly available. Without the sources we cannot hope for the much-anticipated and commonly advertised re-use of all this painstakingly collected and prepared content in innovative research, visualization, and popularization.'

previous projects? Is this always a desirable aim for research projects, which tend to address immediate and nuanced questions of their own? Do we as researchers inherently prefer to start with a fresh slate (if that is ever possible), rather than inheriting the nuances or 'biases' from previous research projects? And are these nuances in the data not in fact what make our research valuable (and fundable) in the first place? Or should we be asking new kinds of questions which can take full advantage of the available datasets? Of course, interoperability is not black and white – it is possible to build a nuanced resource that is also interoperable. Nevertheless, it is important to ask these questions when envisaging the long-term aims of our digital research outputs and their potential relationship with other resources and research.

Let us consider another aspect of digital research data. Datasets are not usually published 'raw' online. Typically, they are mediated through a web interface of some kind. Such 'front end' interfaces or applications can provide context to a dataset, but they can also be complex and costly to design and maintain.[31] Moreover, interfaces might even be said to limit the way that users can 'view' the data.[32] Sustainability might therefore seem to be more achievable if we dispense with those interfaces and applications which require regular upkeep. An example that is moving in this direction is PoMS, where the funding for the regular maintenance of the public interface[33] with its search and browse facility is due to come to an end on 7 January 2028, but the data itself has been converted into an RDF format allowing for greater longevity – if, of course, users are able to learn how to 'query' the database of 'triples' using a particular querying language known as 'SPARQL'.[34] For many resources, however, the data and the interface cannot be so easily separated. An example of this is eDIL where the interface, with its bespoke search engine, is key to the data's functionality for users. The relationship between the interface and the underlying 'data' is therefore something to be carefully considered.

A range of guidance has recently been developed in order to aid researchers approaching the technological sustainability of their projects. One of the best cross-disciplinary examples is the FAIR principles. These are 'Guiding Principles for scientific data management and stewardship' that aim to 'improve the Findability, Accessibility, Interoperability, and Reuse of digital assets'.[35] 'Scientific' here is meant to be all encompassing, though the initiative did initially grow out of concerns within the STEM community specifically. As the authors of the principles acknowledge, these are not intended to be prescriptive measures but rather high-level considerations aimed at helping 'data publishers' to evaluate their choices at the outset of an 'investigation'.[36] The guidance itself is rather technical, but it is important for humanities researchers to also themselves

[31] Regular upkeep of a web application might entail, for example, updating the application's code to the latest version of the 'software framework' (e.g., a JavaScript framework) that it uses. These sorts of updates require the time of a software developer, and are best done little and often rather than in large batches in order to keep pace with framework updates as and when they are published (which can even, in the extreme, be multiple times a year). Not doing so increases security risks and therefore potentially incurs even greater costs down the line.

[32] In relation to digital editions, it has recently been argued that the encoded texts are the core of the edition but that 'far too much discussion of digital editions focuses on the presentation views of the edition': Cummings (2019: 181). See also Turska *et al.* (2016: paragraph 4): 'in digital editions the encoded texts themselves are the most important long-term outcome of the project, while their initial presentation within a particular application should be considered only a single perspective on the data.'

[33] PoMS website. https://www.poms.ac.uk

[34] For examples of what these SPARQL queries look like, see *PoMS: RDF Services Documentation*, 'SPARQL examples'. https://www.poms.ac.uk/rdf/doc/sparql.html [accessed 23 September 2021].

[35] *FAIR Principles*. https://www.go-fair.org/fair-principles/

[36] Wilkinson *et al.* (2016: 5).

reflect on what makes a digital resource or a dataset 'FAIR' (Findable, Accessible, Interoperable and Reusable) within their field. Some general observations might be offered here.

One potentially important principle of sustainable design in DH is simplicity. Most projects begin life with significant aspirations (partly, of course, to satisfy funder criteria). Minimalism, however, has a real long-term value that should not be overlooked.[37] Simple architecture might involve 'modularity' – that is, the separation of functional elements such as the content itself (the data), the user interface, and the code behind this which speaks to the server.[38] This modularity can be crucial for enabling sustainability since elements can be updated independently. It is also important, however, not to simplify a design to the point that a digital resource loses its potential richness or functionality. Indeed, digital resources are often most useful when they are dynamic, searchable and even customisable for their users. Such features can require more bespoke solutions, in addition to well planned mark-up and complex search engines. Indeed, it has been said that: 'In some projects the need for a complex platform might outweigh the need for long term preservation.'[39] While the choice between complexity and sustainability need not be a binary one, it is clear that simplicity of design can be a beneficial principle for users and creators alike. For Edmond & Morselli, simplicity of design in a DH project includes clarity of purpose, having a single point of communication, and simplicity of access.[40]

Second, in order to become FAIR and sustainable for users, a resource must also be readily citable – not just overall but each individual web page (much like the pages of a book or article). A citable resource can more easily engage with wider print scholarship in the longer term. This is something that website creators can achieve by offering users clear guidance for citation. To support this practice, there are now strong recommendations for permanent digital identifiers.[41] For example, a 'Persistent Identifier' is a form of identifier that is 'globally unique', is 'actionable' (it takes a user to the resource or information about it), and is 'managed' to enable a long-lived link to the object or information about it.[42] 'PIDs' are now increasingly used by Cultural Heritage Organisations, notably libraries with large digital holdings. They can be used for a range of items such as publications, manuscripts, people, subjects or organisations (familiar examples of PIDs include DOIs or ORCID iDs for researchers). PIDs are now viewed not only as a significant aspect of data management but also for enabling data reuse (and therefore facilitating interoperability with other resources), making them a crucial part of any resource's long-term sustainability. Libraries are key to supporting this activity, given their expertise with this form of data management.

Persistent Identifiers also have an intrinsic value for research practice in the digital age. Take the example of the digitised manuscript. Normal conventions would dictate that any reference in published work would be to the manuscript's shelf mark (e.g., London, British Library Harley MS 5280). This is a unique and stable reference. Where a researcher has only viewed a digital

[37] There is now a movement towards 'minimal computing' in DH more generally: see the 'Minimal computing' working group which is part of *Global Outlook::Digital Humanities*. http://www.globaloutlookdh.org/minimal-computing/ [accessed 5 July 2021].

[38] For a discussion of modularity (especially in terms of multiple and separate project outcomes), see Brown *et al.* (2009: paragraphs 7–15). Edmond & Morselli (2020: 1027) also discuss the importance of modularity for ensuring sustainability of a DH project.

[39] Goddard & Seeman (2020: 42).

[40] Edmond & Morselli (2020: 1028).

[41] For more technical details on making digital identifiers permanent, see the advice from the W3C: 'Data on the Web Best Practices, 8.7 Data Identifiers'. https://www.w3.org/TR/dwbp/#DataIdentifiers [accessed 24 September 2021].

[42] Madden & Kotarski (2020: 3).

reproduction, however, best practice would suggest that this is made explicit.[43] Because PIDs can be permanently tied to a specific digital object, a researcher can now use this to refer to the digital images themselves – e.g., using the 'Archival Resource Key' for the same manuscript:[44]

http://access.bl.uk/item/viewer/ark:/81055/vdc_100123802477.0x000001

This explicitly reinforces the distinction between the digital images and the original manuscript, and importantly allows for transparency about which is being cited by the writer – a distinction which is not always clear in current research conventions.

Another important aspect of being citable is ensuring that all those individuals involved, and their efforts, are appropriately acknowledged on the website or within the data itself. This is particularly important for projects which are collaborative, as is the case for almost all DH projects. It also speaks to wider concerns to ensure that contributions to digital resources are recognised as an important form of research in their own right. This feeds into the notion of sustaining not just the resource but also the careers of those involved.[45]

Another consideration when applying the FAIR principles in order to achieve a sustainable resource is the choice of what digital solutions to use. The development of shared standards and frameworks has been an important part of DH generally, partly for this reason. Highly successful examples include the Text Encoding Initiative (TEI)[46] for processing text, and more recently the International Image Interoperability Framework (IIIF)[47] for processing images.[48] Given the time and money invested in supporting frameworks such as these, as well as their ubiquity, any digital resource built using them certainly enhances its chances of sustainability in the future.[49] When designing with sustainability in mind, therefore, a potential tension can emerge between eye-catching innovation and well established solutions. There are obvious advantages to using tried and tested tools, methods and technologies rather than bespoke design. One question we might contemplate, however, is whether convergence on certain standards actually limits aspects of DH innovation.[50] In other words, is sustainability achieved at the expense of experimentation or originality? Will these standards begin to confine how we visualise or design digital resources and their functionality in the future? Are these limitations, at the end of the day, a necessary compromise in the pursuit of sustainability? A tangible case study of this conundrum can be found in the work of King's Digital Lab (KDL) at King's College, London. As a result of the innovative projects and extensive collaborations of KDL and its predecessors, the team have come to be responsible for around 160 DH projects, 100 of which were 'legacy projects' created before the establishment

[43] For illustrations of how digital images can enhance or distort aspects of a manuscript, see Stanford (2018) and Morawetz (2019).

[44] For more on the British Library's use of 'ARKs', see Madden & Kotarski (2020: 8–13).

[45] See, for example, the recommendations about career structures (for DH practitioners generally, not just research software engineers) from the recent report by Bergel *et al.* (2020).

[46] *Text Encoding Initiative.* https://tei-c.org

[47] *International Image Interoperability Framework.* https://iiif.io

[48] The important repository of digital images *Irish Script on Screen* (ISOS), for example, has recently begun upgrading its images into IIIF format in order to keep up with user expectations.

[49] Though for discussion of some looming issues with TEI XML encoding, see Schmidt (2019).

[50] For the relationship between sustainability and standardisation, see Drucker (2021: 5). See also the comments in Kálmán *et al.* (2019: 127): 'While research thrives on innovative solutions with fast-paced development progress, the requirements for software maintainability for the long run are directly contrary'. Goddard & Seeman (2020: 42) also note this issue: 'One complicating factor is that sustainability and interoperability are largely at odds with a researcher's freedom of choice when it comes to decisions about platforms, tools, and data models.'

of the Lab in 2015 – some as early as the 1990s.[51] For the Lab itself to be sustainable, an important aspect of its management is the 'homogenisation' of their 'technical stack'. By sticking to a defined set of programming languages and web frameworks, they are able to manage the legacy estate as well as provide support for new projects in the future. While this means that KDL uses a more limited range of technologies, this still allows for some freedom to experiment with new tools and methods without having to maintain too much diversity across the board.

Another prominent context for pursuing this question has been in the world of editing. It hardly needs to be said that access to high quality editions of historic texts is fundamental to so much work in the humanities. Elena Pierazzo has noted that to attract a funder, a digital edition of a text must be innovative, but that:

> … the specialisation of such editions makes their long-term preservation particularly complex and expensive, and in part as a consequence of this, digital editions are perceived as unstable and not worthy of investment … [P]hilologists, who were amongst the first and most keen adopters of digital methods, are still torn between the wish to take advantage of the opportunities offered by the digital medium and the safety offered by a print publication.[52]

There have been recent calls to develop a single, shared infrastructure for publishing digital scholarly editions, by Pierazzo and others.[53] Pierazzo draws a useful distinction between bespoke, experimental digital editions (which is what many digital editions are today) and those which are produced using standard 'frameworks' that do not reinvent the wheel in a technical sense. In proposing the developing of more shared infrastructures for digital editions, she points out that: 'The main problem is that creating such an infrastructure requires a degree of standardisation and implies a loss of originality in the edition itself', where the 'digital' part of the work is seen to be sidelined, reduced to merely a 'tool' rather than an innovative part of the project itself. She continues with a reassuring alternative: 'However, it might well be that not all digital editions are intended to be ground-breaking in terms of the digital solutions they offer: in many cases, textual scholars want to be ground-breaking with the texts they publish, and not with the publishing solutions they offer'.[54] Especially where sustainability is the core concern, then, widely established, even 'off-the-shelf' solutions are surely advantageous for any editorial project.

Another crucial problem in relation to technological sustainability is how to manage the 'after life' of the resource and its data. Digital projects that have moved beyond their short term funded period are faced with the challenge of how to alter or update the resource's content, and how to communicate this to users. This is a particularly prevalent issue since it is not unusual for a project's data production to be incomplete at the end of the funded period.[55] In their 2009 article 'Published yet never done', Brown *et al.* discuss the common practice of 'incremental' releases, which of course is a feature which sets digital publishing apart from the traditional print world where

[51] Smithies *et al.* (2019).
[52] Pierazzo (2019: 210).
[53] Cummings (2019: 190–1): 'Instead of creating solutions that are individual to any specific project's needs, we need collaboratively to build small modular improvements on top of a generalised infrastructure for the creation, publication, and analysis of scholarly digital editions.'
[54] Pierazzo (2019: 214).
[55] Pierazzo (2019: 211) has noted that often 'at the end of the funding period, many editions are not quite ready for the public eye or do not meet the exacting standards of textual scholars; nevertheless, published they must be, as funding runs out and the funder must be shown that something has been done with the resources they provided. The result is that many editions are put on the web with a "Beta" or a "Work in Progress" disclaimer.'

'publication' and 'completion' are essentially the same thing. They suggest that 'designing projects to incorporate such incrementalism by way of staged releases that mark phases of accomplishment or a number of discrete and in some way publishable deliverables, seems a particularly useful way to structure digital projects.'[56] Rather than regarding this as post-project clean-up work, or as a failure which ought to be quietly rectified, the incremental nature of digital projects ought to be embraced. They argue that this is particularly suited to some types of project, notably digital editions where, for example, parts of the text can be released in stages.[57] *The Endings Project* have also been probing the wider question of when DH projects actually 'end', intellectually rather than financially, and offer some principles for what they call 'release management'.[58]

It is useful here to look to those longstanding resources which have gone through successive 'phases' of incremental evolution. The process of maintaining eDIL, for example, has involved various new versions, including the expansion of the marked-up data and technical enhancements of the search engine and its capabilities. As a result, a Supplement (2013) and Companion (2019) have been produced which document the changes to content made in each version.[59] For PoMS, the pattern of funding has meant that the underlying dataset itself has developed in each successive project in order to fulfil a new 'research question'. This has led to the addition of maps and social network analysis visualisations, as well as an extension of the timespan from 1314 to 1371 (but only for a subset of the data: royal documents). What this means is that the original database's scope and structure is no longer obvious (though it is recoverable by limiting the search criteria). All of this highlights the importance of 'accessed on' dates in references, though these are only as useful as the website's documentation of its successive 'versions' and updates.

If digital research tools are to have long lives, perhaps we ought to become accustomed to thinking about updates to their form and contents as 'new editions', much like in the print world. There are important questions to address here: should the original edition (the website or dataset) still be retrievable in some way, as is the case to an extent with eDIL and PoMS? Are there easy ways of being more transparent about this process (e.g., including the equivalent of a new 'preface' for each edition, much like eDIL's Companion)? How much would have to change in order to define a new 'edition' of a digital web resource? Flexible adaptation and updates are arguably one of the strengths of digital resources, not least because it allows the shape of the data and interface to respond to its actual use and emerging interests. Indeed, as time goes on, the greater the technical capabilities of a digital resource the greater will be the user demand for increased functionality.[60] Many websites are now labelled as distinct 'versions' (Version 1.1, 1.2, 2.1, 2.2, etc.), but this is often perceived in terms of technological updates. Perhaps we ought to be embedding an expectation of 'new editions' into our wider understanding of digital research tools and how we use them.

[56] Brown *et al.* (2009: paragraph 9).
[57] Brown *et al.* (2009: paragraph 9). See also Cummings (2019: 183–4).
[58] *The Endings Project*, see 'Principles: Endings Principles for Digital Longevity, Version 2.2, 2022-08-31', nos 5.1–5.5. https://endings.uvic.ca/principles.html#release-management [accessed 4 October 2022].
[59] For the Supplement, see *eDIL: Supplement*. https://dil.ie/supplement [accessed 24 September 2021]. Another example of a web resource which has detailed its 'site updates' is *The Auchinleck Manuscript* (National Library of Scotland, originally published 5 July 2003), Version 1.1, 'Archive of site updates'. https://auchinleck.nls.uk/editorial/archive.html [accessed 24 September 2021].
[60] For longstanding resources such as the British History Online Digital Library (run by the Institute for Historical Research), it has been possible to be responsive to actual user needs, which has made the resource have 'academic sustainability': Blaney & Winters (2010: 99; 104).

Finally, it is worth recognising that another broad area of technological sustainability is 'digital preservation'.[61] Here we can look to libraries and archives since sustaining access to materials is part of their core remit.[62] Both have significant expertise, tools and models specifically for supporting the technicalities of digital preservation.[63] Indeed, it has been argued that the only long-term solution for much digital research is to ensure that the relevant library is a project collaborator, with the project providing support and resources for the archiving process, and even integrating the library's aims into the project's missions.[64]

It is useful to recognise that there can be a theoretical distinction here between digital *preservation* and digital *archiving*. Though individual institutions might define these practices differently, archiving generally implies an element of public access not just storage.[65] It has recently been argued that, because access requires the archived materials to be 'usable' and not just 'available' for consultation, robust descriptive metadata is of the upmost importance when archiving in a web context.[66]

Cultural Heritage Organisations now regularly have to rethink their digital preservation strategies.[67] As there are significant staff, infrastructure and cost implications to this process, it requires careful planning and review. The National Library of Scotland (NLS) now manages this, for example, by creating 'digital preservation and access plans' for each distinct type of digital material that they hold – whether moving images, audio files, magnetic tapes, digital images or other types of digital data.[68] Though storage costs are generally coming down, they can still be a significant outgoing for an institution and, importantly for a legal deposit library such as the NLS, the amount of storage space required will expand greatly over time. The National Archives (TNA) have been very active in their outreach in this area, most recently through their project 'Plugged In, Powered Up' which provides a range of information, tools and training for archivists in order to build their capacity to meet their own institution's digital ambitions.[69]

For individual humanities research projects, attempts have been made in the past to address the problem of preservation, a notable example being the establishment of the Arts and Humanities

[61] For a short history of digital preservation, see Bradley (2007: 151–6).

[62] See Abbey (2012: 91): 'The process of facilitating the survival, or sustainability, of the cultural record is at the very heart of what archivists do. It is a vital part of the raison d'etre of archival science.'

[63] Goddard & Seeman (2020).

[64] Kretzschmar & Potter (2010). Eschenfelder et al. (2019: 193) have also noted the potential role of libraries: 'Greater attention to the closure of projects, the process of gracefully winding things down and the responsible transfer of collections could be very helpful to LIS professionals.'

[65] See Vlassenroot *et al.* (2019: 86): '"Web archiving", therefore, refers to the whole process, whereas "web preservation" is one of the steps in the process of archiving the web.'

[66] Vlassenroot *et al.* (2019: 103).

[67] The Digital Preservation Coalition (of which many libraries, archives and HE institutions in the UK and Ireland are members) have produced a handbook as well as a set of reports and papers offering guidance on matters of digital preservation for organisations: *Digital Preservation Coalition: Digital Preservation*. https://www.dpconline.org/digipres [accessed 20 February 2022].

[68] This is one part of the National Library of Scotland's *Digital Preservation Policy and Plan* (April 2021). https://www.nls.uk/media/uo4b1zfj/digital-preservation-policy-and-plan.pdf [accessed 1 October 2021].

[69] See *The National Archives*, 'Plugged In, Powered Up'. https://www.nationalarchives.gov.uk/archives-sector/projects-and-programmes/plugged-in-powered-up/ [accessed 3 September 2021]. The programme includes a training course ('Novice to know how'), a mentoring scheme, advocacy resources, guidance on digital preservation workflows, and case studies of successful digital strategy implementations. TNA have also developed PRONOM, a resource which provides technical information about those file formats which archives typically use, and the software products which support them, in order to help archivists make decisions about file formats by looking to the future to try and predict potential migration requirements: https://www.nationalarchives.gov.uk/PRONOM/Default.aspx

Data Service which aimed to 'collect, preserve and promote the electronic resources which result from research and teaching in the arts and humanities'. Unfortunately, this service was disbanded in 2017.[70] Rockwell *et al.* discuss ways to deposit the data at the end of a project, what they describe as 'properly burying' the project which should 'die gracefully'. This is not just about depositing the data somewhere; it is about 'leaving as a legacy the research data developed in a form usable in the future'.[71] When done properly, this process can take years, as in the case of their Globalization and Autonomy Online Compendium.

For current DH research projects based in Higher Education Institutions, the good news is that these institutions are increasingly supporting the preservation of this kind of digital scholarship. One result of this is that localised hosting can now be more readily quantified in terms of real costs. Though digital preservation practices might be unique to an individual institution, it is likely that some standard practices will emerge as we continue to share experiences and best practice. KDL, for example, has recently formalised the kind of post-project management that it offers to digital resources. This includes Service Level Agreements (SLAs) which typically last around five years; migration; static conversion; dataset deposit;[72] and minimal archiving and storage.[73] While progress is clearly being made in this area, it should be acknowledged that many projects remain reliant on the willingness and time of those involved in the original project who have retained the skills – and also the access – to provide basic ongoing maintenance.

There are, of course, many more technical considerations for sustainability than have been discussed here, and each point that has been raised has much greater depth to it than has been offered in this survey. What can be seen, however, is that there are important issues for humanities researchers to be aware of: the current emphasis on research as 'data' production and the need to plan for its quality and contextualisation at the outset (especially as a result of the growing expectations around 'reuse' and 'interoperability'); the role of high-level guidance such as the FAIR principles and what this can mean in context (such as simplicity of design based on tried and tested technologies, as well as citability and Persistent Identifiers); the potential to embrace incrementalism in digital research publishing and to think about digital resources in terms of 'editions'; and the role of institutions in digital preservation and archiving of projects. Despite being of a generally 'technical' nature, humanities researchers have something important to offer in all of these areas, especially given our expertise with research publications and their long-term uses and influences. If sustainability in DH is indeed an ongoing activity, then it is one in which the 'humanities' ought to have just as loud a voice as the 'digital'.

[70] The project itself has now been archived: *Arts and Humanities Data Service: Enabling Digital Resources for the Art and Humanities*. https://ahds.ac.uk/ [accessed 21 February 2022].

[71] Rockwell *et al.* (2014: 31). They discuss what aspects of a project should be deposited, where, and some of the issues with this process including funding and the instability associated with project members moving jobs.

[72] An example of a dataset deposit (or 'dump', as the process is sometimes ungracefully referred to) is the recent transition of a database related to PoMS, *People of Northern England (1216-1286) database archive*: https://researchdata.gla.ac.uk/1126/ [accessed 1 October 2021]. This has now been archived in raw SQL format by the University of Glasgow Library.

[73] See a description of these options at *King's Digital Lab*, 'Archiving and Sustainability'. https://kdl.kcl.ac.uk/our-work/archiving-sustainability/ [accessed 24 September 2021]. The University of Oxford is another institution which is formally tackling this issue. In August 2020 they launched a 'Digital Sustainability service' for Oxford DH projects: *Digital Humanities @ Oxford*, 'Digital Sustainability at Oxford'. https://digital.humanities.ox.ac.uk/digital-sustainability-oxford#/ [accessed 5 July 2021].

Funding digital sustainability

It is widely recognised that one of the major obstacles for digital sustainability is funding. While there is a natural tendency to regard web publishing as 'free', the reality for DH projects can be far from this. Large digital projects are funded by different routes: the *Thesaurus Linguae Hibernicae* and CELT are two examples which began life with private funding, for example. When the finances for a project derive from the public purse, however (as is more often the case), certain sustainability requirements come to the fore, notably the imperative to make the research results (including the 'data') available to a wide audience and beyond the project's funded lifetime.

A widely shared concern about the current funding model is its 'short-termism'. It is entirely natural, of course, that funding is structured by defined pots of money assigned to finite research projects (typically three to five years for the larger projects). For DH projects, however, this is particularly troublesome. One institution which has keenly felt the problems with this model is KDL. Many of its inherited projects are now in a funding 'limbo', having gone beyond their funded period but still being used by a community. The team at KDL have identified a range of problems associated with funding models, with short-termism as one of the most prominent given that it appears to 'incentivise researchers to produce "orphan" projects with uncertain futures'.[74] As such, the KDL team have been keen to point out the 'human cost' of this model since it is often the 'goodwill of colleagues' which saves these resources.[75] High costs of maintenance are not, therefore, usually storage space or security certificates or hosting, but developer time, especially if significant updating or re-engineering of a site is involved.

There is a central tension for humanities researchers seeking to apply for money for a digital project. Funders naturally look for field transforming resources which will last; they also, however, want projects that are led by a specific 'research question', a question which must be answered by the creation of that digital resource. This is not a tension if the resource is designed to be 'of the moment', a tool to answer a very particular question. If, however, longevity is desirable or necessary, then applicants must think carefully about how to package the digital resource's primary purpose as a short-term objective. Inevitably, all digital research tools are 'of the moment', in their conception if not in their stated aim. Returning to the example of digital editions, Pierazzo has noted that the short-termism of the funding model 'does not allow for the discovery of new materials and/or the kinds of complications which are often encountered in work involving manuscripts and other types of primary sources'.[76] In other words, despite all the flexibility offered by the digital environment, the nature and parameters of any digital resource typically have to be determined before the research itself has been carried out. Given all this, it is not surprising if humanities researchers are deterred from opting for the digital as a medium for their published work.

Moreover, once a project comes to the end of its funded period, there are no funding pots (or at least none that are well known) simply for updates to an existing website or digital research tool. Indeed, funders often explicitly exclude this kind of activity.[77] It has been said that there are instead

[74] Smithies *et al.* (2019: paragraph 35).
[75] Smithies *et al.* (2019: paragraph 15).
[76] Pierazzo (2019: 211).
[77] The current guidance for the AHRC's Follow-on Funding for Impact and Engagement Scheme, for example, states that it 'Cannot be used to support resource enhancement activities or to develop or extend an existing website or resource.' UKRI, *AHRC Research Funding Guide*, Version 5.7 (July 2022), p. 30. https://www.ukri.org/publications/ahrc-research-funding-guide/

'[c]ultural biases that celebrate program creation' rather than continued support and maintenance of existing resources.[78] Thanks to the perseverance of the team behind PoMS, it is a rare case where the web resource has been kept alive by a succession of funded projects (Arts and Humanities Research Council and The Leverhulme Trust). The continuation of a digital resource's funding is not, however, the same as being 'sustainable'. For PoMS, each separate project did not primarily fund the database for its own sake but rather (as noted earlier) to answer its own new 'research question'.

If funding is being sought for a project beyond its original funded lifetime, it is useful to be able to show that the resource is still providing value to a community of users. Currently, however, far less time and energy tends to be invested in assessing the use of resources than in developing and launching them. According to Lorna Hughes: 'There is a perception that revisiting a digital resource to understand its impact, assess its user base, and enhance its embedding in teaching and research, are luxuries that institutions can rarely afford.'[79] This she links to the short-termism of funding models and the lack of early assessments of how user engagement will be achieved (which should ideally be addressed before seeking funding). In 2013, a report was published by Ithaka S+R entitled *Sustaining Our Digital Future: Institutional Strategies for Digital Content*. The study (funded by the Jisc-led Strategic Content Alliance) looked primarily at the humanities and social sciences and included recommendations for funders, academic project leaders, and managers at universities and cultural institutions. A notable point of emphasis in their report was the current lack of post-project management, especially in relation to user engagement. Indeed, the report's definition and measure of what is a sustainable resource centres on the 'value' that it delivers to its intended users.[80] If digital research tools are to be truly sustainable, therefore, it is important that their hosts understand their continuous use, not just the scale of users but the nature of the community itself. KDL, for example, undertakes a post-project assessment of a resource's impact which includes not only its user traffic but also its research value and impact.[81] Currently, however, such assessments can be difficult for individuals to conduct.[82] Information tends to be anecdotal, not least because a resource's use is not always apparent in the footnotes of published work. This may be a result of digital sources seeming to be 'unstable' and it therefore appearing to be wiser to refer to the original manuscript or printed edition. Measuring success might therefore require the resource creators to find more direct and creative ways of engaging with their community of users. One framework to support this (developed by the Oxford Internet Institute in 2009, with funding from Jisc) is the *Toolkit for the impact of digitised scholarly resources* (TIDSR), which 'offered a mix of qualitative and quantitative methods to measure impact, including a wide variety of case studies demonstrating impact in action'.[83]

[78] Eschenfelder *et al.* (2019: 193).

[79] Hughes (2014: section 6).

[80] Pickle & Maron (2013). For the recommendations, see pp. 48–9. Their full definition of sustainability (at p. 12) is: 'The ability to identify the resources needed to keep a digital resource or service delivering value to the users it is intended to serve.'

[81] Smithies *et al.* (2019: paragraph 29).

[82] An example of a study of a digital resource's predisposition to use or neglect using log analysis as well as qualitative methods of questionnaires and interviews is Warwick *et al.* (2008).

[83] Jisc, *Toolkit for the impact of digitised scholarly resources*: https://www.jisc.ac.uk/guides/toolkit-for-the-impact-of-digitised-scholarly-resources [accessed 13 October 2021]. This project has now been archived, but Jisc have added its findings to another guide, *Making your digital collections easier to discover*, which develops 'techniques to increase reach of your digital collections, optimise them for discovery and evaluate their usage and impact': https://www.jisc.ac.uk/guides/making-your-digital-collections-easier-to-discover [accessed 13 October 2021].

Overall, greater attention ought to be paid to our methods for measuring the use and impact of digital resources, especially beyond their funded lifetimes.[84] Anyone strategically minded might conclude that there is an incentive to produce 'high traffic' digital resources since the case for financially sustaining them will be easier to make in the long run. It is yet to be seen how far this thinking will affect the kind of digital resources we choose to build in the longer term. It will also be incumbent on the creators and/or hosts of the resources to continuously measure how they are being used and by whom – a crucial aspect to sustaining a digital resource and its community of users.

Another crucial dimension to funding is institutional hosting, serving and ongoing support. For many institutions – and especially those with inflexible budgets – the key challenge for supporting digital resources on a continuous basis is that it is difficult to predict what will be necessary in the future. Local university or library finances, for example, are unlikely to be able to manage this volatility for a large and growing number of projects, even for those produced 'in house' by DH departments. There is also a risk that each institution will end up reinventing the wheel to some extent for every individual project. Instead, our collective ambition for sustainable digital research means that solutions need to look beyond individual projects or local institutions.

This leads to the emerging role of national and international research infrastructure consortia. Current examples of frameworks and resources which serve a range of disciplines and areas include: the European Strategy Forum on Research Infrastructures (ESFRI); the Social Sciences and Humanities Open Cloud (SSHOC) and the European Open Science Cloud (EOSC); the Digital Research Infrastructure for Arts and Humanities (DARIAH); and the Digital Repository of Ireland (DRI).[85] Many of these have been explicitly set up to address the challenges surrounding digital sustainability, including depositing and disseminating research results, ensuring they adhere to certain standards and licensing conditions, and enabling interoperability. A core aim of DARIAH, for example, is 'to help developing sustainability models for Digital Humanities (DH) projects and their data collections, especially to ensure the longevity of such projects after the direct funding period has run out'.[86] Such work is increasingly described as 'data stewardship'. Education for researchers about the role of these infrastructures is now imperative, but so also is the role of the researcher as an advocate for such infrastructure given its significance in supporting research projects and their future sustainability.

Other matters related to funding could readily be addressed in the near future. A practical issue is the lack of specificity required in project funding applications themselves. For the Arts and Humanities Research Council, funding applications previously included a 'Technical Appendix', though a consultation in 2009/2010 recommended this be redesigned in its scope and conceptualisation, later giving way to the Data Management Plan.[87] It has recently been argued that, today, application forms and data management plans do not ask for enough information, including that relating to 'system quality, infrastructure, or lifecycle management'.[88] Given the importance of

[84] See Hughes (2014: section 6): 'It may be that we have placed too much emphasis on technical sustainability, the provision of research infrastructures and technical management of the digital life cycle as topics for research, and that more investigation of the use of digital content in scholarship is necessary in order to safeguard our valuable digital assets.'

[85] For the recent development of a tool that helps researchers in the humanities identify an appropriate data repository, see Buddenbohm *et al.* (2021).

[86] Kálmán *et al.* (2019: 116).

[87] McLaughlin *et al.* (2010).

[88] Smithies *et al.* (2019: paragraph 35). See also paragraph 27: 'While it is unreasonable to expect funding agencies to provide ongoing funding for all projects, it does seem reasonable to ask their support for projects that are managed according to

sustainability for the wellbeing of research, it seems logical that this should become a higher priority, or should at least be reframed, within the criteria for funding. Indeed, this was a recommendation of the 2013 *Sustaining Our Digital Future* report.[89] If sustainability is regarded as an essential requirement for achieving funding in the first place, clearer guidance needs to be in place for what this entails and how it might be realised. For example, there is a real need for more costed models for sustaining digital resources in the humanities.[90] These would allow for all involved to gain a clearer sense of potential institutional roles and responsibilities at the outset of any given project.[91]

When planning for a digital research project's future, it need not be that one person or one part of an institution takes full responsibility for the entire project's sustainability. There is potential for many to share the responsibilities – and the benefits – of continuously supporting a project. This is sometimes now referred to as the 'distributed' or 'decentralised responsibilities' of sustainability.[92] There is even potential for cross-institutional, inter-dependent hosting of a project, with various partners supporting (and therefore funding) different aspects of it. This, however, needs to be carefully designed, defined and managed. An eye also needs to be kept on whether the work of building a support infrastructure is being unnecessarily replicated across many institutions.

When it comes to funding sustainability, a change in mindset is occurring at an individual and an institutional level. There has been some positive movement here, with libraries and digital research centres taking a more active role in preserving digital research projects and developing frameworks for doing so. This will hopefully be propelled in future by investment in national and international infrastructures. Indeed, this change in mindset could be linked to the growing realisation that sustainability is an ongoing activity, not a hole that can be filled by a finite pot of money.

Environmental aspects of digital sustainability

It must be said that the impact on the environment is perhaps the most ominous dimension of digital sustainability. The relationship between digital activities (especially data storage) and the human carbon footprint is becoming a more prominent part of the mainstream conversation.[93] Environmental issues vary: the extraction of raw materials for the production of technological devices; the manufacture of such materials and the waste created in the process; transportation impacts; the storage of digital data (especially the demands on electricity and cooling facilities when this is done at scale); e-waste from discarded devices; as well as the impact of buildings and human occupation of them.

---

transparent processes and accepted frameworks, that include a range of archival approaches, and integration into Research Data Management (RDM) systems that leading research agencies advocate greater use of.'

[89] Pickle & Maron (2013: 48). Their recommendations for funders were as follows: 'Clarify sustainability needs when awarding grants', 'Challenge unrealistic impact statements' and 'Encourage – or require – grantees to think creatively about how their content will best reach an audience'.

[90] For an example of a detailed study of a costing model in a digital context, see Palaiologk *et al.* (2012). For general discussion of models for funding financial sustainability, see Blaney & Winters (2010: 101–2).

[91] See also Smithies *et al.* (2019: paragraph 35).

[92] Pickle & Maron (2013: 49).

[93] Oberhaus (2019).

Some of the DH community have begun to discuss in earnest the relationship between scholarly work and the environment, though inevitably this is mostly in theory rather than in practice.[94] In the library, museum and archive sectors, literature on climate action has been relatively well established since the 1970s. The emphasis, however, has tended to be on the organisation's own environmental impact, including institutional facilities management (energy efficiency, water conservation, recycling practices), collections management processes (including storage strategies and materials), and outreach practices (such as using sustainable materials in exhibitions).[95] When digital activity has been discussed in the literature relating to libraries, it has typically been regarded as a separate concern from environmental sustainability.[96] One exception is the work of Gobinda Chowdhury who has written about the carbon footprint of the library sector and ways to reduce it, such as using cloud computing.[97] In 2019, the Climate Heritage Network was established which is 'a voluntary, mutual support network of arts, culture and heritage organisations committed to aiding their communities in tackling climate change and achieving the ambitions of the Paris Agreement.'[98] There is, however, no explicit mention of digital activities and their environmental impact in their mission statement.

A recent comprehensive analysis of digital environmental sustainability in the context of Cultural Heritage Organisations was published in 2019 by Pendergrass *et al*. The authors state the problem clearly:

> As cloud and other networked storage systems become more prevalent, digital storage is increasingly disassociated with its physical impacts by end-users, leading to greater, and less judicious, use. This has serious consequences for environmental sustainability, due to the large support infrastructure needed for usable cloud and networked storage, and the raw materials and energy that this infrastructure requires.[99]

The innocent metaphor of 'cloud' storage is not as harmless as it may seem. Servers themselves are often simply out of sight but still require enormous amounts of energy for power (and cooling). Pendergrass *et al*. therefore propose some practical ways forward for Cultural Heritage Organisations, including: basic energy efficiency measures while still meeting user needs; scheduling tasks for off-peak times; and moving to clean energy sources (while being conscious of the location of the facilities and their local climate). More profoundly, however, they argue that for real progress to be made, a paradigm shift is necessary. They highlight three aspects in particular. One is archivists paying more attention to their 'appraisals' of their holdings (i.e., being more selective; including environmental impact as a criterion for digital preservation strategies; avoiding unnecessary duplication; considering file sizes; and undertaking regular reappraisals). Second, the notion of 'permanence' ought to be reassessed (i.e., institutions should calculate and embrace acceptable levels of loss, as with physical materials; they should take a tiered approach in their prioritisation; and they should consider carefully their storage technologies, whether local,

[94] Nowviskie (2015); Drucker (2021: 6).
[95] Abbey (2012).
[96] Bradley (2007: 156–9). For Bradley, the main barrier to digital sustainability stems from how to finance it. See also Eschenfelder *et al.* (2019) who have shown the ways that Library and Information Sciences are concerned with 'organisational sustainability', mainly the technical, financial and management elements of the organisation, rather than its environmental sustainability.
[97] E.g., Chowdhury (2014).
[98] See the *Climate Heritage Network*. https://climateheritage.org [accessed 1 September 2021].
[99] Pendergrass *et al.* (2019: 173–4).

distributed or 'cloud' based). Third, and perhaps most radically, we ought to re-examine the default assumption of immediate access, which would include a re-assessment of the value or otherwise of mass digitisation programmes and potentially a move towards on-demand digitisation and access. While all of these are questions for Cultural Heritage Organisations themselves to consider, each potentially has a direct impact on the work of humanities researchers, who may even need to reorientate their understanding of the archive's very function.

The impact of our collective digital activity and storage on the environment is certainly alarming. The other side of the coin is, of course, the implications of reduced travel to archives and libraries as a result of greater access to materials online, especially digital images of archival objects and print publications. It is almost certainly too complex to calculate the exact balance of environmental impact of travel versus provision of digital resources. What is clear is that access to the physical objects and access to complementary online materials are now equally important aspects of scholarship and public engagement, and we can therefore almost certainly expect both to continue long into the future. The challenge for sustainability is to make both as 'green' as possible.

While digital environmental sustainability is a much wider conversation, and one which humanities scholars might follow from a distance, it is important to recognise that advocacy at a local level also has the potential to be a tool for driving change here. Individuals and organisations in all contexts can exert pressure on cloud storage providers to measure and declare their greenhouse gas emissions. Organisations could even integrate this factor into their conditions for procurement, or funders might add this to their application process, making environmental considerations a routine part of the assessment criteria.

'Human' aspects of digital sustainability

One of the most important ingredients in the overall recipe for sustainability is human time. It is also one of the most unpredictable aspects in the mix. Indeed, the survival of many digital resources has often come down to personalities and individual goodwill to keep a project going.[100] Personal enthusiasm is, of course, not something that can easily be anticipated or written into a business case or funding application. Nevertheless, instances of emotional engagement have often been the crucial factor leading to a resource's long-term success. What can be controlled is the development of an engaged community, supported by adequate promotion and training. Training itself is not always explicitly conceptualised as part of planning for sustainability.[101] It is clearly a significant factor, however, for maximising a digital resource's vitality in the medium- and long-term future.

When developing DH training opportunities, it is important to appreciate the context, whether it is for researchers from a particular disciplinary background, for librarians and archivists, for the general public, or for postgraduate and undergraduate students, to name just some groups. It should also not be assumed that training is only necessary for individuals at the earliest stages of their careers, given how quickly DH can develop and how pervasive digital research tools can become for research, teaching and curation. It may also be that formal training is not always the best approach: internships or secondments can be a very effective way of steadily developing new

[100] Drucker (2021: 3); Smithies *et al.* (2019: paragraph 15).
[101] Bergel *et al.* (2020) emphasise the importance of training for sustaining DH research generally.

skillsets and they can also help to embed these digital skills within an institution at large, especially Cultural Heritage Organisations. Another major context for digital training is funded research projects, not just for the Research Software Engineers (RSEs) but also any collaborating analysts and humanities researchers. Individual career development is a central concern that must be balanced alongside the development of the resource itself. This makes good sense if we recognise that sustainability of digital resources is linked to the sustainability of individual careers.[102] Such investment also has obvious institutional benefits since it can ensure continuity of expertise and domain knowledge.[103]

There is already a large menu of DH training courses on offer. These range from intensive summer schools which allow participants to learn about a specific technology (such as the *Digital Humanities Summer Institute* at the University of Victoria or the *Digital Humanities at Oxford Summer School*) to online tutorials which individuals can work through at their own pace (such as the *Programming Historian*).[104] With so much choice, humanities researchers can naturally feel overwhelmed. In navigating this, it may be unhelpful to think of ourselves as having a skills 'gap', which suggests a finite space that can eventually be filled. In such a fast-moving environment, what constitutes 'essential' digital knowledge and skills never stands still. It is important for humanities researchers in particular to recognise this, and to view upskilling as an ongoing process. Indeed, in the context of digital editions, it has been questioned whether or not textual scholars ought to possess the technical skills needed to create a digital edition, or whether what is needed is a careful division of labour in a collaborative partnership.[105]

It is useful to distinguish between 'targeted' training in the use of particular tools and 'holistic' training in DH generally. Breadth of exposure and experience is a particular virtue for humanities researchers, especially at an early career stage. Indeed, many universities now offer taught courses specifically in DH. One approach that can be successful is to make this training less focused on acquiring generic technical skills (in the abstract hope that these will come in handy in future) and more focused on project specific experiences of DH technologies in which skills are learned in a real and particular context. Importantly, this reframes 'training' as about more than just acquiring the technical know-how but also as an ongoing experience of collaboration and development.

Another crucial distinction can be drawn between training to *build* a digital resource and training to *use* one. Training programmes might often be a mixture of these two, but the emphasis tends to be on the former (e.g., how to write your own TEI, how to create RDF triples, how to write python scripts). Given the nature of our digital material, however, training that is explicitly targeted at a resource's users is fundamentally important. Today, web design is broadly such that the interface and its functionality are intended to provide a seamlessly intuitive 'user experience', whether the user is shopping, browsing social media, searching a repository, or reading the news. For academic research, often our material – both our primary sources and the data we create – is highly nuanced. It does not always allow for entirely 'intuitive' use. Instead, it ought to be consumed slowly. The internet, by contrast, is not designed for slow use. Perhaps the emphasis of 'user

[102] The support of researchers has now been formalised in the UK under the *Researcher Development Concordat* (https://researcherdevelopmentconcordat.ac.uk/). See also the recommendations in Bergel *et al.* (2020).
[103] For RSEs in particular, KDL has been pioneering pathways for their career development: see Smithies (2019).
[104] *Programming Historian*. https://programminghistorian.org
[105] Pierazzo (2019: 212, 217).

training', then, should be on the slow use of digital resources for research.[106] In this, we should not forget the role of printed books. Many digital resources are based to some extent upon previous printed works or editions (including eDIL, CELT, *Monasticon Hibernicum*, and PoMS). Such material objects can act as a reminder of the value of slow and methodical consumption of scholarly work. In other words, we should be encouraging users to 'read' a digital research tool and not just 'use' it. Overall, perhaps we need a new narrative about training in (or rather 'experiencing') DH, one that is not just about acquiring technical skills but also involves learning something new about the material with which we work, and how we can best work with it and with others through it. This would allow us to more firmly centre the 'humanities' within DH training.

Another emerging theme in relation to the 'human' aspect of sustainable DH resources is how to manage 'transitions' for major digital projects. What happens, for example, when the original Principal Investigator is no longer able to continue championing the resource and locating funds for it (most obviously in the case of retirement or if they move from the host institution)? How can their expert role be filled, which often includes a unique mix of scholarly and digital skills and experience?[107] Such 'generational change' is a great concern not only for users but also those hosting or managing digital resources. It might be that more formalised training is an important part of succession planning.[108] Currently, however, transitions are left to an ad hoc arrangement, usually reliant on a mixture of individual enthusiasm and fortunate circumstances – a further reminder of the fragility of sustainability when 'human' factors are taken into consideration.

Is digital sustainability necessary?

The above discussion has endeavoured to deepen our understanding of 'digital sustainability' in relation to humanities research. A potentially controversial question ought now to be considered: is sustainability always necessary, or desirable, for digital research resources? It is fair to say that most DH projects begin life with aspirations of changing the future of scholarship and access for a generation if not longer. But not all digital resources need to fulfil this particular ambition. There may be space for resources which satisfy a relatively short-term research objective.[109] Projects can then perhaps be more readily archived, as is now done with the rest of the web by the likes of the Internet Archive[110] and also national libraries and archives.[111] This would involve a reorientation of how we think about digital research tools, with some as explicitly short-term and 'of the moment'.[112] This need not be a sign of failure: it may, in fact, be the sign of success, since it could be argued that a digital research tool will have fulfilled its purpose if it enables researchers to take their field in entirely new directions, rendering the resource itself obsolete. The crucial point is that

[106] This links to other ideas of 'slow research', 'slow digitisation' and 'slow archives': e.g., Christen & Anderson (2019); Prescott & Hughes (2018).

[107] An example of a recent transition is *Thesaurus Linguae Hibernicae*, originally led by Patricia Kelly and now voluntarily taken over by Fangzhe Dimurjan Qiu.

[108] One potential approach, currently being pioneered by Pádraic Moran, is to establish an 'editorial board' for a digital resource that creates a community of individuals with the knowledge and interest in maintaining it.

[109] This is also discussed in Smithies *et al.* (2019: paragraph 26).

[110] *Internet Archive.* https://archive.org

[111] Vlassenroot *et al.* (2019).

[112] Varela (2016) has argued that performance theory can be helpful for understanding digital archives as 'traces', not the real thing, and therefore something transient and performative.

sustainability should not be sought simply because it is possible to keep digital resources alive. This is implicit in the passing reference by Pendergrass *et al.* (quoted above) to 'less judicious use' of digital resources spurred by the allure of our growing digital capacity.[113]

This idea of purposefully transient digital resources has not received much attention in the DH world, where the emphasis is firmly (and justifiably) on supporting existing and future projects to become sustainable.[114] We might instead consider here the work of archivists, who have more of an established relationship with the 'temporality' of records and information. In an important article from 1989, James O'Toole discussed the concept of permanence by exploring the meaning of archival 'preservation' – its deep roots and its changing meaning for archivists across the modern era, from the preservation of the 'information' in the records to the preservation of the physical document itself as a result of improved conservation techniques. Subsequently, he concluded, the archivist now understands permanence as something 'relative' rather than 'absolute':

> Having become convinced [in the mid-twentieth century] of how important it was to preserve their physical holdings permanently, archivists began to realize how impossible it would be to do precisely that. … [A]rchivists began to grow uncomfortable with the apparently limitless commitment that adherence to a notion of absolute permanence implied, and they began to view questions of appraisal and preservation in much more relative terms.[115]

Much of this statement could apply directly to digital research tools in the 21st century. As libraries and archives have now become responsible for 'born digital' materials, they have increasingly turned their attention to the question of digital permanence. In the journal *Library Trends*, Kevin Bradley pointed out that: 'Not all data will, or should, be sustained in perpetuity … Certain datasets or learning objects may have intellectual, teaching, or research value for only a short period of time'.[116] Sustainability might not, therefore, always be necessary for quality research experiences. We should also not forget that there are different levels of input required for sustainability: some digital resources might involve 'ongoing growth and investment' while other digital content might require more straightforward 'maintenance' as part of a larger platform.[117]

Now, this is not to say that no resource or dataset deserves long-term investment. The question is whether sustainability needs to be an objective for every DH project. Should we, as humanities scholars, more firmly embrace short-termism in the digital environment, rather than trying to fight against this? Perhaps as well as shifting our working definition of digital sustainability, we also need to reorientate our expectations of what digital resources themselves are seeking to achieve, and what is meant by 'successful longevity' in the digital environment. Would we conceive of digital research projects differently if, for example, we anticipated our audience would be limited to one generation at most? If so, then DH activities generally might be more firmly regarded as part of the transitory infrastructure and methodologies of research, rather than the 'product' of it.

It may be useful to offer a parallel here. Scholars already think in terms of sustainability in relation to print. Printed books and journals are expected to survive for centuries, but they can easily be neglected; we therefore also 'sustain' our published work through presentations, through

[113] Pendergrass *et al.* (2019: 174). See above, p. 111.
[114] One exception is the concept of 'graceful degradation' for DH projects proposed by Bethany Nowviskie and Dot Porter who conducted a survey of DH projects in 2010 investigating the issues associated with 'transition and decline': Nowviskie & Porter (2010).
[115] O'Toole (1989: 21).
[116] Bradley (2007: 159).
[117] As pointed out in Pickle & Maron (2013: 5).

teaching, by nurturing interested communities, building careers, and doing outreach work with the wider public. It could be said that, in the print world, we are fully accustomed to the fact books and journal articles cannot be changed after their publication, but that the research itself is not 'static'. We are simply comfortable with print as a 'tool' – what it can do and what it cannot do, what it is appropriate for, and how we should use it to our best advantage. It is important, therefore, that we become collectively accustomed to the nature of digital resources and publications as one of many 'tools' for research and teaching.

Conclusion

For many humanities researchers, the sum total of concerns about sustainability will amount to a cautionary tale that warns against taking on the responsibility of creating any new digital research tool. Sustainability is difficult to plan, implement, and measure the success of. It is therefore right that we should be wary of diving into Digital Humanities projects without due consideration to all of the implications, both short- and long-term, foreseeable and unforeseeable. We also ought to thoroughly consider whether a new digital research tool is always necessary for achieving our research and dissemination objectives. The issue of sustainability should, therefore, cause us to pause and think carefully about how and why we produce and publish digital research in future. Despite this note of caution, there is no doubt that DH research, and the production of digital resources, will only expand in future, and to the greater benefit and vitality of our disciplines.

It is clear that if we are committed to continuing to build digital research tools, we ought to do this with an informed and shared understanding of what sustainability actually means, and how it can be put into practice. This article has suggested a new working understanding of digital sustainability, one that is rooted in the idea of sustainability as a range of ongoing activities, not a state that can be achieved. Sustainability is a persistent technical activity, given the need for routine updates; it is financially continuous for the host of the resource or dataset; the production and storage of data certainly has an ongoing environmental impact; and human investments (both in terms of user training and career development) are significant long-term requirements. In sum, there is no such thing as a 'sustainable' digital resource in the sense of 'self-replenishing'. All of these aspects need to be actively managed in order to be kept in motion. As a result of this suite of issues, this article has also raised the question of whether sustainability itself is always desirable, and whether it is possible for humanities researchers to become more at ease with the idea of 'ephemeral' digital research tools and methodologies.

There are reasons to be optimistic about the future of sustainable DH. There has been the emergence of some general guidance (such as the 'FAIR' principles or the work of *The Endings Project*) as well as shared standards (including the likes of Persistent Identifiers for digital material, or IIIF for images). Cultural Heritage Organisations and Higher Education Institutions are now more in tune with digital preservation concerns and 'distributed responsibilities' for sustainability, which is coupled with the emerging role of national and international research infrastructure consortia as 'data stewards'. There are also early signs of moves to reassess our work in relation to its environmental impact, including introducing this as a criterion for procurement, rethinking presumptions of immediate access, and potentially being more selective in 'appraisals' of what digital material is retained. In all this, the role of the humanities scholar is not side-lined but is heightened, given the emphasis on contextualised, 'good quality' data that is worth preserving.

There is also potential for treating digital research tools more like the print publications with which we are familiar, in terms of their standards of citability and transparency, and to regard their incremental growth in terms of 'editions'.

For individual research projects, what is clear is that there is no one size for creating a sustainable digital resource. Sustainability ought to be considered on its own terms and inbuilt from the outset.[118] Early and comprehensive discussions with partners is crucial for any project, but should now include as standard a conversation about long-term sustainability aspirations. This need not be simply an exercise in writing a 'data management plan'. Questions of sustainability can also contribute to understanding the nature of the research itself, its publication and 'use'. This means that project leaders ought to take into consideration the principles outlined above, with more of an initial concern for standards, hosting and maintenance, long-term funding options, environmental impacts, continuous engagement with and training of users, and therefore the project's overall sustainability. It is also worth remembering, however, as a guiding principle, that there are aspects of sustainability that cannot be controlled. Projects ought to allow for flexibility since very often the best solution will only emerge during the course of a tool's extension and use. Moreover, individual personalities and goodwill – neither of which can be easily planned or assumed – have proven in the past to be the defining factor in maximising a resource's life expectancy.

It may be that, moving forward, the aftermath of 2020's sudden 'online pivot' provides greater opportunities for more continuous conversations about issues relating to sustainability, whether via online workshops or training sessions or establishing communities of users for a particular resource. Such conversations are surely part of the solution to recognising sustainability as an ongoing concern, allowing individuals to take stock and hear about approaches in other digital research projects and among other disciplines and institutions. Facing the challenge of sustainability is not something that need be done in isolation, especially if we embrace it as a fundamental part of our ongoing and collective 'digital citizenship' as humanities researchers. Indeed, digital research methods and the digital publication of research provide an opportunity to rethink what it means for research itself to be 'sustainable'. It is not necessarily about our books or articles or project databases and websites lasting into eternity. It can also be something much greater but less tangible: sustainable research as contributing critical perspectives and practices that enrich all future work in the humanities.

Acknowledgements

This article is an outcome from the AHRC- and IRC-funded network *A Digital Framework for the Medieval Gaelic World* (August 2020 – September 2021, Grant reference AH/V00235X/1). It develops the discussion from the fourth and final network workshop, held online on 10 and 11 June 2021, hosted by the University of Glasgow. The article is indebted to all those individuals who attended the workshop and contributed to its discussions. Workshop presenters and facilitators were: Claire Breay (The British Library), Dauvit Broun (University of Glasgow), Arianna Ciula (King's College, London), Beatrix Färber (University College Cork), Deborah Hayden (National University of Ireland, Maynooth), Lee Hibberd (National Library of Scotland), Cillian Joy (National University of Ireland, Galway), Rachael Kotarski (The British Library), Barbara McCormack (Royal Irish Academy), Máire Ní Mhaonaigh (University of Cambridge), Pádraic Moran (National University of Ireland, Galway), Orla Murphy (University College Cork), Fangzhe

[118] As acknowledged by the likes of Edmond & Morselli (2020) and Smithies *et al.* (2019).

Qiu (University College Dublin), David Stifter (National University of Ireland, Maynooth), Greg Toner (Queen's University Belfast), and Nora White (Queen's University Belfast). I am also very grateful to Dauvit Broun for reading various drafts of this article, and to the anonymous reader for their very helpful suggestions. All errors or oversights remain the author's.

References

Abbey, H.N. (2012), 'The green archivist: a primer for adopting affordable, environmentally sustainable, and socially responsible archival management practices', *Archival Issues*, 34: 91–115.

AHRC (Arts and Humanities Research Council) ICT Methods Network (2006), Report on digital sustainability (2006). www.methodsnetwork.ac.uk/activities/es7report.html [accessed 21 February 2022].

Barats, C., Schafer, V. & Fickers, A. (2020), 'Fading Away… The challenge of sustainability in digital studies', *Digital Humanities Quarterly*, 14(3).

Bergel, G., Willcox, P., Armstrong, G., Baker, J., Ciula, A., Cole, N., Nyhan, J., Ridge, M., Seip, O., Taylor, C., Thornton, P., Williamson, E. & Winters, J. (25 September 2020), *Sustaining Digital Humanities in the UK* (Software Sustainability Institute). https://zenodo.org/record/4046267#.Yp5Yni8w2S4

Blaney, J. & Winters, J. (2010), 'The British History Online Digital Library: a model for sustainability?', *Bulletin of the Belgian Royal Historical Commission*, 176: 95–106. https://doi.org/10.3406/bcrh.2010.1078

Bradley, K. (2007), 'Defining digital sustainability', *Library Trends*, 56: 148–63. https://doi.org/10.1353/lib.2007.0044

Brown, S., Clements, P., Grundy, I., Ruecker, S., Antoniuk J. & Balazs, S. (2009), 'Published yet never done: the tension between projection and completion in Digital Humanities research', *Digital Humanities Quarterly*, 3(2).

Buddenbohm, S., Jong, M., Minel, J. & Moranville, Y. (2021), 'Find research data repositories for the humanities – the data deposit recommendation service', *International Journal of Digital Humanities*, 1: 343–62. https://doi.org/10.1007/s42803-021-00030-7

Carlin, C. (2018), 'Endings: concluding, archiving, and preserving digital projects for long-term usability', *KULA: Knowledge Creation, Dissemination, and Preservation Studies*, 2(1): 19. https://doi.org/10.5334/kula.35

Chowdhury, G. (2014), 'Sustainability of digital libraries: a conceptual model and a research framework', *International Journal on Digital Libraries*, 14: 181–95. https://doi.org/10.1007/s00799-014-0116-0

Christen, K. & Anderson, J. (2019), 'Toward slow archives', *Archival Science*, 19: 87–116. https://doi.org/10.1007/s10502-019-09307-x

Cummings, J. (2019), 'Opening the book: data models and distractions in digital scholarly editing', *International Journal of Digital Humanities*, 1: 179–93. https://doi.org/10.1007/s42803-019-00016-6

Drucker, J. (2021), 'Sustainability and complexity: knowledge and authority in the digital humanities', *Digital Scholarship in the Humanities*, fqab025. https://doi-org.eres.qnl.qa/10.1093/llc/fqab025

Edmond, J., & Morselli F. (2020), 'Sustainability of digital humanities projects as a publication and documentation challenge', *Journal of Documentation*, 76(5): 1019–31. https://doi.org/10.1108/JD-12-2019-0232

Eschenfelder, K.R., Shankar, K., Williams, R., Salo, D., Zhang, M. & Langham, A. (2019), 'A nine dimensional framework for digital cultural heritage organizational sustainability', *Online Information Review*, 43(2): 182–96. https://doi.org/10.1108/OIR-11-2017-0318

Ghosh, P. (2015, 13 February), 'Google's Vint Cerf warns of "digital Dark Age"', *BBC News*, https://www.bbc.co.uk/news/science-environment-31450389 [accessed 22 September 2021].

Goddard, L. & Seeman, D. (2020), 'Negotiating sustainability: building digital humanities projects that last', in Crompton, C., Lane, R.J. & Siemens, R. (eds), *Doing More Digital Humanities: Open Approaches to Creation, Growth, and Development* (London, Routledge), 38–57. https://doi.org/10.4324/9780429353048-4

Hughes, L.M. (2014), 'Live and kicking: the impact and sustainability of digital collections in the humanities', in Mills, C., Pidd, M. & Ward, E. (eds), *Proceedings of the Digital Humanities Congress 2012: Studies in the Digital Humanities* (Sheffield, The Digital Humanities Institute). https://www.dhi.ac.uk/openbook/chapter/dhc-hughes

Kálmán, T., Ďurčo, M., Fischer, F., Larrousse, N., Leone, C., Mörth, K. & Thiel, C. (2019), 'A landscape of data – working with digital resources within and beyond DARIAH', *International Journal of Digital Humanities*, 1: 113–31. https://doi.org/10.1007/s42803-019-00008-6

Kretzschmar Jr., W. & Potter, W. (2010), 'Library collaboration with large digital humanities projects', *Literary and Linguistic Computing*, 24.4: 439–45. https://doi.org/10.1093/llc/fqq022

Madden, F. & Kotarski, R. (2020), *Persistent Identifiers at the British Library* (London, British Library). https://doi.org/10.23636/1242

McLaughlin, J., Meredith, M., Pidd, M. & Rogers, K. (2010), *A Review of the AHRC Technical Appendix and Recommendations for a Technical Plan: Compiled by the Humanities Research Institute, University of Sheffield on behalf of the Network of Expert Centres TA Working Group*, 26 March 2010. https://www.dhi.ac.uk/projects/technical-appendix/ [accessed 30 March 2022].

Morawetz, L. (2019, 6 December), 'A first-hand encounter with MS. Don. e. 248', *Manuscripts from German-Speaking Lands: A Polonsky Foundation Digitization Project*. https://hab.bodleian.ox.ac.uk/en/blog/blog-post-17/ [accessed 30 September 2021].

Nowviskie, B. & Porter, D. (2010), 'The graceful degradation survey: managing Digital Humanities projects through times of transition and decline', *Digital Humanities 2010*. http://dh2010.cch.kcl.ac.uk/academic-programme/abstracts/papers/html/ab-722.html [accessed 1 October 2021].

Nowviskie, B. (2015), 'Digital Humanities in the anthropocene', *Digital Scholarship in the Humanities*, 30: Supplement 1. https://doi.org/10.1093/llc/fqv015

Oberhaus, D. (2012, 12 October), 'Amazon, Google, Microsoft: here's who has the greenest cloud', *Wired*. https://www.wired.com/story/amazon-google-microsoft-green-clouds-and-hyperscale-data-centers/ [accessed 5 July 2021].

O'Toole, J.M. (1989), 'On the idea of permanence', *American Archivist*, 52: 10–25. https://doi.org/10.17723/aarc.52.1.3x85283576r43387

Palaiologk, A., Economides, A., Tjalsma, H. & Sesink, L. (2012), 'An activity-based costing model for long-term preservation and dissemination of digital research data: the case of DANS', *International Journal on Digital Libraries*, 12.4: 195–214. https://doi.org/10.1007/s00799-012-0092-1

Pendergrass, K.L., Sampson, W., Walsh, T. & Alagna, L. (2019), 'Toward environmentally sustainable digital preservation', *The American Archivist*, 82: 165–206. https://doi.org/10.17723/0360-9081-82.1.165

Pickle, S. & Maron, N.L. (2013, 29 January), *Sustaining Our Digital Future: Institutional Strategies for Digital Content*. http://sr.ithaka.org/?p=22547 [accessed 1 October 2021].

Pierazzo, E. (2019), 'What future for digital scholarly editions? From Haute Couture to Prêt-à-Porter', *International Journal of Digital Humanities*, 1: 209–220. https://doi.org/10.1007/s42803-019-00019-3

Pitti, D. (2004), 'Designing sustainable projects and publications', in Schreibman, S., Siemens, R. & Unsworth, J. (eds), *A Companion to Digital Humanities* (Oxford, Blackwell Publishing), 471–87. https://doi.org/10.1002/9780470999875.ch31

Posner, M. (2015, 25 June), 'Humanities data: a necessary contradiction', *Miriam Posner's Blog: Digital humanities, data, labor, and information*. http://miriamposner.com/blog/humanities-data-a-necessary-contradiction/ [accessed 2 July 2021].

Prescott, A. & Hughes, L. (2018), 'Why do we digitize? The case for slow digitization', *Archive Journal*, Special issue 'Digital Medieval Manuscript Cultures', ed. Hanrahan, M. & Whearty, B. https://www.archivejournal.net/essays/why-do-we-digitize-the-case-for-slow-digitization/

Rockwell, G., Day, S., Yu, J. & Engel, M. (2014), 'Burying dead projects: depositing the globalization compendium', *Digital Humanities Quarterly*, 8(2).

Schmidt, D. (2019), 'The current state of the Digital Scholarly Edition and three challenges', pre-publication paper. https://www.academia.edu/37585331/The_Current_State_of_the_Digital_Scholarly_Edition_and_Three_Challenges [accessed 24 September 2021].

Smith, A. (2004), 'Preservation', in Schreibman, S., Siemens, R. & Unsworth, J. (eds), *A Companion to Digital Humanities* (Oxford, Blackwell Publishing), 576–91. https://doi.org/10.1002/9780470999875.ch37

Smithies, J. (2019, 7 February), 'The continuum approach to career development: research software careers in King's Digital Lab', *King's Digital Lab Blog*. https://kdl.kcl.ac.uk/blog/rse-career-development/ [accessed 7 September 2021].

Smithies, J., Westling, C., Sichani, A-M., Mellen P. & Ciula, A. (2019), 'Managing 100 digital humanities projects: digital scholarship and archiving in King's Digital Lab', *Digital Humanities Quarterly*, 13(1).

Stanford, E. (2018, 31 July), 'Booksquashing', *Medium. The Bodleian Libraries Blog*. https://medium.com/@bodleian/booksquashing-1d3fa800c253 [accessed 30 September 2021].

Stifter, D., Cnockaert-Guillou, N., Färber, B., Hayden, D., Ní Mhaonaigh, M., Tucker, J. & Yocum, C.G. (2022), *Developing a Digital Framework for the Medieval Gaelic World: Project Report* (Irish Research Council/Arts and Humanities Research Council Digital Humanities Research Network). https://www.qub.ac.uk/schools/ael/Research/ResearchinLanguages/imdorus/Publications/ [accessed 1 May 2022].

Turska, M., Cummings, J. & Rahtz, S. (2016), 'Challenging the myth of presentation in digital editions', *Journal of the Text Encoding Initiative*, 9. https://doi.org/10.4000/jtei.1453.

Varela, M. E., (2016), 'The archive as repertoire: transience and sustainability in digital archives', *Digital Humanities Quarterly*, 10(4).

Vlassenroot, E., Chambers, S., Pretoro, E.D., Geeraert, F., Haesendonck, G., Michel, A. & Mechant, P. (2019), 'Web archives as a data resource for digital scholars', *International Journal of Digital Humanities*, 1: 85–111. https://doi.org/10.1007/s42803-019-00007-7

Warwick, C., Terras, M., Huntington, P. & Pappa, N. (2008), 'If you build it will they come? The LAIRAH study: quantifying the use of online resources in the arts and humanities through statistical analysis of user log data', *Literary and Linguistic Computing*, 23.1: 85–102. https://doi.org/10.1093/llc/fqm045

Wilkinson, M.D. *et al.* (2016), 'The FAIR Guiding Principles for scientific data management and stewardship', *Scientific Data*, 3:160018, 1–9. https://doi.org/10.1038/sdata.2016.18

To cite the article: Tucker, J. (2022), 'Facing the challenge of digital sustainability as humanities researchers', *Journal of the British Academy*, 10: 93–120. https://doi.org/10.5871/jba/010.093

Journal of the British Academy (ISSN 2052–7217) is published by
The British Academy, 10–11 Carlton House Terrace, London, SW1Y 5AH
www.thebritishacademy.ac.uk