Digital Scholarship in the Humanities, 2022, **00**, 1–22
https://doi.org/10.1093/llc/fqac050
Full Paper

OXFORD

Digital History and the Politics of Digitization

Gerben Zaagsma *

Center for Contemporary and Digital History, University of Luxembourg, Luxembourg

*Correspondence: Gerben Zaagsma, Center for Contemporary and Digital History, University of Luxembourg, 11 Porte des Sciences, 4366 Esch-sur-Alzette, Luxembourg. E-mail: gerben.zaagsma@uni.lu

Abstract

Much has been made in recent years of the transformative potential of digital resources and historical data for historical research. Historians seem to be flooded with retro-digitized and born-digital materials and tend to take these for granted, grateful for the opportunities they afford. In a research environment that increasingly privileges what is available online, the questions of why, where, and how we can access what we can access, and how it affects historical research have become ever more urgent. This article proposes a framework through which to contextualize the politics of (digital) heritage preservation, and a model to analyse its most important political dimensions, drawing upon literature from the digital humanities and history as well as archival, library, and information science. The first part will outline the global dimensions of the politics of digital cultural heritage, focusing on developments between and within the Global North and South, framed within the broader context of the politics of heritage and its preservation. The second part surveys the history and current state of digitization and offers a structured analysis of the process of digitization and its political dimensions. Choices and decisions about selection for digitization, how to catalogue, classify, and what metadata to add are all political in nature and have political consequences, and the same is true for access. The article concludes with several recommendations and a plea to acknowledge the importance of digital cataloguing in enabling access to the global human record.

1 Introduction

Much has been made in recent years of the transformative potential of digital resources and historical data for historical research. Historians seem to be flooded with retro-digitized and born-digital materials and tend to take these for granted, grateful for the opportunities they afford. As the late digital history pioneer Roy Rosenzweig suggested two decades ago, historians 'may be facing a fundamental paradigm shift from a culture of scarcity to a culture of abundance' (Rosenzweig, 2003, p. 739). Yet when historians talk about digital abundance, they usually imply the Global North, specifically Western Europe and North America, and even there this abundance is still rarely questioned and qualified, let alone contextualized in time and space. Put simply, historians all too seldom reflect on why, where, and how we can access, what we can access in the digital age, how it affects historical research, and what its global dimensions might look like.

Few historians would deny that archives and libraries are repositories of carefully selected and curated collections and thus far from neutral: 'No archive is innocent', as Elizabeth Yale (2015, p. 332) succinctly observed. By the same token, the digitization of historical sources is not neutral. In a research environment that increasingly privileges what is available online, where non-digitized analogue archives are more and more referred to as hidden, and old-fashioned browsing is replaced by surgical discovery, we would do well to start imagining what a world of historical scholarship based upon digital resources looks like. The documentary record affects historical research in myriad ways, something historians readily acknowledge when traditional archives are concerned. Yet there seems to be a discrepancy between the uptake of digital resources by historians and their appreciation and understanding of how these are created and how, more generally, our online evidentiary basis is constituted and shapes historical scholarship.

Most historians acknowledge the benefits and opportunities afforded by increased access to retro-digitized sources yet the effects of digitization on historical research go beyond access. As digitization entails a selection of already selected analog materials, they find themselves facing old questions as they navigate new and unfamiliar digital environments. How do digital resources shape the historical themes, topics,

© The Author(s) 2022. Published by Oxford University Press on behalf of EADH.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Downloaded from https://academic.oup.com/dsh/advance-article/doi/10.1093/llc/fqac050/6702047 by Consortium Luxembourg, gerben.zaagsma on 16 September 2022

and debates that can be researched, and how might they influence research agendas more broadly? In what ways can they enable us to ask new research questions or open avenues of inquiry that challenge existing master narratives? Can digital resources facilitate research into transnational histories when most digitization projects are still nationally framed? To be sure, the basic questions of why, where, and how we can access what we can access, and which histories can (and cannot) be told with them, and by whom, have not changed after the digital turn. Yet as historians increasingly make use of digital resources in their research, they have become ever more urgent.

Increased access to retro-digitized sources does not imply completeness, even in the case of mass digitization. From the global perspective, heritage digitization is highly unequal and even in the Global North, many materials are not, and never will be, digitized. Digitization thus, first and foremost, entails selection. Archives, libraries, museums, and other heritage institutions select materials to be digitized based on a range of criteria. These include the need to preserve fragile materials, providing easy access to collection highlights and frequently used materials, the research value of specific collections and academic research agendas. Memory politics, public discourses on the past, and the articulation of a country's imagined national identity are of similar importance. At the same time, legal, ethical, and copyright issues also frame and constrain digitization strategies. Finally, given the costs involved, the availability of funding plays a crucial role in enabling digitization projects in the first place (Zaagsma, 2013a). In addition to the role these factors can play in the preservation efforts of official actors, bottom-up initiatives such as community archiving have their own agendas and can be animated by different concerns.

Against this backdrop, what are the politics of digitization and what are its implications for historical research? Many aspects of digitization can be considered political, from selection to modes of access to the political aspects of infrastructure. None of these is specific for our digital age and historical context is crucially important. Digitization is only the most recent technology of heritage preservation and reproduction, which has a history that dates back to the first uses of photography for research purposes in the early 20th century (Krumbacher, 1906). Similarly, historians, philosophers, archival, and heritage scholars have discussed the politics of heritage and the political dimensions of heritage preservation, as well as the relation between archives, social memory, knowledge, and power since at least the 1970s. And as long as archives have existed, the question of access has been key in determining who writes history. Finally, the politics of digitization should also be linked to discussions about censorship and the repression or persecution of historians, and the connections between history and human rights more broadly (De Baets, 2009, 2019).[1]

This article contextualizes the politics of cultural heritage digitization and outlines its most important political dimensions with the aim to encourage further debate on its implications for historical research.[2] In the first part, I seek to outline the global dimensions of the politics of digital cultural heritage, focusing *inter alia* on developments between and within the Global North and South, framed within the broader context of the politics of heritage and its preservation. The second part surveys the history and current state of digitization and offers a structured analysis of the process of digitization and its political dimensions. As an analytical framework, I will draw on the dimensions of archival and curatorial work outlined by Brown and Brown-Davis (1998) in their seminal 1998 article 'The Making of Memory'. As they argued, every element in the archival process—from collections allocation to cataloguing, classifying, and providing access—is inherently political. These elements can be easily transposed to the digital domain; choices and decisions about selection for digitization, how to catalogue, classify, and what metadata to add are all political in nature and have political consequences. The same is true for access. While many digital resources are freely available online, many a fruit of digitization still hides behind a paywall or is not accessible to those whose heritage is concerned. In the final section, I provide a number of recommendations and make a plea for more attention to the importance of digital cataloguing in enabling access to the global cultural record.

My intention in this article is not to critique digital history but to extend the critical questions that historians have always asked about their sources to the digital realm. The points it raises should be seen within the broader context of what could be called the 'critical turn' in digital humanities (Berry and Fagerjord, 2017), a shift that has taken place in the past decade from technological determinism to more attention for the methodological and epistemological consequences of the use of new digital technologies and approaches in humanities and historical research (for a concise overview of recent developments with regard to the latter, see Romein *et al.*, 2020). Thus, in recent years, scholars have discussed tool criticism (Van Es *et al.*, 2021), data criticism (Verbruggen *et al.*, 2017; Lee, 2021; Ortolja-Baird and Nyhan, 2021; Edmond *et al.*, 2022), algorithmic criticism (Verhaar, 2016; Burkhardt, 2017; Kerssens, 2017), code criticism (van Zundert and Haentjens Dekker, 2017), interface criticism (Ehrmann *et al.*, 2019), and digital source criticism (Föhr, 2017; Fickers, 2020, Wurthmann and Schmidt, 2020). In digital history, this critical turn has been conceptualized in

Downloaded from https://academic.oup.com/dsh/advance-article/doi/10.1093/llc/fqac050/6702047 by Consortium Luxembourg, gerben.zaagsma on 16 September 2022

different ways, from digital historiography (Sternfeld, 2011) to digital hermeneutics (Fickers and van der Heijden, 2020) to digital literacy more broadly (Jensen, 2021).

The literature that is used in this article was collected in the course of several years by (1) tracking the contents of a range of journals in digital humanities and history as well archival, library, and information science to identify relevant articles; (2) performing targeted literature searches on specific topics that are adressed in the various subsections in major literature databases such as Worldcat, JStor, and Google Scholar; and (3) through snowball search. In addition, the history of digitization section draws upon research in the Wayback Machine to reconstruct the International Federation of Library Associations and Institutions (IFLA)/UNESCO Directory of Digitized Library Collections. The aim was to identify key dimensions of the digitization of cultural heritage and its political dimensions from a range of disciplinary perspectives. Nonetheless, this article does not claim to be comprehensive as the literature on each of the issues it touches upon is vast. Morever, it does not escape its own subject matter: my own linguistic capabilities shaped this article in that most of the cited literature was written in English, though by Anglophone as well as non-Anglophone scholars, and to a lesser extent German and Dutch. The main ambition of the article is thus rather more modest, namely, to propose a framework through which to contextualize the politics of digitization and provide a model to analyse its most important political dimensions, and thus contribute to further discussion on what I regard as a key topic for the direction of historical scholarship in the 21st century.

2 The Politics of Heritage and Its Preservation

What societies choose to designate as heritage and thus deem worthy of preservation is profoundly political. By its very nature, the notion of heritage is intimately tied to the never-ending project of constructing the nation. As Stuart Hall argued in an influential lecture about heritage in Britain two decades ago: 'what the nation "means" is an ongoing project, under constant reconstruction. We come to know its meaning partly through the objects and artefacts which have been made to stand for and symbolise its essential values' (Hall, 1999, p. 5). Hall deliberately spoke of the Heritage as he sought to diagnose its political function in British society, while simultaneously observing that its rigid or fixed character was beginning to 'un-settle'. In a similar vein, Smith (2006) coined the term 'authorized heritage discourse' to draw attention to the intrinsic connection between heritage and the nation. At the same time, she highlighted the existence of competing discourses about what constitutes and comes to be regarded as heritage (Smith, 2006). Both Hall and Smith emphasized that heritage is never neutral but always negotiated and contested and is thus inherently political. In the past decade, the growing attention to the politics of heritage, generated by their work and that of many others, has resulted in the new and very much global field of critical heritage studies (Moore and Whelan, 2007; Harrison, 2013; Gentry and Smith, 2019; Smith, 2021).

Many actors are involved in the preservation and curation of heritage. Archives, as the heritage repositories most relevant to historians, fit within a much broader constellation that transcends the galleries, libraries, archives, and museums (GLAM) sector to include many public and private institutions, foundations, and countless individual and communal bottom-up projects. While national, regional, and municipal archives are mandated to collect official governance-related materials, community and counter archives are motivated by the desire to fill gaps and silences in the official record and to promote new and different perspectives. If national archives represent one end of the imaginary scale from official and authorized to unofficial and unauthorized, local community and counter archives represent the other.

Although a topic of debate ever since Foucault's *Archeaology of Knowledge*, the archival turn of the 1990s brought with it a renewed interest in the politics of the archive and its role in the production and reproduction of knowledge and power (Foucault, 1972; Derrida, 1995; Manoff, 2004; Ketelaar, 2017). As a result, it is now commonly recognized that archives enable certain (re-)constructions and visions of the past and, as such, are powerful actors, and potential gatekeepers, in the production of historical knowledge. This recognition has led to increased attention to archival silences (Carter, 2006; Cook, 2011; Thomas *et al.*, 2017) as part of a general broadening of the scope of archival scholarship to include its historical, social, cultural, and global contexts (Gilliland, 2016). At the same time, archives are not simply repositories or purveyors of memory, as so much contemporary discourse would have it. Instead, they provide (part of) the raw material that feeds into its construction. The notion that archives equate to memory obscures the role of mediation in how archival materials are converted into reconstructions of the past and the many ways in which this process influences individual and social memory. Indeed, while the importance of the nexus between archives and memory is commonly recognized, the role of archives in the formation of memory is fiercely debated among archivists themselves

Downloaded from https://academic.oup.com/dsh/advance-article/doi/10.1093/llc/fqac050/6702047 by Consortium Luxembourg, gerben.zaagsma on 16 September 2022

(Jacobsen, 2013). Be that as it may, few would disagree that 'archivists are active agents in constructing social and historical memory' (Cook, 2010, p. 170).

Historians have also debated the relationship between heritage, history, and memory, especially since Pierre Nora's *Les Lieux de mémoires* (Nora, 1984), often focusing on constructions of the nation and national identity (see, for instance: Berger and Lorenz, 2010; Berger, 2013). Crucial contributions to this debate have come from Francis Blouin and William Rosenberg, who published two seminal volumes that discussed archives and social memory (Blouin and Rosenberg, 2010) and the nexus between archives, historians, and authority (Blouin and Rosenberg, 2011). As they have argued, 'archives are not only places where historical narratives are woven from the documentation at hand. They have narratives of their own that need to be carefully if figuratively "read" before their materials can be fully appreciated and most effectively used'. And as they noted about digital archives, the archives' own narratives 'may be totally opaque, disguised by elegantly designed web-based locators and catalogues that are easy to "surf"' (Blouin and Rosenberg, 2011, pp. 209–210), a point that will be further explored below.

Such observations notwithstanding, debates about the nature of the historical record in the digital age, and its effects on historical research, are still mostly limited to digital history practitioners. Broader disciplinary conversations about the political aspects of digitized cultural heritage concerning questions of power, memory, and the nation often seem to be missing among historians, the very group that should be most concerned with it. This lack of debate is especially noticeable in the Global North, notably in Western Europe and North America. To be sure, this observation applies predominantly to retro-digitized sources, which are often seen as uncomplicated surrogates of their analogue originals, and much less to born-digital resources, whose radical newness as a historical source has perhaps made a reckoning inevitable (Milligan, 2019; Story *et al.*, 2020; Öhman, 2022).

This relative absence of debate among professional historians might be partly due to some early suggestions that the importance of the nation was diminishing in the digital age. Rosenzweig (2003), for instance, suggested almost two decades ago that 'some key aspects of the digital present [...] do not follow national boundaries and, indeed, erode them. If national archives were part of the projects of state-building and nationalism, then why should states support post-national digital archives?' (pp. 752–753). Yet the potential 'outcome' of digitization, the ability of digital archives to facilitate research that transcends national borders (Putnam, 2016), should not be mistaken for how digitization comes about. Rosenzweig (2003) wrote his words at a distinct moment in space and time, the USA, at the turn of the millennium. Twenty years later, there can be little doubt that national concerns and involvement still play crucial roles in efforts to digitize and (re-)construct the past.

All of this is not to suggest that only the national level matters nor is it intended as a reification of the nation. Yet we cannot escape its role in shaping the context in which digitization often takes place nor its importance as a cognitive framework of reference. The imagined nation is still of paramount importance as a focal point around which the construction of memory and identity, as well as struggles for the past, coalesce and play out, and debates about what heritage should be valued and preserved are had. This holds true even if we recognize that heritage is linked to the construction of multiple, and intersecting, identities, and social memory on local, regional, national, global, and supranational levels. Within countries, within and between regions, and between the Global North and South, we can observe how digital resources are always embedded in broader cultural and political configurations. The extent to which they are intended to serve or push specific political agendas varies considerably when looked at globally, but one thing seems clear: people's engagement with the past, and the struggle for 'memory' increasingly take place online and digital heritage has become an integral part of it.

3 The Politics of Digital Cultural Heritage

How can we conceive of the political dimensions of digitized cultural heritage against this broader backdrop of the politics of heritage? Digitization and the advent of the digital archival and library turn have led to old questions about heritage, power, and the production and construction of historical knowledge being asked anew (see, for instance: Ketelaar, 2008). Over the past decade or so, a burgeoning literature has developed that discusses topics such as the coloniality of the archive and its influence on historical knowledge production (Stoler, 2010; Gordon, 2014), displaced archives in various contexts and locales (Lowry, 2017; Sela, 2018; Anziska, 2019), and ethical approaches to indigenous knowledge creation and production (CARE Principles for Indigenous Data Governance, 2019; Guiliano and Heitman, 2019; Garaba, 2021). This literature has also provided a much-needed corrective to universalist ideas about what constitutes heritage or the archive (Garaba, 2015). The question of how *digitized* heritage features within that broader context is now firmly on the agenda in heritage studies, archival science, as well as the digital humanities (Lewi *et al.*, 2019; Mizruchi, 2020; Kessler *et al.*, 2021; Münster

Downloaded from https://academic.oup.com/dsh/advance-article/doi/10.1093/llc/fqac050/6702047 by Consortium Luxembourg, gerben.zaagsma on 16 September 2022

et al., 2021). Scholars in postcolonial digital humanities, for instance, have warned that 'the digital cultural record is in danger of telling the story of humanity from the perspective of the Global North alone' (Risam, 2019, p. 6; see also Grallert, 2021). Leaving aside the question of how one defines the digital cultural record, it is evident that the current state of digitization, when considered globally, is highly uneven and varied. At the same time, as we shall see, the fault lines cannot be drawn so neatly between the Global North and South, no matter how appealing that analysis might seem.

The politics of digitization also have their own ethical dimensions and implications. Zinaida Manžuch has discussed these in terms of their potential adverse effects on individuals and communities. A salient example is how selection processes, the organization, and the provision of access to indigenous heritage can conflict with the views and interests of originating communities. Inequalities in access become especially evident when originating communities cannot access their own digitized heritage. Conversely, ease of access can also pose its own challenges in protecting individual privacy (Manžuch, 2017; see also Rehbein, 2016). More generally, some scholars have noted that 'accessibility is more than just availability and findability' for experts (Francomano and Bamford, 2022, p. 24). Such issues notwithstanding, digitization can also allow for an 'ethics of the algorithm' in which voices beyond the canonical become heard, as Presner (2016) has argued in his thoughtful discussion of the USC Shoah Foundation's Visual History Archive (VHA).

The VHA's adherence to a preservation ethics involving bit-level preservation also raises another question, namely how compromises in quality (for reasons of funding or access) might result in losing the richness of testimony (Smith, 2016), an issue akin to the problem of working with interview transcripts, which tend to obscure silences, emotions, and other non-verbal elements. While the VHA is committed to creating digital master files of the highest possible quality, not every institution can commit the necessary resources to do so, raising further questions about what constitutes authenticity in the context of creating digitized objects. As Manžuch (2017) notes, authenticity is not only technologically but also culturally contextualized (pp. 9–10). Ma's (2020) recent study of a small-scale project digitizing Chinese marriage documents is a case in point. In foregrounding digitization ethics in the project, Ma highlights the importance of the concept of *translation* in cross-cultural digitization work.

The still limited debate among professional historians about how the online record is constituted in Western Europe (partly due to a comparative embarrassment of digital riches) could easily suggest that politics play a subordinate role in digitization. Yet this would be mistaken. In the European Union, the promotion of cultural heritage is inextricably linked to strengthening a sense of European identity (Whitehead *et al.*, 2020) and so is its digitization (European Commission, 2011; Niggemann, 2011; European Commission, 2019; Jakubowski *et al.*, 2019). The EU's flagship Europeana portal aggregates content from hundreds of GLAM institutions and 'produces a new form of cultural memory politics that converge national and supranational imaginaries with global information infrastructures' (Thylstrup, 2018, p. 57). Like Europeana, the more recent Time Machine project sees itself as contributor to a common European identity and history, a role that it unfortunately has not adopted without resorting to certain rather crude stereotypes. In a recent marketing video, the Time Machine project juxtaposed its laudable mission of rescuing Europe's cultural heritage, framed as akin to the efforts of the Enlightenment *encyclopédistes*, to various threats, including the destruction of heritage in the Muslim world by Islamists (with Islamism referred to as 'the wayward ideology affecting our times'). How the latter relates to *European* heritage is never explained, however.[3]

In general, most European national digital archives and libraries highlight their function as guardians of national heritage and promoters of the national past. This is especially striking in post-communist Central and Eastern Europe (Zaagsma, 2013a, pp. 20–21; Zaagsma, 2013b, pp. 207–208), where digitization can function as part of a reckoning with the communist past. At the same time digitization practices also reflect how difficult pasts are confronted in different European countries. A good example is the digitization of Jewish heritage in Germany and Poland in recent decades. Germany is the only European country where a concerted effort has been made to save the printed Jewish past for the digital future, funded largely by its national funding agency, the DFG (Zaagsma, 2013b). By contrast, private institutions support most of the digitization of Polish-Jewish heritage, only some of which receive public funding (Zaagsma, 2017). More broadly, the politics of digital heritage are highly visible in Eastern Europe and Russia, where historians face increasing political pressures and censorship (Golubev, 2021). Against this backdrop, the role of digital heritage in battles for the past in formerly communist Eastern Europe and Russia has recently attracted greater academic interest and was the theme of a special *Politics of e-Heritage* conference in Marburg in 2019.[4]

The way in which contested pasts manifest online can also be observed in other parts of the Global North. Ringel (2020), for example, has analysed the

Downloaded from https://academic.oup.com/dsh/advance-article/doi/10.1093/llc/fqac050/6702047 by Consortium Luxembourg, gerben.zaagsma on 16 September 2022

National Library of Israel's Time Travel digital resource, finding that 'the selection of archival records for digital preservation, the design of the search interface, and the crowdsourcing of metadata collection are all directing archive users toward certain narratives about Israeli history and away from others' (p. 1). Shammout (2018) has shown how digitization can play a role in engaging with displaced and seized heritage in his discussion of Palestinian audiovisual heritage. These examples also highlight the complications in discussing the politics of digitization in terms of an all too narrowly conceived Global North/South divide, which foregrounds economic development and thus the availability of resources. Yet while the latter might go some way to explain differences in levels of digitization overall, it does little to illuminate variations within the Global North resulting from questions about how the imagined nation is constituted, especially between West European nation-states (many of which are former colonizers), East European ones (which tend to see themselves as having been colonized by the Soviet Union), and settler-colonial states such as the USA, Canada, Australia, New Zealand, and Israel.

In the post-colonial Global South, the absence of a geographical overlap between nation and state has produced a complex duality. New postcolonial states in Africa, for instance, established new national archives based mainly upon the written legacy of colonial administrations. Indigenous knowledge and social memory were largely preserved in parallel through oral traditions which had long preceded and transcended colonial-era geographical boundaries (Katuu, 2015). The relation of these very different records of the past to the imagined nation and the real state is complicated, but now digitization enables the preservation of both. The interplay between digitization and (postcolonial) nation-building and national identity in postcolonial (South-)Africa is not straightforward, however, as Kahn and Tanner have pointed out. They have called for the development of 'digital collections that reflect an indigenous African identity, not an imagined Westernised one' (Tanner and Kahn, 2014, p. 125), following Lalu (2007), who argued forcefully for a 'politics of digitisation that [expands] what can be said about the history of liberation struggles in Southern Africa' (p. 42). Such calls point to the much broader post-colonial context in which digitization in Africa should be situated: the involvement of Western actors, both public and private, questions of access, privilege, and ownership interact in complex ways to create distinct power asymmetries between the Global North and South. As a result, the question becomes whose interests are served by digitization and to what extent more generally 'digitization is an extension of Western hegemony' (Ringel, 2020, p. 6) which helps reproduce power assymetries of knowledge production. Within this context, Chamelot *et al.* (2020) have described the various phases that digitization in Africa went through, as 'digital imperialism' in the 2000s, the 'complex of the digital savior' in the 2010s, and the appropriation of the discourse on 'endangered archives' most recently.

While post-colonial legacies can thus impact local heritage digitization efforts in adverse ways, digitization can conversely function as a means to address some of those legacies. A key characteristic of the historical record of much of the Global South is dispersal and displacement, and the fact that significant parts of it reside in the Global North. Digital repatriation, which allows 'surrogates of cultural heritage materials to be returned to [former colonial] source communities' (Christen, 2011, p. 187), has been used as a remedial approach to provide access, enabled by the ability of digitized heritage to cross national borders (it should be stressed, as Christen also does, that digital repatriation should *not* be seen as an alternative to replacing the return of actual physical objects). That same ability has allowed colonial history to become a key example of the transnational research potential of digitization, as colonial pasts can only be fully explored by digitizing materials from the Global South alongside relevant archival holdings in former colonizing countries of the North. A good example of this is the SlaveVoyages database, the oldest digital resource on Trans-Atlantic (and Intra-American) slavery, which has gone through several online versions since its first publication as a CD-ROM in 1999.[5]

Indeed, the transformative potential of digitization might be nowhere more salient than in the study of slavery (Rusert, 2017; Nowatzki, 2021) where sources are globally dispersed and located in both the Global North and South. A recent example of how digitization can enable broader perspectives that transcend the national is the Dutch National Archive's collation of slavery-related digitized sources from institutions in the Netherlands, England, Guyana, and Suriname.[6] This example also illustrates Marlene Manoff's observation that digital technologies 'vastly expand the possibilities for both creating and redressing archival silence' as digitization can do both (Manoff, 2016; Levi, 2021; Ortolja-Baird and Nyhan, 2021). In relation to this, Jeurgens and Karabinos (2020) have argued that digitization of colonial archives, in general, can push the boundaries of, and help decolonize, colonial archival legacies. One way of doing this is to involve the public, as the Surinamese Slave Registers Crowdsourcing Project has done (Galen, 2019). At the same time, 'digital history and digitized archives might facilitate truth and reconciliation processes', as Foley (2018) has argued in a recent paper on the history and memory of

Downloaded from https://academic.oup.com/dsh/advance-article/doi/10.1093/llc/fqac050/6702047 by Consortium Luxembourg, gerben.zaagsma on 16 September 2022

American slavery. It is to be hoped that the current momentum to explore the transnational history of slavery digitally will also help focus greater digital attention on sources that highlight the vast richness of African culture and history in general.

Finally, digitization can also assist in documenting heritage in areas of war and conflict, as the Syrian Heritage Archive Project has done since 2013 (Ballouz, 2018).[7] At the same time, digitized heritage itself can be fragile and susceptible to destruction. Following the Russian invasion of Ukraine in February 2022, an international initiative was set up to safeguard online Ukrainian cultural heritage (SUCHO, 2022).

Notwithstanding the potential outlined above, it is important to recognize that digitization requires difficult choices over the allocation of (often) scarce resources and should never be viewed as a simple panacea for all the documentary needs of historians, especially in the Global South. Chamelot *et al.* (2020, p. 113) have issued a stark warning in this regard, highlighting the crucial and continuing importance of prioritizing cataloguing work and non-digital preservation:

> There is now a greater risk that archives which have not been previously classified and inventoried will be lost because the slow work of digitization projects monopolizes the time of many archivists. There is then the very real prospect that the process of digitizing will endanger not only the activities of archivists but to a certain extent even their very profession. In other words, Northern digital aspirations, with the complicity of Southern governments driven by short-term financial interests, might ruin longer-term Southern efforts at analogue preservation.

4 The History and Current State of Digitization

As a means of preservation and reproduction, digitization is only the latest in a long line of technologies. The introduction of mechanical aids to preserve and reproduce archival and library materials dates back to the late 19th and early 20th centuries, though they only became ubiquitous in the 1930s and 1940s (Dommann, 2008). Since the introduction of photography in archives and libraries in the late 19th century, the photostat, microfilm, photocopying, and digitization have all been deployed in efforts to preserve and reproduce heritage materials (Friedrich, 2014). And then as now, selection determined what could be researched on the basis of those materials that were deemed worthy of reproduction. In that sense, the question of how digitization affects historical research could be more accurately rephrased to ask how heritage preservation and reproduction have always affected it.

No museum, library, or archive has ever captured the entire documentary record and preservation choices have always been made, even if these were not always well understood (Baker, 2002). Official archives are mandated to collect materials that document the governance and heritage of their respective societies, and archival appraisal processes determine what is kept and what is not. Microfilming or retrodigitization thus entails the preservation or reproduction of a selection of an already existing selection. When, for example, the practice of using microfilm for the preservation and distribution of research materials took off in the 1930s in the USA, archivists also confronted choices about which holdings should be prioritized (Morris, 1941). Newspapers were high on the list, and archival materials were also identified as a priority, especially by the Historical Records Survey that ran between 1935 and 1942 (Morris, 1941; Cady, 1990; Foster, 2010).

World War II provided a further incentive to these endeavours. Already in 1940, American archivists developed plans for the large-scale microfilming of European archival materials spurred on by the destruction of documentary heritage in Europe, while the early postwar years saw an increasing drive to microfilm in Europe itself (Proceedings: Conference on Microcopying Research Materials in Foreign Depositories, 1940). In the 1950s and 1960s, microfilm also began to be used extensively in archive-building in Africa, used by colonial administrations and postcolonial governments alike. During this period, research libraries in the former European colonial powers and the USA also sought to build up microfilm archives of colonial records (Curtin, 1960; Boylan, 1986; Howell and Scheven, 1997).

In all these microfilming endeavours, research value and preservation concerns were the leading selection criteria, in contrast to current digitization practice where access is often prioritized (Gertz, 1999).[8] When digitization took off in the 1990s, it ran parallel to microfilming and was not seen as a preservation technique that could ensure longevity in the way microfilm could. That access and user demand were prioritized is unsurprising given the prohibitive costs of digitization compared with microfilming, especially early on. In short, early digitization was about access, not preservation, and its value came to be assessed by more than its potential for facilitating (academic) research. Importantly, selection for both microfilming and digitization raised and raises questions about how existing silences, dominant narratives, and power relations in the existing archives were and are reproduced, and what new perspectives they might enable that allow these to be challenged.

Downloaded from https://academic.oup.com/dsh/advance-article/doi/10.1093/llc/fqac050/6702047 by Consortium Luxembourg, gerben.zaagsma on 16 September 2022

Given the above, however, what do we know about digitization today? Despite nearly three decades of activity, it remains surprisingly difficult to assess the global state of cultural heritage digitization as statistics are either incomplete or non-existent. In Europe, where heritage digitization is central to the European Commission's digital growth strategy, surveys have been conducted since 2007, which provide a reasonably comprehensive aggregate picture. Surveys exist for a number of countries in other parts of the world too, though different methodologies can complicate comparisons and we lack comparable data on a global scale. The first attempt to gain a truly global perspective was the IFLA/UNESCO Survey on Digitization and Preservation conducted in 1998, at the dawn of the era of (mass) digitization. As the published results were not split out by region, the report only provided an aggregate picture, and it was never followed up (Gould and Ebdon, 1999). The findings showed that perceived historical or cultural value, boosting access, academic importance, limiting damage, and preservation were the main criteria for digitization.

The survey formed part of an IFLA project that also created a Directory of Digitized Library Collections ('Digicol'), which was handed over to UNESCO in 2002. By the time it ceased development in 2006, the directory listed 293 collections (it can now be accessed through the Internet Archives' Wayback Machine).[9] The institutional category listed 242 projects, as shown in Table 1.

Bearing in mind that the survey data were self-reported, they clearly show the unequal global distribution of digitized resources at that time. The few resources from Africa in the directory, including one South African project on post-Apartheid periodicals and one Guinean project on music recordings, also indicate that the scale of digitization was limited compared with large European projects, such as the French digital library Gallica. Digicol was discontinued in 2006, however, and while UNESCO organized a conference in 2012 on *The Memory of the World in the Digital Age*, it was the first and last of its kind (Duranti and Shaffer, 2012; Edmondson *et al.*, 2020).

The most recent attempt to chart the state of digitization globally is the Economist Intelligence Unit's 2016 report *A New Age of Culture*. Commissioned by Google, it includes data from 243 heritage institutions located in twenty-two countries (McCauley, 2016). As the report notes, cultural heritage digitization 'is extremely uneven across countries and regions, and even within countries'. The Cultural Digitization Scorecard, which was developed to track the progress of digitization, measured the 'digital reproduction of artefacts or documents', and the definition of cultural digitization also encompassed 'cultural institutions' online presence, their web and mobile interaction with audiences, and the electronic access they accord to their archives, and the digital educational initiatives they undertake with external communities' (p. 3). As one might expect, Europe and North America rank highest. However, there are important nuances. For example, China and Indonesia rank highly on archive digitization and access, and digital access to cultural heritage in Mexico, Peru, and Thailand is also assessed as comparatively good.

As mentioned, Europe is the only region with comprehensive and systematically gathered information on the state of digitization, drawn from the so-called Numeric and eNumerate surveys conducted between 2007 and 2017.[10] Among the key findings of the last eNumerate Core Survey 4 from 2017, we find the following aggregate statistics:

- On average, 58% of the heritage collections has been catalogued in a collection database.
- 22% of the heritage collections has been digitally reproduced.
- 54% still needs to be reproduced.
- For an estimated 24% of the collections, there is no need to digitally reproduce collection items (as phrased in the survey itself).

These figures highlight several important points. First, European cultural heritage is not fully digitized and will likely never be. They also point to an important problem: much of our cultural heritage cannot even be discovered digitally through institutional collection databases. As the figures show, many relevant materials for historical research are not digitized and even discovering those materials online is often impossible. This highlights the often-overlooked point that access to cultural heritage in the digital age is not just about digitization but starts with access to their descriptive metadata and cataloguing.

5 Modelling the Politics of Digitization

In the previous sections, I have sketched several key aspects of the politics of digitization globally and briefly discussed the global state of cultural heritage digitization. As indicated, heritage is highly political in nature, and this is no different in the digital realm where the struggle for the past increasingly takes place. This plays out in both the Global North and South, a division that does have explanatory value when assessing the availability of resources for digitization, the effects of colonialism, and global asymmetries of knowledge production, yet should not obscure significant internal variation. Heritage digitization might have progressed most in Western Europe but even there

Downloaded from https://academic.oup.com/dsh/advance-article/doi/10.1093/llc/fqac050/6702047 by Consortium Luxembourg, gerben.zaagsma on 16 September 2022

Table 1. The Directory of Digitized Library Collections institutional listings split out by region

Continent/region	Archives	Libraries	Museums	Other	Totals	%
Africa	0	1	0	2	3	1
Arab States	0	0	0	0	0	0
Asia and the Pacific	0	32	0	4	36	15
Europe and North America	12	142	13	31	198	83
Latin America and the Caribbean	0	2	0	1	3	1
	12	177	13	38	240	100

Table 2. Political dimensions of archival and curatorial work (Brown and Davis-Brown, 1998) and their digital equivalents

Political dimensions of archival and curatorial work	Digital equivalents
Collections are allocated to different depositories, libraries, or archives in the name of efficiency in avoiding redundancy = allocation of control.	Which institutions digitize and control digital collections?
Collection development refers to decisions concerning what is and is not collected, what is merely stored but not catalogued (and hence made intellectually accessible), and what is thrown.	What is digitized and why? What is metadated and OCR'ed?
Cataloguing and classification refer to the organizational and intellectual description of what is held. Whose schema will be used?	How is it classified, and how is it metadated?
Circulation and access refer to decisions about who gets to see what, which is shaped in part by the classification system or categorial order.	How is access provided and mediated?

much remains to be digitized and politics shape the outcomes. As noted, what is digitized shapes the stories we can tell about the past, just as heritage preservation in general has always affected historical research and our engagements with that past. Nevertheless, digital technology amplifies these impacts in various ways, which will be discussed in the following section.

To perform a more structured analysis of the process of digitization and its political dimensions, I draw on the scheme proposed by the sociologist Richard Harvey Brown and the librarian Beth Davis-Brown in their seminal 1998 article 'The Making of Memory'. In their analysis, they explored four ideological and political functions of archival and curatorial work 'as these are understood by professional librarians and archivists in the United States' and showed how these necessarily entail 'deployments of power' (Brown and Davis-Brown, 1998, p. 22). These functions can easily be transposed to the digital realm, as shown in Table 2

As will be clear, this model follows the logic of the archival process and in that sense offers a broader and more expansive take than earlier important discussions upon which I also draw (Sternfeld, 2011; Jensen, 2021).[11] It also reflects Western notions of the analogue archive that dominate the Global North but have also been adopted in parts of the Global South. Nonetheless, I would argue that its essential points are relevant to global discussions about the digitized human record. Though it focuses on retrodigitization, most of the issues raised could also be applied to born-digital materials.

5.1 Who digitizes and controls digital collections?

If heritage photography and microfilming used to be the preserve of libraries and archives, a wider variety of institutional and non-institutional actors became involved as heritage digitization took off in the 1990s. Some of these, such as libraries and archives, were simply continuing their existing reproduction and preservation work with new digital means. Others, such as academics, lay enthusiasts, and community archivists, were new to the game, in part because digitization became increasingly affordable. In general, national, regional, and local GLAM institutions have always played important roles in driving digitization efforts but, as we saw, significant differences exist between and within the Global North and South depending on the extent to which heritage and its (digital) preservation are prioritized and funding is available. Retro-digitization of heritage can be a luxury that many countries cannot easily afford. To take a salient example, the European Union's policies of promoting digital cultural heritage as an economic asset contrast sharply with the more limited resources that states on the African continent can invest and where priorities concerning heritage preservation can be very different. Moreover, as will be discussed later, Barringer *et al.* (2014) have noted that 'political and economic inequality between North and South [...] has shaped not only the form and content of digital libraries, but also access on the continent to material about the continent' (p. 2).

Downloaded from https://academic.oup.com/dsh/advance-article/doi/10.1093/llc/fqac050/6702047 by Consortium Luxembourg, gerben.zaagsma on 16 September 2022

The Global North is home to several large-scale digitization projects and repositories, such as the non-profit Project Gutenberg (since 1971), the Internet Archive (1976), HathiTrust and Europeana (both 2008), as well as commercially driven or owned mass digitization projects such as Ancestry.com (launched in 1997), Google Books (2004), and Google Arts and Culture (2011).[12] Private actors, non-profit or otherwise, thus often play key roles in the creation of digital repositories with clear political consequences. Kizhner *et al.* (2021), for instance, have analysed the biases in aggregated digitized content on the Google Arts and Culture platform, the most global of these projects in outlook and content, and concluded that it 'inherits cultural biases of the print era' and, unintentionally, tends to favour the 20th century Global North. De Groot (2020) has suggested that historical information companies such as Ancestry.com 'have replaced the academy as gatekeepers and controllers of access to knowledge' (p. 25).

As Thylstrup (2018) has shown, the Google Books project was an important catalyst for the Europeana project (2008), which was also partly infused by French anti-American sentiment and inspired by one of the earliest national large-scale digitization initiatives, the French Gallica platform (1997) (pp. 11–15). The kind of mass digitization that Google Books represents was a typical phenomenon of, in, and for the Global North, whereas digitization has generally been more selective and small scale in the Global South. As a result, political aspects can become more salient when selection for digitization is involved. The age of mass digitization in the Global North seems by and large over, however (Crymble, 2021, pp. 54–87). Google Books no longer adds content and progress within the framework of Europeana, which aggregates digital content from European heritage institutions, has advanced piecemeal in recent years. In fact, the Europeana strategy 2020–25 is directed towards improvement and realizing the potential of digitized heritage.[13] The exception to these developments might well be Ancestry.com, which continues digitization at scale in private–public collaboration with heritage institutions.

A wide assortment of non-state GLAM institutions, NGOs, and charitable foundations are also involved in digitization efforts. The Endangered Archives Project (EAP), which is administered by the British Library, 'facilitates the digitization of archives around the world that are in danger of destruction, neglect, or physical deterioration', and is perhaps the best-known global actor.[14] It is tempting to see the EAP as a unique enterprise that could only exist in our digital age. Yet, its mission resembles the efforts of the International Institute for Social History in Amsterdam (founded in 1935), which for decades played a crucial role in preserving the often-endangered archives of the global labour movement.

Digitization, as we saw, is also a relatively recent addition to the decades old practice of community archiving which has ranged from small, local efforts to mass digitization projects and similarly reflects a long tradition of bottom-up preservation (Bastian and Flinn, 2020). The Marx Memorial Library in London, for example, began as a community-led initiative in 1933. And during World War II, the famous Oyneg-Shabes underground archive in the Warsaw Ghetto led by Polish-Jewish historian Emmanuel Ringelblum functioned as a counter archive, existing to document and resist erasure from the 'official' Nazi record (Kassow, 2007). Digital community archives are not necessarily small as is shown by the YIVO Vilna Online Collections Project which has digitized the pre-war New York and Vilnius library and archival holdings of the YIVO Institute for Jewish Research (itself a community project dating back to 1925), thereby creating the world's most important digital archive of East European Jewish History.

The Vilna Collections project is also an excellent example of how the dilemma of transnational heritage preservation (who takes responsibility when heritage is often engaged with on national levels) can be confronted in the digital age and communities themselves step in to take charge. Access to and the ability to mobilize resources, however, differ across communities. Online Roma/Sinti heritage, for example, is scarce and mostly enabled by public, often European, funding. At the same time, small-scale community archiving can grow into bigger endeavours drawing upon diverse funding resources, as the example of the Black Cultural Archives in London shows.[15] As might be clear from the latter example, community archiving can be seen as a form of counter archiving which can be part of a wide range of activist causes and agendas (Vukli and Gilliland, 2021) and might include academic collaborators too. When digitisation took off in the 1990s, some scholars and universities began to work with communities to digitise their heritage, such as the Multicultural Archives database created at Virginia Commonwealth University (Whaley, 1994). Recent examples in the digital realm are the Turkish Geziparkarsiv, created in the wake of the harshly suppressed Gezi Park protests in Istanbul in the Spring and Summer of 2013 (Ozban, 2014), and the crowd sourced Daraa archive that documents the Syrian War (Saber and Long, 2017).

One of the oldest forms is the type of small-scale digitization performed by individuals or groups without academic motivation. One need look no further than the many websites created by collectors and fans of all stripes. While such endeavours are sometimes

Downloaded from https://academic.oup.com/dsh/advance-article/doi/10.1093/llc/fqac050/6702047 by Consortium Luxembourg, gerben.zaagsma on 16 September 2022

dismissed as non-professional, it is too often forgotten what foundational and extraordinary work was and is done by hobbyists. 'Amateur' archives have recently become the subject of increasing academic attention, for instance in Abigail De Kosnik's work on 'rogue archives', and Sarah Baker's analysis of do-it-yourself approaches in her work on popular music heritage (De Kosnik, 2016; Baker, 2018). Finally, on the smallest scale, we have private archiving in the form of digital photography kept for personal research, a practice that is, in fact, older than is often assumed. In 1906, for example, the German classicist Karl Krumbacher published a booklet in which he lauded the potential uses of photography in the humanities, focusing on (art) history and classical studies (Krumbacher, 1906; Fretwell, 1908). Similarly, microfilming was enthusiastically promoted by some as a means to gather a personal library, along the lines of what Krumbacher had advocated, resembling today's use of digital photography in archives (Smisor, 1942). Such private digitization can become public when scholars start sharing their photographs online or upload them to repositories such as the former Parallel Archive, launched in 2008 and developed by the Open Society Archives (OSA) at Central European University.[16]

5.2 What is digitized and why? What is metadated and OCR'ed?

Having considered who digitizes and what forms of digitization exist, the next question to ask is what is being digitized and why, and how it is made intellectually accessible? As discussed, digitization first and foremost means 'selection' informed by criteria ranging from preservation, user demand, and research value to broader discourses about a society's past and memory politics, all of which feed into funding decisions. In their discussion of digital historical newspaper archives, Hauswedell *et al.* (2020) call this 'explicit' selection (intentional in- or exclusion) as opposed to 'implicit' selection or 'the way that users' engagement with digitized material is mediated by the search and retrieval possibilities that are or are not made available to them via a given interface' (n.p.). The latter will be discussed separately in the section on access and mediation, while in the following I will focus on explicit selection.

While scale can easily obscure the fact, it bears remembering that retrodigitized archives are selections of the human record or, as Mills (2015) has called them, 're-selections'. In an important recent article, Coburn (2020) has suggested that historians are quite aware of issues about what he calls 'digital selectivity', broadly defined as an awareness of the limitations of digital resources. His user study is an important empirical corrective to some of the earlier warnings about historians' lack of critical awareness when working with digital resources. However, awareness among historians when they work with specific digital resources is different from asking how their digital source base is constituted and selected in the first place, and how digital resources might influence and shape research agendas (Hobbs, 2013; Milligan, 2013).

Selection works on different levels. Europeana, for instance, does not digitize but aggregates the digitized content that others selected and, as Scholz *et al.* (2017) notes, 'in principle publish[es] all types of content that museums, libraries, archives, and audiovisual institutions based on their digital strategy consider [...] important enough for publication in Europeana'. Nonetheless, the political ambitions in its last published content strategy (2017) are reflected in the aim to showcase objects that reflect a 'European cultural heritage and are testimony to a common European civilization' and confirm its ambition for the 'digital construction of Europe's collective memory' (Valtysson, 2012).[17] Selection is subsequently outsourced to participating European GLAM institutions that showcase their national, regional, or local digital heritage on the platform while simultaneously reframing it as European, in line with Europeana's mission. A good example is the Croatian Cultural Heritage project which highlights the importance of digitization for 'the networking and presence of the Croatian cultural heritage in the European and regional cultural contents networks, for the preservation of cultural diversity, and for the use of cultural contents in education, tourism, and other services activities'.[18]

Many countries in Europe have national digitization strategies that invariably stress the importance of the 'national' in selection procedures, albeit to varying degrees. The national Dutch Metamorfoze programme, for instance, funds the digitization of collections that 'concern the Dutch language, culture, and history in the broadest sense of the word and are of national or international importance'.[19] There is no publicly available information, however, on how the latter is assessed. This absence of transparency is often striking when surveying national digital archives and libraries, although some clearly articulate their aims. The National Library of Poland's digital library Polona, for instance, claims to offer the 'most valuable treasures of the Polish culture and history' in its mission statement before adding that 'there are also objects representing an entire cross-section of subject fields and epochs'.[20] In many other cases, such as the digital collections of the National and University Library in Zagreb or the US Library of Congress and National Archives, to name only two examples, we find not much more information than the remark that historical value is a criterion for selection. Even if more information on national or institutional digitization strategies can be found,

Downloaded from https://academic.oup.com/dsh/advance-article/doi/10.1093/llc/fqac050/6702047 by Consortium Luxembourg, gerben.zaagsma on 16 September 2022

these often tend to be vague, just as the mission statements and about sections of the digital resources they fund. As the recent example of the National Library of Scotland shows, however, some institutions seek to increase transparency with regard to the criteria used for selection for digitization (National Library of Scotland, 2018). Nonetheless, much more revealing of what heritage is considered of special value or importance, and how the imagined community is constructed online, are the shifting thematic collections and exhibitions offered by many national digital libraries and archives.

The above prompts the question of which canons are reproduced online, a question to which scholars are now increasingly turning in Western Europe too.[21] The 'canon question 2.0' is not entirely new, however. In 2012, for example, Prescott (2012) provocatively warned of the danger of digital humanities being 'party to a concerted attempt to reinstate an outmoded and conservative view of the humanities', and Hitchcock (2012) noted that 'if we are not careful, we will see the creation of a new "naturalisation" of human thought based on the narrowest sample of the oldest of dead white males'. Against this backdrop, Mihai (2015) has shown how selection, negotiation, and framing are reflected in the online works of Shakespeare. In addition, in a short article on digitization and social history, Gomez (2019) analysed how British working-class history fares online, pointing out how 'patron interest' and appeal can determine what gets digitized or not, especially when budgets are limited or shrinking. Gomez (2019) also pointedly notes that 'the old politics of who deserved to be historicized has become the new politics of who deserves to be digitized' (p. 400).

Needless to say, the more the (national) past is contested, the more its digitization becomes a site of contestation. This logic can be observed in two distinct locations: the post-communist states of Eastern Europe and Russia, and the postcolonial Global South, for instance, in post-Apartheid South Africa. As Golubev (2021) has shown, in Russia, the 'state has sought to use digital archives to firmly re-establish itself through its institutions as the main authority in the production of knowledge of Russian history' with the Federal Archival Agency (Rosarhiv) a key player in selection (p. 365). In other cases, such as the Czech Republic, the digitization of communist police archives became a highly contested realm and site of reckoning with the communist past in which various political actors and parties sought to assert their influence (Gjuričová, 2019).

Such struggles for the past are compounded when resources are scarce and digitization strategies might be absent (Balogun, 2018; Pandey and Kumar, 2020). In a seminal article addressing key issues around selection for digitization in Africa, Pickover (2014) has argued that 'content selected for digital projects should not be a new form of colonial "discovery" of the African "other", who through the selection of content for digitization are relocated from an invisible space to one that is hyper visible' (p. 9). She explores the example of the Digital Innovation South Africa (DISA) project aimed at digitizing liberation struggle materials and the 'awkward one-dimensional repression/resistance narrative' that it seemed to push that is 'mainly aimed at an undergraduate studies audience in the USA' (p. 10). This points to the danger of contested heritage becoming a playground for academics from the Global North who might project their own agendas on countries abroad without properly considering the actual concerns of those communities whose history is at stake. All of this also points to the political and ethical issues that can emerge when foreign funders and partners become involved in digitizing African heritage, not only when selection is concerned but also access (Anderson and Hart, 2016; Chamelot *et al.*, 2020).

In the Brown and Davis-Brown scheme, collection development refers not only to the question of what is stored but also to what is (or is not) catalogued and thus made intellectually accessible. Viewed from the perspective of digitization, this is akin to having digitized materials that are not further described, structured, or metadated, and Optical Character Recognition (OCR) could be included here too. The digital equivalent is not entirely similar to the analogue one, though, since some form of cataloguing will always take place in the form of minimal description or structuring. One can think here of digitized books in the Google Books or Hathi Trust corpora which are stored but cannot be accessed (or, if so, only in snippets) for copyright reasons. Another example is that of newspaper databases, which only contain non-OCR'ed scans of the digitized pages and are thus not text searchable. Most prominently, we see this with archival documents written by hand or in languages for which OCR does not (yet) exist, or with documents containing images. In both cases, access depended until very recently upon descriptive metadata. However, as we shall see below, several technological developments, often subsumed under the header Artificial Intelligence (AI), are rapidly changing how digital archival content can be opened up (Romein *et al.*, 2020).

5.3 How is it classified and how is it metadated?

As Brown and Davis-Brown (1998) assert, 'classifications never emerge solely from the material to be classified since our ways of defining the material itself are shaped by the dominant intellectual or political paradigms through which we view it' (p. 25). Well-known archival principles—such as *respect des fonds* and *respect de l'ordre* (original order)—and library

Downloaded from https://academic.oup.com/dsh/advance-article/doi/10.1093/llc/fqac050/6702047 by Consortium Luxembourg, gerben.zaagsma on 16 September 2022

classification schemes are neither neutral nor universal but have a long history (Trace, 2020a). It is now commonly acknowledged that classification schemes help produce and reproduce knowledge and the cultural, social, and political ideas and positions that underpin them (Rafferty, 2001; Youn, 2016; Turner, 2017).

In the process of conversion from the analogue library or archive to its digital representation, an already existing classification is reproduced, re-classified, or both, and is accompanied by the 'challenging task for archivists [...] to bring archival context and digital content into an organic whole in digital archival representation' (Zhang, 2018, p. 100). If digitization entails the digital reproduction of a selection, its subsequent classification entails reclassifying the already classified. This process is not to be confused with the inherent bias of the original order of the analogue archive; historians know and understand the constructed nature of the archives they work with. Yet the digital archive is a new and different type of construct that imposes new orders on the material. How structure is provided; thematic collections are chosen, titled, and arranged; and keywords and other types of metadata are chosen and added—as well as how extensive these are—all reflect social, cultural, and political dispositions. How, then, can digital archives and libraries be (re)classified without reproducing the power relations and political visions embedded in analogue classifications while, quite literally, maintaining a workable order?[22] In the case of a digital library, we must ask whether any classification can (or should) be achieved that resembles the carefully curated (though inevitably biased) collections of analogue libraries that allow for other relevant books on the same subject to emerge. How to digitally enable serendipity?

The translation of classification schemes into metadata schemes is a critical factor in mediating user access to the contents of digital resources. In its most basic form, metadata determine findability: if a picture of a boat does not contain the word 'boat' somewhere in the descriptive metadata, no user will find it (unless image recognition is part of the search implementation)'. Thus, the quality of the metadata matters. The 2016 EIU global survey noted that poor quality metadata prevents users from locating and searching relevant materials (McCauley, 2016; see also Traub *et al.*, 2015; Masenya, 2021). The example of the boat might seem simplistic, but it can illustrate a much bigger point about the coloniality of metadata and the ways in which certain narratives are reproduced. Jeurgens (2019), for example, has discussed how the Dutch military intelligence service gathered large amounts of materials from Indonesians during the independence war of the late 1940s. These can be found in the Dutch National Archives in The Hague and are mostly described with the terminology of the intelligence services, such as a section in the inventory on 'Republican gangs, terror, sabotage, etc.' (Jeurgens, 2019). Needless to say, archival description and metadata thus impose a distinct view of the past, in this case, that of the former colonizer. That view will be reproduced when digitization simply adopts the analogue classification.

There are many examples of the social, cultural, and political biases inherent in the digital reproduction of existing classification schemes and the coloniality of metadata (Fernandez, 2018; Baker, 2019). Jeurgens (2019) also provides an example of how to address these issues: in the Canadian-based Reciprocal Research Network, heritage institutions collaborate with originating communities who can assign their own descriptive metadata and also determine who gets access. Indeed, engaging with communities is crucial in addressing such issues, as another example, the UK-based Making African Connections Digital Archive, shows.[23]

Recent advances in AI offer much potential for curating digital archival data (Colavizza *et al.*, 2021) and could bypass some of the human biases discussed above, even if challenges abound (Venkata *et al.*, 2021; Van der Werf and Van der Werf, 2022). Rapid developments in handwritten text recognition are a potential game-changer in opening up information contained in handwritten documents (Muehlberger, 2019) while computer vision techniques make images and drawings accessible for search and analysis (Stork *et al.*, 2019; Wevers and Smits, 2020). The new field of computational archival science now promises to address archival data curation challenges at scale (Stančić, 2018; Marciano *et al.*, 2019; Hedges *et al.*, 2022). A key question here is what type of technology is being employed to assist in tasks like classification and what, if any, domain expertise still plays a role in their application (Rolan, 2019). In the latter case, it is easy to see how previous schemes (with all their potential biases) might be reproduced. If not, the question becomes how the algorithms were trained and which potential bias this might introduce (Trace, 2020b).

5.4 How is access provided and mediated?

The critical achievement of digitization is undoubtedly the vast expansion of the user base for historical source materials, from dedicated researchers and enthusiasts working in the archive or library to much broader and more diverse online audiences of various stripes. Of course, the question of what can be accessed by whom is hardly new, but digitization has thoroughly reshaped it. As Winters (2019) put it, 'historians have always had to work with what is available to them, but in the sense of what has survived rather than what is easiest for them to access and use' (p. 295). Access, which I

Downloaded from https://academic.oup.com/dsh/advance-article/doi/10.1093/llc/fqac050/6702047 by Consortium Luxembourg, gerben.zaagsma on 16 September 2022

understand here to include how it is mediated through interface design and search and retrieval options, comes in many forms. Digital resources can be available online for free or through a paywall. Access can also be mediated through library membership, which allows access to the digital resources that the library subscribes to. In the case of academic libraries, access will be more privileged than when public libraries are concerned, given that membership is not open to all. Another question is whether access is provided online or only through an archival or library reading room, which would necessitate a trip to the institution. Finally, yet another dimension comes in the form of decisions being made by librarians about which digital archives from commercial publishers to subscribe to. Here again, it is only library subscribers who subsequently have access.

If digital resources are made freely accessible online, this allows for easy access to heritage beyond national borders. Of course, this can also apply to microfilm collections, but these still must be consulted in libraries or archives (unless they are digitized). Access is, of course, also about who has access to technology: personal computers and the Internet are often common goods in the Global North, yet this can be very different in the Global South. We are currently in the paradoxical, and ethically highly problematic, situation where digitization can rectify archival silences concerning the Global South or the post-colonial record, yet its fruits can often only be fully enjoyed and accessed in the Global North.

Apart from the fundamental question of who can access which digital resources where, we also need to ask how interface design, the implementation of metadata schemes and search and retrieval options mediate access to the contents of digital resources. Furthermore, the question of 'which' contents are made available, and in what ways, is becoming increasingly important as 'collections as data' approaches, that open up digital contents to computation, are gaining ground (Padilla *et al.*, 2019; see also Moss *et al.*, 2018). In the past decades, the traditional role of archives and archivists, libraries and librarians, as mediators has been extended and, to various extents, even supplanted, by online catalogues and inventories, digital libraries, and archives. This shift from 'human-mediated archives [...] to computer-mediated archives' does not only alter the role of archivists/librarians; these new modes of documentary access actively shape historical research (Hedstrom, 2002, p. 22). If access used to be mediated through inventories, card catalogues, and archivist knowledge, all of which were equally 'guilty' in the construction of archives as power-laden sites of the record of the past, metadata schemata, search functionality, and interfaces now demand our attention as the new mediators.

How knowledge is represented visually through the interface is also highly political. To illustrate this, Drucker (2016) has provided an illuminating thought experiment of

> a faceted interface that allows us to see a collection of artefacts from various perspectives. Consider an online museum displayed through a set of filters that organize and select objects according to different criteria: the knowledge of an original collector, a scheme of standard metadata from a western perspective, in accord with a classification scheme from an indigenous community, and so on. Each structuring organization offers a different argument, [and a] different set of hierarchies and values in its presentation (p. 62).

A good example of integrating indigenous knowledge systems into online collections is the Mukurtu CMS, collaboratively developed with the Warumungu Aboriginal community in Australia. The Mukurtu project can be seen as part of a broader turn 'to the development of digital platforms, databases, interactive augmented and virtual reality exhibitions [...] and virtual repatriation [...] to return cultural information and objects to their respective indigenous communities' (Whaanga and Mato, 2021, p. 448). The project was meant to answer several urgent questions, as explained by one of the academics involved, Christen *et al.* (2018): 'How do we create structures that allow for those cultural protocols, the knowledge systems, that already exist? [...] how do we create archives that respect those different ethical codes for managing, circulating, sharing, and duplicating what's in those records?' (p. 11). The complexities of addressing these concerns are well illustrated by the history of the Popol Vuh Online project, focusing on Central American indigenous knowledge (Espinosa de los Monteros, 2019). Such projects also touch upon questions and concerns about the ethics of making available online those materials that document violent pasts, colonial, or otherwise, and how to do so (Presner, 2016; Odumosu, 2020).

Search is perhaps the most important mediator and, as is often argued, almost always follows the Google paradigm of the single search box. The problem of search and the question of how it affects historians' research practices has received much scholarly attention among digital humanists in recent years (Huistra and Mellink, 2016). Kerssens (2017) has explored the positivist 'ideological nature of algorithmic search' and the supposed objectivity of search engines in general. Jensen (2021) has pointed to the divergence of the

Downloaded from https://academic.oup.com/dsh/advance-article/doi/10.1093/llc/fqac050/6702047 by Consortium Luxembourg, gerben.zaagsma on 16 September 2022

interests of general users, who might favour the single search box, and those of academic researchers (p. 258). Moreover, Coburn (2020) notes as one of the results of his interviews the opinion 'that ease of use could be a bad thing for rigorous historical research' (p. 10).

A fundamental problem here is the loss of context when surgical discovery enables immediate access to the micro-level without first passing through the meso/macro levels (see also Zaagsma, 2013a, pp. 26–27). At the same time, as Coburn (2020) notes, 'digital archival collections prevent the kind of general browsing that historians perform in physical archives. This, in turn, reduces the opportunities for chance finds' (p. 7). Crucially, this was a concern only among the older historians in Coburn's sample who contrasted such serendipity with 'the more surgical form of data retrieval that occurs with digital archives' whereas 'early career researchers [...] were more inclined to talk about the benefits of keyword searches and surgical discovery' (pp. 7–8). This indicates that selection bias is compounded by a loss of (awareness of) context. If younger generations tend to gravitate more towards digital resources and, being used to surgical discovery, do not see the value of browsing materials, then teaching the crucial importance of context becomes paramount. Within this context, Coburn even suggests that digital platforms may want to provide 'an interface that makes a difficult information—retrieval process a feature of the user experience' (p. 10).

One way to address this issue is to design interfaces that convey more information about the composition of digitized holdings and reveal the biases inherent in digitized collections. The Netherlands Institute for Sound and Vision, for instance, maintains a 'Visualizing the Archive' page, which provides insights about the distribution of archive items over collection types for the whole archive, over carrier types, the proportion of archive items that has been digitized, and their distribution over time.[24] Special pages subsequently provide more information on television, radio, and music. Digital humanities scholars are also experimenting with interface designs that reflect scholarly uses. For example, the interface of the Impresso project's database with 200 years of digitized newspapers from Switzerland and Luxembourg was co-designed in close collaboration with historians.[25]

6 Concluding Remarks

What should we make of the questions and issues described and raised above? The gap between the availability of an increasing number of digital resources and tools and their rather haphazard uptake by historians is by now well known. At the same time, professional historians make increasing use of digitized cultural heritage (Sinn, 2013; Sinn and Soares, 2014). Leaving aside the question of how online digital resources shape public perceptions of history, the crucial issue is how the availability of and access to digital cultural heritage affect historical research. As we have also seen, the democratizing potential of digitization should not lead us to ignore the obvious challenges and broader ethical questions about global inequity when access to digital cultural heritage is concerned. Yet, it is important to move beyond listing pitfalls and overly simplistic takes. Digitization is not inherently good or bad, democratic or undemocratic. Digitization can be used by states as a means to promote official master narratives, it can reinforce global power and knowledge assymetries, but it can also enable untold and ignored stories to be heard and told. Nonetheless, as many examples cited in this article highlight, even in the case of manifestly unheard or silenced heritage, struggles for power and control can still, and often do, abound. Obvious as it may seem, it is thus first of all necessary to underscore the basic point that professional historians should approach the analogue and digital cultural record in the same critical manner.

The challenges I have discussed can be addressed in several ways. To start with, and echoing the call of Hauswedell *et al.* (2020), we urgently need what could be called 'digital cultural heritage transparency guidelines' that encourage institutions to provide more information about how their digital resources are constituted and were created. Who funded the project? What selection criteria were applied? What classification and metadata schemes were used? How is access provided? What choices were made concerning interface and search options? And crucially (if often ignored), what relevant offline sources could be of interest to researchers? An increasing number of institutions is beginning to provide (parts of) this information, a development that will hopefully see wide adoption across the GLAM sector. This would ideally also include collection overviews based on an analysis of collections data such as provided by the Netherlands Institute for Sound and Vision mentioned above (note 25) to indicate the biases present in a given digital resource.

Much work also remains to be done in the realm of education to promote and achieve what Jensen (2021) has called digital archival literacy. The engagement with the digital and digital resources in several recent skills manuals for history students, for example, is still decidedly mixed: Kamp *et al.* (2016, 2018) and Eckert and Beigel (2019) pay relatively scant attention, while Bérard *et al.* (2020) and Corfield and Hitchcock (2022) seek to integrate digital resources and their speficities in more comprehensive ways, Meanwhile, textbooks such as the recent *Doing Digital History* fill a crucial

Downloaded from https://academic.oup.com/dsh/advance-article/doi/10.1093/llc/fqac050/6702047 by Consortium Luxembourg, gerben.zaagsma on 16 September 2022

gap in this respect for professional historians (Blaney *et al.*, 2021). It is also vital that historians engage much more with digital archivists, librarians, and public historians who are at the forefront of dealing with and thinking through how digital resources change the online heritage landscape. Such engagement might include involving historians in co-designing interfaces soliciting more detailed insights from users with regard to the information they need to better assess digital resources and contextualize their search results.

Furthermore, and perhaps paradoxically, I would suggest that we need to pay much more attention to the challenges facing the analogue archive. We currently face a paradoxical situation concerning digitization and heritage in both the Global North and South. It might be increasingly common to describe non-digitized heritage as hidden, but that label suggests digitization as a miracle cure that will solve all supposed problems of accessibility. The real problem, however, would seem to be that much of our cultural heritage cannot even be discovered digitally through institutional collection databases. More attention should be paid to online cataloguing before digitization, in the case of materials where neither is done, as well as linking online archives to catalogues or descriptive information about offline resources. As we have seen, the challenge of cataloguing remains of fundamental importance, not only in many parts of the Global South, but in the Global North too.[26]

This points towards a final consideration. Research on digitization tends to focus on specific national, or perhaps supranational, cases but rarely adopts a global comparative perspective. In view of this, it seems high time to ask how we can better connect conversations about digital heritage on a truly global level, and include a broader range of actors, perspectives, and concerns than currently seems the case. Above all, then, it is important to start rethinking how we value and integrate both the global and the non-digital into our digital history discourses and practices.

Notes

1. See also the crucial documentary work undertaken by De Baets through the Network of Concerned Historians: http://www.concernedhistorians.org/content/home.html.
2. The present article is not focused on the politics of digital humanities as such. Instead, it seeks to assess the impact of digitization on historical research practice. It is important to differentiate this question from the much broader question of cultural criticism in the digital humanities and its linguistic and geopolitical dimensions (Fiormonte, 2012, 2017; Liu, 2012) or issues surrounding the politics of citation (Wodziński, 2022). This article is not an attempt to advance a certain politics for the digital humanities, nor do I wish to promote the idea that scholars in the field have a special advocacy role to play. Instead, I would argue that all those working in the humanities should be mindful of the ethical and political implications of their work in relation to the discipline as a whole and society at large.
3. See the trailer on the Time Machine website: https://www.timemachine.eu/trailer/.
4. *Politics of E-Heritage: Production and Regulation of Digital Memory in Eastern Europe and Russia*, 3–4 June 2019, Herder Institute for Historical Research on East Central Europe in Cooperation with the Aleksanteri Institute and the University of Glasgow, organized by Eszter Gantner and Olga Dovbysh. https://www.herder-institut.de/veranstaltungen/workshops/2019/politics-of-e-heritage-production-and-regulation-of-digital-memory-in-eastern-europe-and-russia.html (accessed 30 June 2021). See also Gantner and Dovbysh (2019). For Eastern Europe and Russia more generally see, for instance, Vaypan and Nuzov (2021).
5. See https://www.slavevoyages.org/about/about#history/1/en/.
6. See https://www.nationaalarchief.nl/en/slavery (accessed 3 May 2021).
7. See https://syrian-heritage.org/ (accessed 29 July 2022).
8. See, for instance, Evans (1970), an internal staff information paper of the US National Archives and Records Service. https://www.archives.gov/preservation/formats/microfilm-prep.html.
9. See https://web.archive.org/web/20060612205929/ http://www.unesco.org/webworld/digicol/.
10. Following a hiatus after surveys in 2012, 2014 and 2015, and 2107, the project has recently been restarted, see https://pro.europeana.eu/page/enumerate. The old eNumerate project website (enumerate.eu), however, now redirects to European Group on Museum Statistics (EGMUS) which focuses only on museum statistics.
11. Sternfeld focuses on selection, search, and metadata, respectively (Sternfeld, 2011); Jensen (2021) discusses '(1) predefined subject categories and tags, (2) algorithms/search fields and sorting/display of results and (3) metadata' (p. 257).
12. Project Gutenberg officially started in 1971 but only began larger-scale digitization in the 1990s. Ancestry.com was launched in 1997 but can only be characterized as a mass digitization project since the 2000s.
13. Europeana strategy 2020–2025. Empowering digital change: https://op.europa.eu/en/publication-detail/-/publication/2174be9e-ac58-11ea-bb7a-01aa75ed71a1/ (accessed 3 May 2021). The strategy is aligned with the 2019 EU Declaration of Cooperation on Cultural Heritage: https://digital-strategy.ec.europa.eu/en/news/eu-member-states-sign-cooperate-digitising-cultural-heritage (accessed 3 May 2021).
14. See https://eap.bl.uk/ (accessed 3 May 2021). See also Kominko (2015).
15. https://blackculturalarchives.org/.
16. See http://www.parallelarchive.org/ (accessed 7 July 2021).

Downloaded from https://academic.oup.com/dsh/advance-article/doi/10.1093/llc/fqac050/6702047 by Consortium Luxembourg, gerben.zaagsma on 16 September 2022

17. Ibidem. The Europeana Strategy 2020–2025 does not discuss content strategy but focuses on the overall strategy until 2025.
18. http://www.kultura.hr/eng/About-us/The-Croatian-Cultural-Heritage-project (accessed 30 December 2020).
19. https://www.metamorfoze.nl/english/approach (accessed 30 December 2020).
20. See https://polona.pl/page/about-polona/(accessed 30 December 2020).
21. See, for instance, the UK–German workshop 'Digital Archive and Canon', during which the 'canon question 2.0' was discussed in March 2021. https://www.digitales-archiv-und-kanon.de/.
22. For an exploration of the tension between the hierarchical principle of original order and non-hierarchical keyword search in digital archives, see Zhang (2012).
23. https://makingafricanconnections.org/s/archive/page/index. For the final report, see Reem *et al.* (2021).
24. https://archiefstats.beeldengeluid.nl/flight-over-the-archive.
25. https://impresso-project.ch/. For a general exploration of the challenges and risks of interfaces in relation to historical newspapers, see Pfanzelter *et al.* (2020).
26. Cataloguing is a fundamental precondition for enabling access to heritage, and without the ability to even find out about important archival holdings online, the question of whether they are digitized or not becomes moot. This is especially true in historical research, where knowledge about the existence of materials that might form part of one's evidentiary base is central to framing research designs, and where justifying the choice of materials that are to be used in a given research project, whether these can be found online or offline, is essential before actual research even begins. Not knowing what sources exist introduces fundamental epistemological biases in historical research.

Acknowledgements

I would like to thank Sinem Adar and the two anonymous peer reviewers for their thoughtful, incisive, and constructive comments, as well as Simon Watmough for English proofreading. Earlier versions of the article were presented in various conferences and workshops and benefitted greatly from discussions, feedback, and suggestions from many colleagues.

References

Anderson, S. and Hart, G. (2016). Challenges in digitising liberation archives: a case study. *Innovation: Journal of Appropriate Librarianship and Information Work in Southern Africa* (53): 21–38.

Anziska, S. (2019). Special document file. The Erasure of the Nakba in Israel's archives. *Journal of Palestine Studies*, **49**(1): 64–76.

Baker, N. (2002). *Double Fold: Libraries and the Assault on Paper*. London: Vintage.

Baker, S. (2018). *Community Custodians of Popular Music's Past: A DIY Approach to Heritage*. London: Routledge.

Baets, A. de (2009). *Responsible History*. Oxford: Berghahn Books.

Baets, A. de (2019). *Crimes against History*. London: Routledge.

Baker, J. (2019). *Decolonial Futures for Colonial Metadata, 1838–Present*. Institute of Historical Research Digital History Seminar, London, *21 May 2019*.

Ballouz, I. (2018). Die digitale Rettung des syrischen Kulturerbes. *Flächennutzungsmonitoring IX. Nachhaltigkeit der Siedlungs- und Verkehrsentwicklung*? **73**: 303–15.

Balogun, T. (2018). The Nexus between digitization, preservation and access in the context of selection of materials for archives. *Library Philosophy and Practice*: 1–15.

Barringer, T., Wallace, M., and Damen, J. (eds) (2014). *African Studies in the Digital Age: DisConnects?* Leiden: Brill.

Bastian, J. A. and Flinn, A. (eds) (2020). *Community Archives, Community Spaces: Heritage, Memory and Identity*. London: Facet Publishing.

Bérard, R., Girault, B., Rideau-Kikuchi, C., and Boucheron, P. (2020). *Initiation aux Études Historiques*. Paris: Nouveau Monde Éditions.

Berger, S. and Lorenz, C. (eds) (2010). *Nationalizing the Past: Historians as Nation Builders in Modern Europe*. New York: Palgrave Macmillan.

Berger, S. (2013). The role of national archives in constructing national master narratives in Europe. *Archival Science*, **13**(1): 1–22.

Berry, D. M. and Fagerjord, A. (2017). *Digital Humanities: Knowledge and Critique in a Digital Age*. Cambridge: Polity.

Blaney, J., Winters, J., Milligan, S., and Steer, M. (2021). *Doing Digital History. A Beginner's Guide to Working with Text as Data*. Manchester: Manchester University Press.

Blouin, F. X. and Rosenberg, W. G. (eds) (2010). *Archives, Documentation, and Institutions of Social Memory: Essays from the Sawyer Seminar*. Ann Arbor: University of Michigan Press.

Blouin, F. X. and Rosenberg, W. G. (eds) (2011). *Processing the Past: Contesting Authority in History and the Archives*. Oxford: Oxford University Press.

Boylan, R. (1986). The cooperative Africana microform project. *Microform & Digitization Review*, **15**(3): 167–71.

Brown, R. H. and Davis-Brown, B. (1998). The making of memory: The politics of archives, libraries and museums in the construction of national consciousness. *History of the Human Sciences*, **11**(4): 17–32.

Burkhardt, M. (2017). Vorüberlegungen zu einer Kritik der Algorithmen an der Grenze von Wissen und Nichtwissen. *Jahrbuch Technikphilosophie*, **3**: 55–67.

Cady, S. A. (1990). The electronic revolution in libraries: Microfilm Déja vu? *College and Research Libraries*, **51**(4): 374–86.

Carter, R. G. S. (2006). Of things said and unsaid: Power, archival silences, and power in silence. *Archivaria*, **61**: 215–33.

Chamelot, F., Hiribarren, V., and Rodet, M. (2020). Archives, the digital turn, and governance in Africa. *History in Africa*, **47**(1): 101–18.

Christen, K. (2011). Opening archives: Respectful repatriation. *The American Archivist*, **74**(1): 185–210.

Christen, K., Davis, L., Griffith, Z., and Neely, J. (2018). Traditional knowledge and digital archives: An interview

Downloaded from https://academic.oup.com/dsh/advance-article/doi/10.1093/llc/fqac050/6702047 by Consortium Luxembourg, gerben.zaagsma on 16 September 2022

with Kim Christen. *DisClosure: A Journal of Social Theory*, **27**(1): 8–14.
Coburn, J. (2020). Defending the digital: Awareness of digital selectivity in historical research practice. *Journal of Librarianship and Information Science*. https://doi.org/10.1177/0961000620918647.
Colavizza, G., Blanke, T., Jeurgens, C., et al. (2021). Archives and AI: An overview of current debates and future perspectives. *Journal on Computing and Cultural Heritage*, **15**(1): 1–15.
Cook, T. (2010). Remembering the future: Appraisal of records and the role of archives in constructing social memory. In Blouin, F. X. and Rosenberg, W. G. (eds), *Archives, Documentation, and Institutions of Social Memory: Essays from the Sawyer Seminar*. Ann Arbor: University of Michigan Press, pp. 169–182.
Cook, T. (2011). 'We are what we keep; we keep what we are': Archival appraisal past, present and future. *Journal of the Society of Archivists*, **32**(2): 173–89.
Corfield, P. L. and Hitchcock, T. (2022). *Becoming a Historian: An Informal Guide*. London: University of London Press.
Crymble, A. (2021). *Technology and the Historian. Transformations in the Digital Age*. Urbana: University of Illinois Press.
Curtin, P. D. (1960). The archives of tropical Africa: A reconnaissance. *The Journal of African History*, **1**(1): 129–47.
De Kosnik, A. (2016). *Rogue Archives: Digital Cultural Memory and Media Users*. Cambridge, MA: MIT Press.
Derrida, J. (1995). Archive fever: A Freudian impression. *Diacritics*, **25**(2): 9–63.
Dommann, M. (2008). Papierstau und informationsfluss: Die normierung der bibliothekskopie. *Historische Anthropologie*, **16**(1): 31–54.
Drucker, J. (2016). At the intersection of computational methods and the traditional humanities. In Simanowski, R. (ed.), *Digital Humanities and Digital Media: Conversations on Politics, Culture, Aesthetics and Literacy*. London: Open Humanities Press, pp. 43–68.
Duranti, L. and Shaffer, E. (eds) (2012). *The Memory of the World in the Digital Age: Digitization and Preservation*. Vancouver: UNESCO.
Eckert, G. and Beigel, T. (2019). *Historisch Arbeiten: Handreichung zum Geschichtsstudium*. Göttingen: Vandenhoeck & Ruprecht.
Edmond, J., Horsley, N., Lehmann, J., and Priddy, M. (2022). *The Trouble with Big Data. How Datafication Displaces Cultural Practices*. London: Bloomsbury Academic.
Edmondson, R., Lothar, J., and Prodan, A. C. (eds) (2020). *The UNESCO Memory of the World Programme: Key Aspects and Recent Developments*. Cham: Springer.
Ehrmann, M., Bunout, E., and Düring, M. (2019). *Historical Newspaper User Interfaces: A Review*. Athens: IFLA World Library and Information Congress.
Es, K. van, Schäfer, M. K., and Wieringa, M. (2021). Tool Criticism and the Computational Turn. A "Methodological Moment" in Media and Communication Studies. *Medien & Kommunikationswissenschaft*, **69** (1): 46–64.
Espinosa de los Monteros, P. (2019). Decolonial information practices: Repatriating and stewarding the Popol Vuh online. *Preservation, Digital Technology & Culture*, **48**(3–4): 107–19.
Evans, F. B. (1970). *The Selection and Preparation of Records for Publication on Microfilm*. Washington: National Archives and Records Service, General Services Administration.
Fernandez, M. V. (2018). *The Coloniality of Metadata: A Critical Data Analysis of the Archive of Early American Images at the John Carter Brown Library*. PhD Thesis. University of Texas, Austin.
Fickers, A. (2020). Update für die Hermeneutik. Geschichtswissenschaft auf dem Weg zur digitalen Forensik. *Zeithistorische Forschungen/Studies in Contemporary History*, **17**(1): 157–68.
Fickers, A. and van der Heijden, T. (2020). Inside the trading zone: Thinkering in a digital history lab. *Digital Humanities Quarterly*, **14**(3).
Fiormonte, D. (2012). Towards a cultural critique of the digital humanities. *Historical Social Research*, **37**(3): 59–76.
Fiormonte, D. (2017). Digital humanities and the geopolitics of knowledge. *Digital Studies/Le Champ Numérique*, **7**(1): 5, 1–18.
Foley, T. (2018). 'Slavery, Memory, and Reconciliation through Digital Archives', in *European Social Science History Conference*, Queen's University Belfast, 4 April 2018.
Foster, C. (2010). Microfilming activities of the historical records survey, 1935–42. *The American Archivist*, **48**(1): 45–55.
Föhr, P. (2017). *Historische Quellenkritik im Digitalen Zeitalter*. Ph.D. thesis, Universität Basel.
Foucault, M. (1972). *The Archaeology of Knowledge*. London: Tavistock Publications.
Francomano, E. C. and Bamford, H. (2022). Whose digital middle ages? Accessibility in digital medieval manuscript culture. *Journal of Medieval Iberian Studies*, **14**(1): 15–27.
Fretwell, J. (1908). Photographic copying in libraries. *Library Journal*, **33**: 223–4.
Friedrich, M. (2014). Vom Exzerpt zum Photoauftrag zur Datenbank. Technische Rahmenbedingungen historiographischer Forschung in Archiven und Bibliotheken und ihr Wandel seit dem 19. *Jahrhundert. Historische Anthropologie*, **22**(4): 278–98.
Galen, C. W. V. (2019). Creating an audience: Experiences from the Surinamese slave registers crowdsourcing project. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, **52**(3): 178–94.
Gantner, E. and Dovbysh, O. (2019). Digitization of memory and politics in Eastern Europe. *EuropeNow*. https://www.europenowjournal.org/2019/09/09/on-digitalization-of-memory-and-politics-of-e-heritage-in-eastern-europe/ (accessed 14 June 2021).
Garaba, F. (2015). 'The Timbuktu Manuscripts: A Model for Preservation in Africa', in *IFLA WLIC 2015*, Cape Town, South Africa.
Garaba, F. (2021). Migrated archives. *Journal of the South African Society of Archivists*, **54**: 1–11.
Gentry, K. and Smith, L. (2019). Critical heritage studies and the legacies of the late-twentieth century heritage canon. *International Journal of Heritage Studies*, **25**(11): 1148–68.
Gertz, J. (1999). Selection for preservation in the digital age. *Library Resources & Technical Services*, **44**(2): 97–104.
Gilliland, A. J., McKemmish, S., and Lau, A. J. (eds) (2016). *Research in the Archival Multiverse*. Clayton: Monash University Publishing.

Downloaded from https://academic.oup.com/dsh/advance-article/doi/10.1093/llc/fqac050/6702047 by Consortium Luxembourg, gerben.zaagsma on 16 September 2022

Gjuričová, A. (2019). European history reloaded: The case of Czechoslovak communism. *EuropeNow*. https://www.europenowjournal.org/2019/09/09/european-history-reloaded/ (accessed 14 June 2021).

Golubev, A. (2021). Digitizing archives in Russia: Epistemic sovereignty and its challenges in the digital age. In Gritsenko, D., Wijermars, M. and Kopotev, M. (eds), *The Palgrave Handbook of Digital Russia Studies*. Cham: Springer International Publishing, pp. 353–69.

Gomez, A. (2019). The making of the digital working class: Social history, digital humanities, and its sources. In Gold, M. K. and Klein, L. F. (eds), *Debates in the Digital Humanities 2019*. Minneapolis: University of Minnesota Press, pp. 399–405.

Gordon, A. A. (2014). *Eurocentric Archival Knowledge Production and Decolonizing Archival Theory*. Ph.D. thesis, York University, Toronto. https://yorkspace.library.yorku.ca/xmlui/handle/10315/29935 (accessed 10 April 2022).

Gould, S. and Ebdon, R. (1999). *IFLA/UNESCO Survey on Digitisation and Preservation*. Boston Spa: IFLA/UNESCO.

Grallert, T. (2021). Catch me if you can! Approaching the Arabic Press of the Late Ottoman Eastern Mediterranean through digital history. *Geschichte Und Gesellschaft*, **46**(1): 58–89.

de Groot, J. (2020). Ancestry.com and the evolving nature of historical information companies. *The Public Historian*, **42**(1): 8–28.

Guiliano, J. and Heitman, C. (2019). Difficult heritage and the complexities of indigenous data. *Journal of Cultural Analytics*, **4**(1): 1–25.

Hall, S. (1999). Whose heritage? Un-settling 'The Heritage'. Re-imagining the post-nation. *Third Text*, **13**(49): 3–13.

Hauswedell, T., Nyhan, J., Beals, M. H., Terras, M., and Bell, E. (2020). Of global reach yet of situated contexts: An examination of the implicit and explicit selection criteria that shape digital archives of historical newspapers. *Archival Science*, **20**(2): 139–65.

Harrison, R. (2013). *Heritage: Critical Approaches*. London: Routledge.

Hedstrom, M. (2002). Archives, memory, and interfaces with the past. *Archival Science*, **2**(1–2): 21–43.

Hitchcock, T. (2012). A five minute rant for the consortium of European research libraries. *Historyonics*, 29 October 2012. http://historyonics.blogspot.com/2012/10/a-five-minute-rant-for-consortium-of.html (accessed 7 October 2019).

Hobbs, A. (2013). The deleterious dominance of the times in nineteenth-century scholarship. *Journal of Victorian Culture*, **18**(4): 472–97.

Huistra, H. and Mellink, B. (2016). Phrasing history: Selecting sources in digital repositories. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, **49**(4): 220–9.

Howell, J. B. and Scheven, Y. (1997). Guides, collections and ancillary materials to African archival resources in the United States. *Electronic Journal of Africana Bibliography*, **1**: 1–82.

Jacobsen, T., Punzalan, R. L., and Hedstrom, M. L. (2013). Invoking "collective memory": Mapping the emergence of a concept in archival science. *Archival Science*, **13**(2–3): 217–51.

Jakubowski, A., Hausler, K., and Fiorentini F. (eds) (2019). *Cultural Heritage in the European Union: A Critical Inquiry Into Law and Policy*. Leiden: Brill.

Jensen, H. S. (2021). Digital archival literacy for (all) historians. *Media History*, **27**(2): 251–65.

Jeurgens, C. (2019). *Geschiedenisschaamte? Het Witte Archief Onder Vuur*. Ketelaarlezing 2019. Den Haag: Nationaal Archief.

Jeurgens, C. and Karabinos, M. (2020). Paradoxes of curating colonial memory. *Archival Science*, **20**(3): 199–220.

Kamp, J., Legêne, S., Rossum, M. V., and Rümke, S. (2016). *Geschiedenis Schrijven! Wegwijzer voor Historici*. Amsterdam: Amsterdam University Press.

Kamp, J., Legêne, S., Rossum, M. V., and Rümke, S. (2018). *Writing History! A Companion for Historians*. Amsterdam: Amsterdam University Press.

Katuu, S. (2015). The development of archives and records management education and training in Africa—challenges and opportunities. *Archives and Manuscripts*, **43**(2): 96–119.

Kassow, S. D. (2007). *Who Will Write Our History? Emanuel Ringelblum, the Warsaw Ghetto, and the Oyneg Shabes Archive*. Bloomington: Indiana University Press.

Kerssens, N. (2017). When search engines stopped being human: Menu interfaces and the rise of the ideological nature of algorithmic search. *Internet Histories*, **1**(3): 219–37.

Kessler, M., Merz, U., and Schibler, B. (2021). *DIGIARCH 2021: Kulturerbe im digitalen Zeitalter. Patrimoine culturel à l'ère numérique*. Basel: Schwabe Verlag.

Ketelaar, E. (2008). Archives as spaces of memory. *Journal of the Society of Archivists*, **29**(1): 9–27.

Ketelaar, E. (2017). Archival turns and returns. Studies of the archive. In Gilliland, A. J., McKemmish, S., and Lau, A. J. (eds), *Research in the Archival Multiverse*. Clayton: Monash University Publishing, pp. 228–68.

Kizhner, I., Terras, M., Rumyantsev, M., et al. (2021). Digital cultural colonialism: Measuring bias in aggregated digitized content held in Google Arts and Culture. *Digital Scholarship in the Humanities*, **36**(3): 607–40.

Kominko, M. (ed) (2015). *From Dust to Digital: Ten Years of the Endangered Archives Programme*. Cambridge: Open Book Publishers.

Krumbacher, K. (1906). *Die Photographie im Dienste der Geisteswissenschaften*. Leipzig: Teubner B.G..

Lalu, P. (2007). The virtual stampede for Africa: Digitisation, postcoloniality and archives of the liberation struggles in Southern Africa. *Innovation: Journal of Appropriate Librarianship and Information Work in Southern Africa*, **34**: 28–44.

Lee, B. (2021). Compounded Mediation: A Data Archaeology of the Newspaper Navigator Dataset. *Digital Humanities Quarterly*, **15**(4).

Levi, A. S. (2021). *Keynote Lecture—The Things We Think We Link: The Cascading Violence of Colonial Archives, Linked Pasts VII Symposium*, 13–21 December 2021, Ghent University.

Lewi, H., Smith, W., vom Lehn, D. and Cooke, S. (eds) (2019). *The Routledge International Handbook of New Digital Practices in Galleries, Libraries, Archives, Museums and Heritage Sites*. London: Routledge.

Liu, A. (2012). Where is cultural criticism in the digital humanities? In Gold, M. K. (ed.), *Debates in the Digital Humanities*. Minneapolis: University of Minnesota Press, pp. 490–509.

Lowry, J. (ed.) (2017). *Displaced Archives*. London: Routledge.

Ma, R. (2020). Translational challenges in cross-cultural digitization ethics: The case of Chinese Marriage Documents, 1909–1997. *Libri* **70**(4): 269–77.

Downloaded from https://academic.oup.com/dsh/advance-article/doi/10.1093/llc/fqac050/6702047 by Consortium Luxembourg, gerben.zaagsma on 16 September 2022

Manoff, M. (2004). Theories of the archive from across the disciplines. *Portal: Libraries and the Academy*, **4**(1): 9–25.
Manoff, M. (2016). Mapping archival silence: Technology and the historical record. In Foscarini, F., MacNeil, H., Mak, B., and Oliver, G. (eds). *Engaging With Records and Archives: Histories and Theories*. London: Facet Publishing, pp. 63–82.
Manžuch, Z. (2017). Ethical issues in digitization of cultural heritage. *Journal of Contemporary Archival Studies*, **4**: 1–18.
Marciano, R., et al. (2019). 'Reframing Digital Curation Practices through a Computational Thinking Framework', in *2019 IEEE International Conference on Big Data*, pp. 3126–3135.
Milligan, I. (2013). Illusionary order: Online databases, optical character recognition, and Canadian history, 1997–2010. *Canadian Historical Review*, **94**(4): 540–69.
Mizruchi, L. (ed.) (2020). *Libraries and Archives in the Digital Age*. Cham: Palgrave Macmillan.
Münster, S., et al. (2021). Teaching digital cultural heritage and digital humanities the current state and prospects. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, **XLVI-M-1-2021**: 471–8.
Hedges, M., Marciano, R., and Goudarouli, E. (2022). Introduction to the special issue on computational archival science. *Journal on Computing and Cultural Heritage*, **15**(1): 1–2.
Masenya, T. M. (2021). The use of metadata systems for the preservation of digital records in cultural heritage institutions. *Journal of the South African Society of Archivists*, **54**: 55–70.
McCauley, D. (ed.) (2016). A new age of culture. The digitisation of arts and heritage. *Economist Intelligence Unit*.
Mihai, A. (2015). The politics of online reproduction of Shakespeare. *Revista Transilvania*, **8**: 15–21.
Milligan, I. (2019). *History in the Age of Abundance? How the Web Is Transforming Historical Research*. Montreal & Kingston: McGill-Queen's University Press.
Mills, A. (2015). User impact on selection, digitization, and the development of digital special collections. *New Review of Academic Librarianship*, **21**(2): 160–9.
Moore, N. and Whelan, Y. (eds) (2007). *Heritage, Memory and the Politics of Identity: New Perspectives on the Cultural Landscape*. Aldershot: Ashgate.
Morris, R. (1941). The challenge of historical materials. *The American Archivist*, **4**(2): 91–116.
Moss, M., Thomas, D., and Gollins, T. (2018). Artificial fibers—the implications of the digital for archival access. *Frontiers in Digital Humanities*, **5**.
Muehlberger, G., et al. (2019). Transforming scholarship in the archives through handwritten text recognition. *Journal of Documentation*, **75**(5): 954–76.
Nowatzki, R. (2021). From datum to databases: Digital humanities, slavery, and archival reparations. *The American Archivist*, **83**(2): 429–48.
Niggemann, E., De Decker, J., and Lévy, M. (2011). *The New Renaissance. Report of the Comité des Sages—Reflection Group on Bringing Europe's Cultural Heritage Online*, Brussels.
Odumosu, T. (2020). The crying child: On colonial archives, digitization, and ethics of care in the cultural commons. *Current Anthropology*, **61**(22): 289–302.
Öhman, C. (2022). The case for a digital world heritage label. *Information & Culture*, **57**(1): 82–95.
Ortolja-Baird, A. and Nyhan, J. (2021). Encoding the haunting of an object catalogue: On the potential of digital technologies to perpetuate or subvert the silence and bias of the Early-Modern Archive. *Digital Scholarship in the Humanities*.
Ozban, E. (2014). *The Politics of Archive and Counter-archival Practices: The Case of Geziparkarsiv*. MA thesis, Goldsmiths College, London.
Padilla, T., Allen, L., Frost, H., et al. (2019). Always Already Computational: Collections as Data: Final Report, Libraries at University of Nebraska-Lincoln.
Pandey, R. and Kumar, V. (2020). Exploring the impediments to digitization and digital preservation of cultural heritage resources: A selective review. *Preservation, Digital Technology & Culture*, **49**(1): 26–37.
Pfanzelter, E., et al. (2020). Digital interfaces of historical newspapers: Opportunities, restrictions and recommendations. *Journal of Data Mining and Digital Humanities*, **27**: 1–27.
Pickover, M. (2014). Patrimony, Power and Politics: Selecting, Constructing and Preserving Digital Heritage Content in South Africa and Africa. *IFLA WLIC*.
Prescott, A. (2012). Making the digital human: Anxieties, possibilities, challenges. Keynote Digital. Humanities@Oxford Summer School. *Digital Riffs*, 6 July 2012. http://digitalriffs.blogspot.nl/2012/07/making-digital-human-anxieties.html.
Presner, T. (2016). The ethics of the algorithm. Close and distant listening to the Shoah Foundation visual history archive. In Fogu, C., Kansteiner, W., and Presner, T. S. (eds), *Probing the Ethics of Holocaust Culture*. Cambridge: Harvard University Press, pp. 175–202.
Putnam, L. (2016). The transnational and the text-searchable: Digitized sources and the shadows they cast. The transnational and the text-searchable. *The American Historical Review*, **121**(2): 377–402.
Rafferty, P. (2001). The representation of knowledge in library classification schemes. *Knowledge Organization*, **28**(4): 180–91.
Reem, A. H., et al. (2021). *Making African Connections: Decolonial Futures for Colonial Collections. Initial Findings and Recommendations*. Reports and working papers, University of Sussex.
Rehbein, M. (2016). It's our department: On ethical issues of digital humanities. In Richts K. and Stadler P. (eds), *Ei, dem alten Herrn Zoll' ich Achtung gern: Festschrift für Joachim Veit zum 60. Geburtstag*. München: Allitera Verlag, pp. 631–54.
Ringel, S. (2020). Interfacing with the past: Archival digitization and the construction of digital depository. *Convergence*, **27**(4).
Risam, R. (2019). *New Digital Worlds: Postcolonial Digital Humanities in Theory, Praxis, and Pedagogy*. Evanston: Northwestern University Press.
Rolan, G., Humphries, G., Jeffrey, L., Samaras, E., Antsoupova, T., and Stuart, K. (2019). More human than human? Artificial intelligence in the archive. *Archives and Manuscripts*, **47**(2): 179–203.
Romein, C. A., et al. (2020). State of the field: Digital history. *History*, **105**(365): 291–312.

Downloaded from https://academic.oup.com/dsh/advance-article/doi/10.1093/llc/fqac050/6702047 by Consortium Luxembourg, gerben.zaagsma on 16 September 2022

Rosenzweig, R. (2003). Scarcity or abundance? Preserving the past in a digital era. *The American Historical Review*, **108**(3): 735–62, 39.

Rusert, B. (2017). New world: The impact of digitization on the study of slavery. *American Literary History*, **29**(2): 267–86.

Saber, D. and Long, P. (2017). 'I will not leave, my freedom is more precious than my blood'. From affect to precarity: Crowd-sourced citizen archives as memories of the Syrian war. *Archives and Records*, **38**(1): 80–99.

Scholz, H., et al. (2017). *Europeana Content Strategy. Getting the Right Content to the Right User at the Right Time*. Den Haag: Europeana Foundation. https://pro.europeana.eu/files/Europeana_Professional/Publications/Europeana%20Content%20Strategy.pdf (accessed 30 December 2020).

Sela, R. (2018). The genealogy of colonial plunder and erasure—Israel's control over palestinian archives. *Social Semiotics*, **28**(2): 201–29.

Shammout, B. (2018). *Digitale Erhaltung des Auditiven und Visuellen Kulturerbes Palästinas. Grundlagen und Perspektiven*. Baden-Baden: Tectum Verlag.

Sinn, D. (2013). The use context of digital archival collections: Mapping with historical research topics and the content of digital archival collections. *Preservation, Digital Technology & Culture (PDT&C)*, **42**(2): 73–86.

Sinn, D. and Soares, N. (2014). Historians' use of digital archival collections: The web, historical scholarship, and archival research. *Journal of the Association for Information Science and Technology*, **65**(9): 1794–809.

Smith, S. D. (2016). On the ethics of technology and testimony, In Fogu, C., Kansteiner, W., and Presner, T. S. (eds), *Probing the Ethics of Holocaust Culture*. Cambridge: Harvard University Press, pp. 203–17.

Smisor, G. T. (1942). Collecting a library of rarities on microfilm. *Journal of Documentary Reproduction*, **5**: 98–102.

Smith, L. (2006). *Uses of Heritage*. London: Routledge.

Smith, L. (2021). Critical heritage studies. In *Bloomsbury History: Theory and Method Articles*. London: Bloomsbury Publishing. http://dx.doi.org/10.5040/9781350970885.071.

Stančić, H. (2018). Computational archival science. *Moderna Arhivistika*, **1**(2): 323–330.

Sternfeld, J. (2011). Archival theory and digital historiography: Selection, search, and metadata as archival processes for assessing historical contextualization. *The American Archivist*, **74**(2): 544–75.

Stoler, A. L. (2010). *Along the Archival Grain: Epistemic Anxieties and Colonial Common Sense*. Princeton: Princeton University Press.

Stork, L., Weber, A., Gassó Miracle, E., et al. (2019). Semantic Annotation of Natural History Collections. *Journal of Web Semantics*, **59**.

Story, D. J., Guldi, J., Hitchcock, T., et al. (2020). History's Future in the Age of the Internet. *The American Historical Review* **125** (4): 1337–46.

Tanner, S. and Kahn, R. J. (2014). Building futures: The role of digital collections in shaping national identity in Africa. In Barringer, T., Wallace, M., and Damen, J. (eds), *African Studies in the Digital Age: DisConnects?* Leiden: Brill, pp. 111–27.

Thomas, D., Fowler, S., and Johnson, V. (2017). *The Silence of the Archive*. London: Facet Publishing.

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

Trace, C. B. (2020a). Maintaining records in context: A historical exploration of the theory and practice of archival classification and arrangement. *The American Archivist*, **83**(1): 91–127.

Trace, C. B. (2020b). Maintaining records in context? Disrupting the theory and practice of archival classification and arrangement. *The American Archivist*, **83**(2): 322–72.

Traub, M. C., Van Ossenbruggen, J., and Hardman, L. (2015). Impact analysis of OCR quality on research tasks in digital archives. In Kapidakis, S., Mazurek, C., and Werla, M. (eds), *Lecture Notes in Computer Science*. Cham: Springer International Publishing, pp. 252–63.

Turner, H. (2017). Organizing knowledge in museums: A review of concepts and concerns. *Knowledge Organization*, **44**(7): 472–84.

Valtysson, B. (2012). EUROPEANA: The digital construction of Europe's collective memory. *Information, Communication & Society*, **15**(2): 151–70.

Vaypan, G., and Nuzov, I. (2021). *Russia: 'Crimes Against History'*. Paris: FIDH.

Venkata, S. K., et al. (2021). Alexa, is this a historical record? *Journal on Computing and Cultural Heritage*, **15**(1): 1–20.

Verbruggen, C., Delafontaine, R., Danniau, F., and Chambers, S. (2017). Bedenkingen bij the history manifesto: van Coöperatie naar collaboratie. *Belgisch Tijdschrift Voor Nieuwste Geschiedenis*, **47**(2–3): 202–12.

Verhaar, P. A. F. (2016). *Affordances and Limitations of Algorithmic Criticism*. Ph.D. thesis, Leiden University.

Vukli, V. and Gilliland, A. J. (2021). Archival activism: Emerging forms, local applications. In Filej, B. (ed.), *Archives in the Service of People—People in the Service of Archives*. Maribor: Alma Mater Europea, pp. 14–25.

Wevers, M. and Smits, T. (2020). The visual digital turn: Using neural networks to study historical images. *Digital Scholarship in the Humanities*, **35**(1): 194–207.

van der Werf, T. and van der Werf, B. (2022). Will archivists use AI to enhance or to dumb down our societal memory? *AI & Society*, **37**: 985–8.

Whaanga, H. and Mato, P. (2021). The Indigenous digital footprint. In Hokowhitu, B., Moreton-Robinson, A., Tuhiwai-Smith, L., Andersen, C., and Larkin, S. (eds), *Routledge Handbook of Critical Indigenous Studies*. London: Routledge.

Whaley, J. H. (1994). Digitizing history. *The American Archivist*, **57**(4): 660–72.

Whitehead, C., Eckersley, S., Daugbjerg, M., and Bozoglu, G. (eds) (2020). *Dimensions of Heritage and Memory: Multiple Europes and the Politics of Crisis*. London: Routledge.

Winters, J. (2019). Digital history. In Tamm, M. and Burke, P. (eds), *Debating New Approaches to History*. London: Bloomsbury, pp. 277–301.

Wodziński, M. (2022). Eastern European Jewish studies: The past thirty years. *Jewish Quarterly Review*, **112**(2): 1–23.

Wurthmann, N. and Schmidt, C. (2020). Digitale Quellenkunde: Zukunftsaufgaben der Historischen Grundwissenschaften. *Studies in Contemporary History*, **17**(1): 169–78.

Yale, E. (2015). The history of archives: The state of the discipline. *Book History*, **18**(1): 332–59.

Youn, E. (2016). Investigating socio-cultural aspects of the implementation of an international archival descriptive

Downloaded from https://academic.oup.com/dsh/advance-article/doi/10.1093/llc/fqac050/6702047 by Consortium Luxembourg, gerben.zaagsma on 16 September 2022

standard in Korea. In Gilliland, A. J., McKemmish, S., and Lau, A. J. (eds), *Research in the Archival Multiverse*. Clayton: Monash University Publishing, pp. 789–810.
Zaagsma, G. (2013a). On digital history. *BMGN—Low Countries Historical Review*, **128**(4): 3–29.
Zaagsma, G. (2013b). Using digital sources in historical research. Jewish History on the Internet. In Clavert, F. and Noiret, S. (eds), *Contemporary History in the Digital Age*. Oxford: Peter Lang Publishing Group, pp. 201–13.
Zaagsma, G. (2017). *East-Central European Jewish Heritage Online. Session Presenting "Difficult" History—Jewish heritage in East-Central Europe. 4th Annual Conference of the International Federation for Public History*. Ravenna, Italy, 5–9 June 2017.
Zhang, J. (2012). Original order in digital archives. *Archivaria*, **74**: 167–93.
Zhang, J. (2018). Digital archival representation: Approaches and challenges. In: Angel, C. M. and Fuchs, C. (eds), *Organization, Representation and Description through the Digital Age Information in Libraries, Archives and Museums*. Berlin: De Gruyter, pp. 89–102.
Zundert, J. J. van, and Haentjens Dekker, R. (2017). Code, scholarship, and criticism: When is code scholarship and when is it not? *Digital Scholarship in the Humanities*, **32(Suppl. 1)**: i121–33.

Miscellaneous

CARE Principles for Indigenous Data Governance (2019). Research Data Alliance International Indigenous Data Sovereignty Interest Group. The Global Indigenous Data Alliance. GIDA-global.org.
European Commission (2011). COMMISSION RECOMMENDATION of 27 October 2011 on the Digitisation and Online Accessibility of Cultural Material and Digital Preservation. *Official Journal of the European Union*, 283.
European Commission (2019). Declaration of Cooperation on Advancing Digitisation of Cultural Heritage.
Europeana Strategy 2020–2025: Empowering Digital Change (2020). Brussels: Publications Office of the European Union.
National Library of Scotland (2018). *The Library's Approach to Selection for Digitisation*. https://www.nls.uk/media-u4/1557983/2018-digitisation-selection-approach.pdf
Proceedings: Conference on Microcopying Research Materials in Foreign Depositories (1940). Washington: Library of Congress and Cosmos Club.

Websites

(All websites have last been accessed in July 2021)

Beeld en Geluid—Visualising the Archive. https://archiefstats.beeldengeluid.nl/.
Black Cultural Archives. https://blackculturalarchives.org/.
Croatian Cultural Heritage-Project. http://www.kultura.hr/eng/About-us/The-Croatian-Cultural-Heritage-project.
Digicol—Directory of Digitised Library Collections. https://web.archive.org/web/20060612205929/http://www.unesco.org/webworld/digicol/.
Endangered Archives Program. https://eap.bl.uk/.
eNumerate. https://en.wikipedia.org/wiki/Enumerate_(project).
Impresso Project. https://impresso-project.ch/.
industrie.lu—D'Industriegeschicht vu Lëtzebuerg. https://industrie.lu/.
Making African Connections. https://makingafricanconnections.org/s/archive/page/index.
Metamorfoze. https://www.metamorfoze.nl/english/approach.
Nationaal Archief—Slavernij. https://www.nationaalarchief.nl/en/slavery.
Parallel Archive. http://www.parallelarchive.org/.
Polona. https://polona.pl/page/about-polona/.
Reciprocal Research Network. https://www.rrncommunity.org/.
SUCHO—Saving Ukrainian Cultural Heritage Online.https://www.sucho.org/.
Syrian Heritage Archive Project. https://syrian-heritage.org/.
Time Machine—A Common History for the Continent. https://www.timemachine.eu/trailer/.
UNESCO—Memory of the World. https://en.unesco.org/programme/mow/.
YIVO Vilna Collections Project. https://www.yivo.org/vilna-collections-project.

Downloaded from https://academic.oup.com/dsh/advance-article/doi/10.1093/llc/fqac050/6702047 by Consortium Luxembourg, gerben.zaagsma on 16 September 2022