

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 archive¹

Alexandra Ortolja-Baird

King's College London, Department of History, London, UK

Julianne Nyhan

UCL, London, UK

Abstract

The subjectivities that shape data collection and management have received extensive criticism, especially with regards to the digitization projects and digital archives of galleries, libraries, archives and museums (GLAM institutions). The role of digital methods for recovering data absences is increasingly receiving attention too. Conceptualizing the absence of non-hegemonic individuals from the catalogues of Sir Hans Sloane as an instance of textual haunting, this article will ask: to what extent do data-driven approaches further entrench archival absences and silences? Can digital approaches be used to highlight or recover absent data? This article will give a decisive overview of relevant literature and projects so as to examine how digital tools are being realigned to recover, or more modestly acknowledge, the vast, undocumented network of individuals who have been omitted from canonical histories. Drawing on the example of Sloane, this article will reiterate the importance of a more rigorous ethics of digital practice, and propose recommendations for the management and representation of historical data, so cultural heritage institutions and digital humanists may better inform users of the absences and subjectivities that shape digital datasets and archives. This article is built on a comprehensive survey of digital humanities' current algorithmic approaches to absence and bias. It also presents reflections on how we, the authors, grappled with unforeseen questions of absence and bias during a Leverhulme-funded collaboration between the British Museum and University College London (UCL), entitled 'Enlightenment Architectures: Sir Hans Sloane's Catalogues of his collections'.

Correspondence:

Julianne Nyhan, UCL,
London, UK.

E-mail:

j.nyhan@ucl.ac.uk

1 Introduction

In this age of data abundance, it is difficult to comprehend the existence and extent of data absence, and

the bias with which it can be implicated. Users have sometimes assumed the neutrality and comprehensiveness of the digital data, collections, and resources at their disposal; and creators have sometimes failed to

ethics of absence at large, while the second, which more frequently engages with the early modern period, as is our frame of study, seeks to recover specific instances of absence without necessarily abstracting these tailored approaches to the general level. To date, the application of these approaches is still nascent in the study of memory institutions. While ‘the museum as a cultural form is the ideal space for drawing attention to gaps in the historical record, to the policies of collection and display, and to the limits of the historical narratives’ (Mason and Sayner, 2019) and has been the site of extensive academic, artistic, and curatorial engagements with these issues over the past two decades (Smith, 2006; *Uncomfortable Truths Exhibition*, 2007; Dissengué and Winter, 2016), there have been fewer interventions into the silences created, or perpetuated, by digital museum practices.² While the democratization of digital heritage and its consumption is at the forefront of many institutions’ initiatives, these are generally assessed in terms of their ‘ability to reach larger user numbers, rather than how the discourse itself is created and mediated’ (Taylor and Gibson, 2017). Paired with the general reluctance in heritage studies to critique the ostensibly well-meaning democratizing turn in the heritage sector, this focus has left questions regarding the underlying meaning/s of digital democratization largely unanswered. In this article, we consequently seek to identify digital approaches to historic, especially early modern absences, and extend them to the field of digital cultural heritage, presenting some tentative approaches to both the epistemological and practical issues of absence.

This article is built on a comprehensive survey of digital humanities’ current algorithmic approaches to absence and bias. It also presents reflections on how we, the authors, grappled with unforeseen questions of absence and bias during a Leverhulme-funded collaboration between the British Museum and University College London (UCL), entitled *Enlightenment Architectures: Sir Hans Sloane’s Catalogues of his collections* (Enlightenment Architectures, 2020). Sir Hans Sloane (1660–1753) was a physician, naturalist, Secretary and later President of the Royal Society, as well as of the Royal College of Physicians, whose personal cabinet became one of the foundational collections of the British Museum. He amassed a vast and varied collection from across the globe during the course of his

long life, which entered into his possession through a variety of means, often via a diverse range of intermediaries. As he collected these objects, Sloane and his amanuenses labelled and described them in more than forty volumes of handwritten catalogues, divided between different parts of his collection, such as botanical materials, antiquities, and books and manuscripts. Although these objects were often recorded in intricate detail, information regarding the routes by which items made their way into Sloane’s possession, through which hands they passed, their exact origins and creators, and the means by which they were acquired is frequently sparse. These lacunae are especially profound considering the forms of colonialism in which Hans Sloane participated and the colonial context of much of his collecting, as we will discuss further below.

Digital humanities has been criticized for its positivism and lack of cultural criticism. Yet, drawing on recent scholarship, and the highly representative challenge posed by Sloane’s catalogues, this article will illustrate how existing and innovative digital tools are being, and might be, realigned, and integrated with a wider symphony of methods and actors, to recover or at least better acknowledge the vast, undocumented network of individuals who have been excluded from canonical histories.

2 Overview of Humanities Literature on Absence

Much recent digital humanities scholarship on absence takes inspiration from longstanding humanities debates regarding the social injustices and hierarchies embedded in the archive. Exemplary of this literature is the work of Rodney G.S. Carter, who argued ‘silences are, in part, the manifestation of the actions of the powerful in denying the marginal access to archives and . . . this has a significant impact on the ability of the marginal groups to form social memory and history’ (Carter, 2006, p. 215). More recently, Ariella Aïsha Azoulay has spoken of the ‘taxonomic violence’ in the imperial archive (Azoulay, 2019, p. 173). These dynamics are no less pertinent to the digital world, as captured by Amiria Salmond, ‘more than merely what may be represented, what “is”, in the ontology of the archive, is that which is *recognized* as worthy of

inclusion. Digital archives are no exception' (Salmond, 2012). Many digital humanists have drawn attention to the absences, both perpetuated and created anew, by the increasing digital presence, representation, and interrogation of data, calling for a re-examination of the social and cultural factors that give rise to absences in the digital archive. This is not to say that such absences are irretrievable; extramural communities and grass-roots organizations have sought to harness the affordances of digital technologies as part of their wider project of 'democratising cultural memory'.³ Yet, as recent studies have comprehensively demonstrated, the project of computerization more widely has frequently transposed rather than replaced deficient cultural scripts about gender, sexuality, and race, through algorithmic bias, datafication, and info-tech labour relations (Hicks, 2018; Noble, 2018; Thylstrup, 2019). The ease and ubiquity of digitization and digital technologies have 'given to the oldest of Western canons a new hyper-availability, and a new authority' (Hitchcock, 2013) while, at the same time, when paired with the political importance of current-day whistleblower culture and freedom of information, also served to undermine the sometimes self-chosen absence of marginalized groups (Robertson, 2018).

Reflection on the nature of digital archives has thus been urged along with the position that as digital archives 'are technocultural artifacts, developments in the field of science studies can provide insight into the interdependence and coevolution of the social, cultural and material factors shaping archival silence' (Manoff, 2016). Similarly, by drawing attention to 'blank spots that exist in spaces that are otherwise data-saturated', Mimi Onuoha has shown how data absences and information that resist quantification are both generated and propagated by power hierarchies and social injustices (Onuoha, 2016). Emphasizing the subjectivity of data collection and data documentation, her research echoes Michel-Rolph Trouillot's earlier conclusions regarding the origins of historical silencing (and therein historical violence) in *Silencing the Past: Power and the Production of History*.⁴ Onuoha adapts this framework for an examination of the digital archive, concluding that:

- (1) datasets are determined by the people collecting data;

- (2) data are collected (and thus flattened) to fit patterns (types, models) of collection;
- (3) data outlive their collecting rationale and consequently play different roles in different cultural contexts;
- (4) when combined, datasets result in inferences that are unexpected and unintended; and
- (5) data collection is 'the result of an invisible relationship' of collectors and collected.

And thus: 'the challenge is for us to keep in mind both aspects of data collection, to see systematic as well as human tensions and biases.' (Onuoha, 2016).

Echoes of Onuoha's rubric can be found in Taylor and Gibson's recent work that identifies five ways in which digitization and digital resources reinforce dominant narratives in the museum:

Research: The direction technology takes has a huge influence on which aspects of digital heritage are researched. The eventual use and implementation of digitised heritage is not always a research priority.

Commissioning: Who funds digitisation and digital interaction activities has a big influence on what is done and whose values are represented.

Resources: The time and equipment required for many projects requires institutional expertise, teamwork and a solid infrastructure.

Training: Expertise required to deliver digital content may vary a great deal, but the advancing of pre-created content out-strips the accessibility to content generation.

Choices: The problems above are not insurmountable. . . but all require some awareness of the political implications. Deferring to the implicit bias of a default situation does not mean that a dominant narrative has been avoided, only that one is not aware of those implications (Taylor and Gibson, 2017).

In identifying the human factors and processes, rather than primarily the technical, which led to absences, the works of Onuoha, Taylor, and Gibson tie into an important body of emerging scholarship that pinpoints the gender, racial, and other social biases of digital resources. On this issue, Tara McPherson has stressed that 'we must remember that computers are

themselves encoders of culture’ (McPherson, 2013), while Nieves and Jaksch have called for a ‘global digital humanities’ (Nieves and Jaksch, 2014). Others have criticized the digital humanities for preserving the social inequalities embedded in historical data (see, e.g. Bianco, 2012). In her examination of slavery’s 18th- and 19th-century Atlantic archive, Jessica Marie Johnson has stressed the importance of black digital practice as a means of reassessing the ‘neutrality of the digital’ (Johnson, 2018). She argues that while the digitization of slavery ‘threatens to replicate the death work of the slave ship register’, black digital practice—the taking up of digital humanities by black subjects—enables individuals to ‘hack their way into the system (modernity, science, the West), take root, and live where they were “never meant to survive”’. Speaking about the Nana Project (Nana Project, 2020), Kirstie A. Kwarteng (2019) has voiced concerns regarding the accessibility of digitized African history. She argues that the primarily academic efforts to digitize the African archive has resulted in the public and the African public, in particular, being unable to profit from digitization efforts. Moreover, she draws attention to culturally significant, non-textual products, such as oral histories, which are absent from the archive. Similarly, Julia Gaffield has drawn attention to the new silences of the digital age, stressing how disparities of access and preservation potentially contribute to the further silencing of the past. She asks:

Which nations or communities get to preserve and make available, physically and virtually, their historical record? Whose sources are beautifully preserved in acid-free boxes and digitized according to international standards? Whose records are organized and made accessible at repositories and online? . . . Whose history is uncatalogued and therefore inaccessible to scholars and to publics because of chronic underfunding? (Gaffield, 2018)

Still, as will be explored in this article, many believe that digital technologies can be used cautiously to re-align the “‘retrospective significance’ of each historical narrative’. The importance of digital approaches as a way of recovering absent voices and rethinking the cultural forms that shape the archive and subsequent histories has been thoughtfully captured by Kelley Kreitz:

As scholars of Latinx Studies – as well as those in related fields such as hemispheric studies, black Atlantic studies, and indigenous studies – work to question the assumptions and omissions of our print-dominated past, digitization projects have become sites for recuperating lost voices, for breaking out of the disciplinary formations that have made sense of cultural history to find new patterns, and for increasing participation in the production of knowledge itself (Kreitz, 2017).

2.1 Research context: current (digital) approaches to historical absence

At the time of writing, the global COVID-19 pandemic—the first ‘data-driven pandemic’—has brought the issue of data absences and their digital presentation and communication crashing down into the public and political arena. Though lauded for raising public awareness of the pandemic, the user-friendly visualization tools tracking the disease have also received criticism for concealing and subsequently perpetuating the disparity in what or whom is being measured. They risk, as Alison Powell has argued, ‘becoming unique forms of disinformation because they focus attention on what is measured (or indeed, measurable) at the expense of nuanced considerations of differential risk and harm from disease as well as its control to women, people of colour and elders’ (Powell, 2020). Yet, as the previous section indicates, this is no new phenomenon and, in spite of ongoing data positivism, a growing corpus of projects is building upon the rich theoretical literature sketched above in order to digitally recover absences in historical collections and realign historical narratives. From them, it is possible to sketch out a typology of current digital approaches to historical data absences: network analysis; visualization tools; the digitization and databasing of historical sources; topic modelling; text-mining; and bibliometrics have all been employed, as have entirely new hybrid methods.

Perhaps, the most popular method is the use of network analysis because of its ability to recover lesser-known actors who are rendered invisible, or treated as ‘lesser figures’, when investigated with traditional humanities lenses. Network analysis has the potential to uncover individuals with strong connections, or

Database, 2020), all provide repositories of materials relating to slavery in the American colonies with the intention of recovering the experiences of the enslaved. The linked open data platform, *Enslaved: The People of the Historic Slave Trade*, which facilitates searches across multiple online databases and provides data visualization tools to create maps, charts, and graphs, likewise intends to recover and preserve data about enslaved persons for future generations (Enslaved, 2020).

Other database projects are more experimental. *Invisible Australians, Living under the White Australia Policy* recovers the biographical records of the ‘invisible’ Australians of the early 20th century—non-Europeans, including Chinese, Japanese, Indians, Afghans, Syrians, and Malays—and enables users to ‘explore the records of the White Australia Policy through the faces of those people’ through portraits extracted from government documents using a face detection script (Invisible Australians, 2020).

More quantitative methods are in use too, often with a pronounced hermeneutical framing. Topic modelling, text-mining, and bibliometrics have been used to recover absences from existing, large-scale digital corpora. This approach has been demonstrated by the *Black Women Big Data* project in their attempt to recover the experiences of Black women from within the digitized record from approximately 800,000 books, newspapers, and articles in the HathiTrust and JSTOR Digital Libraries (Brown *et al.*, 2016). Through reflections that seem to evoke the potential of a humanities data science approach, they argue for the transferability of their approach:

if researchers can train algorithmic models to identify and replicate patterns in data, we can train models to project forward where or when topics might reappear, perhaps shifting in how they are spoken about but (re)emerging nonetheless. . . . For example, it maybe possible to discover an algorithmic thread which connects discussions of lynching during Ida B. Wells’ era, state violence against communities of color during the 1960s, and the current Black Lives Matter movement that may allow us to predict when or how future discussions around state-sanctioned violence against communities of

color might manifest (Brown *et al.*, 2016, p. 122).

In her study of Ottoman Algerian women, Ashley Sanders Garcia uses text-mining to decolonize the archive and reposition women properly within the historical narrative. By using named entity recognition to recover the names of prominent men and women, Sanders Garcia argues that, despite their absence in both scholarship and the public record, women were central to the ‘socio-political fabric of Ottoman Algerian society and government’ (Sanders Garcia, 2019). Similarly, in her examination of the work of Mary Wroth and Mary Sidney Herbert, Amanda Henrichs proposes that text-mining and literary stylistics might enable us to reassess longstanding assumptions regarding these women’s networks (Henrichs, 2020). By reclaiming these intertextual gaps ‘at a site where such a gap—according to our assumptions about this group of authors—should not exist’, Henrichs unravels entrenched narratives and reveals new avenues for future study (Henrichs, 2020).

While the aforementioned projects have adapted existing digital technologies to explore archival absences, others have developed entirely new conceptual and digital approaches to address and explore missing data. Perhaps, the most ground-breaking outlook is provided by Lauren Klein, in her exploration of the intrinsic issues of archival absence in the Antebellum period through the database *The Papers of Thomas Jefferson* (Klein, 2013). Stepping away from ‘the damaging notion that African American voices from before emancipation—not just in the archival record, but the voices themselves—are silent, and irretrievably lost’, Klein ‘demonstrates how a set of techniques that derive from the fields of computational linguistics and data visualization help render visible the archival silences implicit in our understanding of chattel slavery today’ (Klein, 2013). Resituating cultural criticism at the heart of her approach to digital humanities, Klein illustrates how using a named entity recognizer and co-appearance analysis can reveal the ‘complexity of the relations among individuals and across social groups’ and therein ‘deform’ the archive. Similarly, creative approaches include that of Scott Weingart, who treats absence as a ‘creative wellspring’ and simulates the different way data ‘might have looked were

the record more aligned with the lived past' (Weingart, 2011). Exploring the concept of 'historical fidelity' and Avezier Tucker's claims that 'some processes tend to preserve in their end states information from their initial state more than others. Fidelity measures the degree to which a unit of evidence tends to preserve information about its given cause.' (*Our Knowledge of the Past*), Weingart reflects on how the subjectivity of archives and scholars over time dilutes the fidelity of documentary evidence, relinquishing vast amounts of sources to 'historical purgatory'. Additionally, as part of the *Eyes on the Past* project (*Eyes on the Past Project*, 2020), Tim Sherratt has introduced the idea of examining the 'seams and edges of our information landscape'—the various people, data, and systems come together to create the digital, however messy—as 'sites of collaboration, negotiation and repair' (Sherratt, 2015). Drawing attention on how our reliance on the search box leads to the 'smoothing over of sticky data to fit our expectations of seamlessness', Sherratt's experimental interface based on facial detection technology, intends to stimulate new questions surrounding how we explore and interrogate digital collections (Sherratt, 2015).

Before moving on to discuss the approaches to absence with which we have started to experiment in the context of *Enlightenment Architectures*, we will now discuss what is currently known about absence in and from Sloane's catalogues.

2.2 Absences in Sloane's collection catalogues

Even by today's standards, the size and breadth of Sir Hans Sloane's collection are difficult to fathom. His diverse objects, sourced from across the globe and ranging from botanical specimens to Roman artefacts, rare manuscripts and marine fossils, along with the catalogues that recorded them, were left in their entirety to the nation in 1753 to form the basis of what are now three national institutions: the British Museum, the British Library, and the Natural History Museum, London. Their number and preservation have facilitated extensive scholarly attention to the wealth of objects that Sloane collected and his detailed cataloguing types (MacGregor, 1994; Walker *et al.* 2012; Delbourgo 2018), which have

shed light, both directly and indirectly, on profound absences in Sloane's records.

A recurring absence is the paucity of accurate or detailed provenance and transit information in Sloane's catalogues of his collections, primarily regarding the non-western or socially 'inferior' contributors to his collection. Much scholarship has consequently focused on recovering, or more modestly acknowledging, the vast network of individuals who helped Sloane build, organize, and document his collection. The importance of such work was emphasized by Marjorie Caygill in her detailed study of Sloane's catalogues, in which she emphasized the role played by the largely undocumented experts, locals, peers, and dedicated assistants in the formation of his collection (Caygill, 2012). Caygill proposed that these individuals could be retrieved to a small degree through studying the correspondence, notes, and diaries of Sloane and his assistants, as well as the early curators and trustees of the British Museum, and by examining the annotations, notes, and references in Sloane's catalogues themselves. This vast undertaking had already been encouraged by Murray-Jones, who stressed the importance of deeper archival research in recovering the role of external agents in the development of Sloane's collection (Jones, 1988); however, it was not until Amy Blakeway's ambitious work on Sloane's series of library amanuenses that such calls were answered (Blakeway, 2011). Blakeway not only identified several previously anonymous amanuenses and recovered their biographies in relation to Sloane, but also demonstrated that their role in shaping Sloane's library collection and documentation thereof was more significant than previous commentators have appreciated. Victoria Pickering's research into Sloane's vegetable substances built on Dandy's earlier work (Dandy, 1958) on the western contributors to Sloane's botanical collection to likewise uncover a diverse set of contributors and intermediaries in the construction of Sloane's herbarium (Pickering, 2017).

More recently, James Delbourgo has looked holistically across Sloane's 'paper empire', arguing that Sloane acted as a centre point around which his collection was built and organized by a network of lesser and now largely forgotten individuals (Delbourgo, 2012, 2018). Sloane, he claims, collected people as much as he did objects, and people likewise collected Sloane.⁵ Like Blakeway, Delbourgo claims that the

2.3 Digital–epistemological reflections raised by absences in Sloane’s catalogues

As discussed elsewhere (see Ortolja-Baird *et al.*, 2019), the overarching aim of *Enlightenment Architectures: Sir Hans Sloane’s Catalogues of his Collections* was to identify and interrogate the highly complex information architecture of Sloane’s catalogues and their intellectual legacies. The project sought to contribute to ongoing conversations in historical, curatorial and museum studies, and digital humanities, by publishing new research and methodologies that could further the decoding of how Sloane organized his collection (see, e.g. Sloan and Nyhan, 2020; Ortolja-Baird, 2020). Focusing on a subset of five of the approximately forty catalogues that were compiled during Sloane’s lifetime, the project placed particular emphasis on the informational units of which the catalogues are composed. At the level of the informational unit, we sought to disambiguate and encode, *inter alia*, catalogue number and entry; changing hands and curatorial interventions; bibliographical references and author name. Conceptually, we viewed the catalogues’ ‘bifocal data’, which we sought to look both ‘at’ and ‘through’ (Sperberg-McQueen, 2018). This involved our privileging of, as far as possible, a historically accurate representation of the informational entities of Sloane’s catalogues over conformance with the views of information that are implicit in 21st-century encoding specifications like TEI, which in turn led to modifications and customizations of TEI. Picking back up the narrative thread of this article, in the following, we examine our experiences of encoding person and place names, in particular.

As we engaged in the task of encoding the personal names that are given in the catalogues, we began to wonder about those not included in the catalogues. Those individuals’ names may be absent but an echo of their agency, and a trace of their presence is, in some nebulous way, enfolded in the catalogues. After all, the existence of the object (in examples of most artificial and some natural items) in itself indicates that it was made, worked, sold, and transported by human beings. As we worked, we began to conceptualize these nameless individuals as presences who ‘haunt’ the catalogues, in the sense that they participate in a dialectic of trace and absence that is detectable from certain viewpoints only, and rarely

anchorable to a specific location of the catalogue. But how can one encode the ghosts and the ‘haunting’ of an early modern archival document? Encoders can usually tag an individual only if they are actually ‘there’ in some fixed way in a text, for example, if they are textually embodied in a person name or metaphor. Though in some cases, it might be possible to view an object name, or category of knowledge as a proxy for their presence, and encode an entry as, for example, ‘unknown maker/collector/agent or gatherer’, this would require further fundamental long-term research and would not result in clear-cut identifications in all instances.

It was in the process of thinking through how absence, and absent individuals and groups, could be modelled, and encoded in the catalogues, that we were alerted to how positivist encoding schemes like TEI can be.⁷ If a feature of a text is present, and recognized as such by the encoder, then they can tag it (directly or with stand-off mark-up) and proceed to study that textual feature in other ways. But what can be done when an anchor point cannot be found and an absence is textually unmoored? And what can be done when we suspect that a milestone in a catalogue should be associated with an individual but their identity is unknown?

These questions may initially be read as being abstruse, and yet there is much at stake in them. The unnamed and silenced individuals who contributed to Sloane’s catalogues were part of his network due to the growth of global trade and imperial expansion, and the forms of colonialism of which Sloane was a part, working for a colonial governor, owning shares in slave-trading companies, and having married into a plantation-owning family. That these individuals were omitted from Sloane’s catalogues are crucial in understanding the socio-cultural and economic contexts of his collecting practice, the hierarchies of esteem and knowledge that his collecting practice participated in and the ideologies of race that overarched his documentation and practices of attribution. The absences in Sloane’s catalogues are caused by personal and societal ideological biases of data selection, and further informed by imperatives for that data to conform to taxonomies of collection. The absences in Sloane’s catalogues thus speak to the inherent subjectivities of data collection and documentation, be it analogue or digital, reminding us of Drucker’s admonition that

data are ‘capta’ (Drucker, 2011). The absences also raise crucial questions about the extent to which such biases, however removed, continue to shape current scholarship and data-driven approaches to the analysis of historical, archival documents. The positivist orientation of TEI (and other XML-based markup languages) to model Sloane’s catalogues risks, however unintentionally, the further perpetuation of historical absences, and indeed, their activation and amplification in new ways as historical datasets are made machine readable and are combined and recombined in new systems and applications. Two questions thus follow: with regard to the particular context of Sloane, how might we use digital tools to recover, rather than re-encode absences in and from his catalogues? From a broader perspective, what steps might be taken to seek to prevent data-driven approaches to historical documents by further perpetuating the silence of individuals who have already been marginalized in the historical record?

Our noticing of individuals who were absent from Sloane’s catalogues followed various prompts raised by the secondary literature, our personal research projects, and the contemporaneous societal problems that reached new prominence as this research was unfolding. Delbourgo’s book *Collecting the World*, which reveals with a new scope and rigour how Sloane’s collection was financed through the profits of the trans-Atlantic slave trade and the infrastructure and dynamics of colonialism and empire, including the coerced labour of individuals and communities on Jamaican plantations, had been recently published. Next to this, Nyhan was at work on a book on the invisibleized and devalued feminized labour that had been contributed to one of the foundational projects of the Digital Humanities, heightening her attention to framings of the subaltern, and the power dynamics of attribution. All these found wider resonance in ongoing social justice movements like Black Lives Matter, and in broader conversations about the ethics and history of museum collections, so that our scholarly and historical interests conjoined with some of the most pressing problems of our times. It was not, however, possible within the timeline or resources of our project to iteratively adjust our encoding strategy in light of this new awareness. As is discussed in the remainder of this article, we did, however, pursue a wide-ranging literature review and environmental

survey, so as to better understand the contexts and import of the issues at stake and to begin a thought experiment as to how to set about computationally modelling the absences in Sloane’s catalogues in a project that we hope to undertake should the funding bid that we have under review at the time of writing be supported by the Arts and Humanities Research Council, UK.

3 Digital Approaches to Sloane’s Catalogues

Hans Sloane’s catalogues are fertile ground for testing the potential of digital approaches to recovering, or at least flagging, absent actors. However, the question remains as to whether and how this might be achieved? As the projects outlined in Section 2 demonstrate, there is no one-size-fits-all digital solution to the issue of data absences. Bearing in mind Onuoha’s dissection of the subjectivities of data collection and processing outlined in Section 1, how do we tailor existing digital tools to move beyond Sloane’s own selection bias when it comes to recording his objects, so as to reconstruct a more accurate view of the creation and origins of his collection? Or, at the very least, how can we alert those who seek to work with the machine-readable versions of Sloane’s catalogues of the presence of such questions? We understand these questions as some of the ‘Grand Challenges’ that Digital Humanities must prioritize going forward, and below we set out some tentative indications of the response to this challenge that the *Enlightenment Architectures* project is investigating.

3.1 TEI encoding

Notwithstanding the criticism of TEI set out above, it has nevertheless proved useful from two perspectives. First, the positivism of TEI can be counteracted, to some extent, by its flexibility. Sperberg-McQueen (1991, p. 34) has sought to decipher the ‘relationship between the theory of texts and the design of electronic text markup’ and offers a series of axioms that explain why TEI needs to be ‘unbounded and unboundable’. Central to this is the view that any representation of a text communicates an opinion regarding ‘what is important in that text’ (Sperberg-

McQueen, 1991). However, as there are many views of texts, which change over time and across contexts, mark-up schemes are required that ‘allow several discrete views of texts’ (Sperberg-McQueen, 1991). TEI, being extensible, thus encompasses a ‘general-purpose representation of texts in machine-readable form’, which embodies ‘a thesis about the kinds of things that are or can be important in texts’ (Sperberg-McQueen, 1991) but allows for these ‘kinds of things’ to grow and adapt. Sperberg-McQueen’s wider point, as we understand it, is to say that no mark-up can be finite and that texts must be viewed as having multiple overlapping and interacting structures. Such digital texts emerge from situated, humanistic contexts, and ‘representations obscure what they do not reveal’ (Sperberg-McQueen, 1991). This adds an additional layer of absences to those already generated by the rigidity of digital tools and the subjectivity and biases of the encoder, and also facilitates understandings of digital texts as situated scholarly interventions, rather than definitive artefacts that can be used unproblematically, for example, as pillars of attempts to induce mathematical models of culture. The digital artefacts that digital humanities make are not necessarily objective formalizations of cultural heritage documents that may be unproblematically enfolded into wider data-driven analyses, or artefacts that exist to serve the interests of a wide range of stakeholders. Secondly, the mark-up of <name> and <placename> in Hans Sloane’s catalogues has enabled us to extract large amounts of data pertaining to Sloane’s socially and geographically vast network. From just two catalogues, we have extracted the names of around 3,000 people and 600 geographical locations. By giving us a clearer picture of the geographical spread of his network, this quantitative approach attests a discrepancy of people and places in colonial and imperial contexts, and hints at just how many persons are absent from Sloane’s network. This does not recover the specific individuals whose identities were left unrecorded, yet it suggests lenses through which to experiment with Onuoha’s call, cited above, to focalize absence amidst abundance.

Our evolving approach to this involves reading the catalogues for what they do not say, as much as what they do say, by analysing patterns in mentions and omissions of person and place names, and setting such observations against the object descriptions

with which they do or do not converse. Simple word frequency analyses of the text encoded with <persName> and <placeName> offer one route into this, showing that in Sloane’s Miscellanies catalogue,⁸ for example, of the ca. 2,168 object entries it comprises, ca. 9% of entries mention both person and place in connection with an object, ca. 17% entries give a place name only, and 11% a personal name only. Of the entries that list place name only, the ten most commonly occurring are:

Place	Number of occurrences based on current data
China	62
Japan	38
East Indies	14
Turkish	14
Guinea	11
East India	9
Virginia	9
Persian	7
Brasile	6
Jamaica	6

Of the entries that mention both person and place name, the ten most commonly occurring place names are:

Place name	Number of occurrences based on present data
China	38
Japan	10
America	6
Carolina	6
East Indies	6
New England	6
Scotland	6
East India	5
Guinea	5
Hudson’s bay	5

This shift from viewing the collection with the canonical museological lens of collector (Sloane) to viewing the collection according to the places whence objects were sourced is a simple but powerful reframing. Probably reflecting the obvious geopolitical dynamics between the majority of those locations, and the UK, the individuals whose names most frequently occur in entries that mention the place names above are all apparently British or European individuals:

place name only, yet include, without attribution, information that may again have drawn on indigenous knowledge: ‘-582. A long Spear with a flaming point & ferill of brasse, and a rest for keeping steady their hands when they discharge their bowes in the East Indies. . . .’. Hints of the extent of indigenous knowledge that is silently enfolded in the catalogues are also suggested by Sloane’s references to ‘their’, which often refers to indigenous people, and occurs one or more times in ca. 2% of entries. The term ‘Indian’ (or Indians and other derivatives) occurs around 200 times, and could be used to refer to Caribbean, West Indies, India, North, and South America people.¹³

Primary research is, of course, required to deepen the perspectives suggested above, yet it is reasonable to propose that attention to the collocation of place and person information, and patterns that may be detectable in the attribution and omission of person and place, and various permutations of that combination, especially when read in conjunction with object descriptions, may open new ways to think about the detection of absence and bias in Sloane’s catalogues. Moreover, we can learn a great deal from the individuals who are recorded in the catalogues. While traditional scholarship has focused on the most prestigious or significant members of Sloane’s network, the reality that the majority of persons we have extracted have no biographical record, indicates that his network was much more socially diverse than we previously imagined. To return to *Miscellanies* again, of the ca. 540 non-unique person names listed in the catalogue, our research to date suggests that the great majority of them do not have corresponding entries in VIAF.¹⁴ This again alerts us to the inherent subjectivities of data collection: individuals included in such Name Authority Lists tend to be those who engaged in formal publication or other activities that are valorized by western societies, and seen as important building blocks of the national identities that are reinforced through discourses of ‘authorised heritage’ (see [Smith, 2006](#)). Again, we may use presence as a mirror to reflect upon the extent of absent or unidentifiable persons, and to think about how such absence can be communicated in digital editions of Sloane or flagged in the open-access datasets of Sloane’s catalogues that the project makes available for further data-driven analysis.¹⁵

3.2 Object-based research

Although Hans Sloane’s collection is object-based, it is his paper archive of catalogues, correspondence, miscellaneous papers, and written and printed secondary sources that dominate provenance research into his objects and their networks. This reliance on textual products is problematic for two primary reasons. First, not all societies have or have been allowed to have a written legacy, and the predominantly textual historical record has commonly precluded the oral, visual, and performative forms of documentation that underpins many cultures. Secondly, the atomization of objects from their written documentation also atomizes, rather than reconciles, the subjectivities and interpretations of the historical record from any potentially conflicting object-based knowledge. Text-based research can consequently entrench further the injustices of the archive by excavating the inherently selective, and subjective written historical record and perpetuating it anew. With its reliance on textual sources, data-driven research is no less text-hindered and at risk of perpetuating the absences of the archive.

One possible antidote to textual bias lies in inverting the research process to prioritize material objects above, or on an equal footing with, their corresponding documentation. What can these objects themselves—their materials, their styles, their craftsmanship—and their relationship to similar and related artefacts, tell us about their makers, traders, communities, social, and cultural functions that corresponding documentation cannot, or does not? To be clear, object-based inquiry is a far from novel methodology: it has long been accepted that objects speak their own languages and consequently require close engagement and physical handling. However, as Amiria Salmond has argued with regards to the Cook-voyage collections, there are limitations to existing object-based approaches:

[the study of artefacts] has tended to be either on establishing authoritative provenances for particular objects by linking them to documentary and pictorial sources, or on describing how these artefacts, and the assemblages of which they are part, fit into larger discussions about Enlightenment collecting and the development of scientific thought. What is often elided is the artefact itself – both as an object of (rather than

merely surrounded by) evidence, and as an instantiation of relationships forged on Pacific beaches that continue to unfold today. (Salmond, 2015, p. 33)

The potential of repositioning objects as evidence for redressing archival silences has recently been acknowledged by a growing number of museum practitioners and museologists. As a theoretical shift, it has been neatly captured by Meyer and Woodthorpe's claims that 'absence occupies a space, that absence can be made present through material objects, and that it has some agency,' (Meyer and Woodthorpe, 2008), but it has likewise been explored through projects which use material culture to rewrite historical narratives. *Artefacts of Encounter* (Museum of Archaeology and Anthropology at the University of Cambridge) in collaboration with current-day Pacific communities, for instance, sought to explore the 'nature and legacy of encounters between European explorers and Pacific islanders' (Boast and Enote, 2013), by examining extant Polynesian objects collected on voyages between 1765 and 1840. When placed in conversation with archives and texts, these artefacts challenged long-standing assumptions regarding indigenous engagement with Europeans, and the meaning and purpose of these artefacts themselves.

How might this methodology be applied to Sloane's collection in the British Museum? Take the instrument commonly known as the 'Akan drum'. Sloane's catalogue describes this object as: '1368. An Indian drum made of a hollowed tree carv'd the top being brac'd wt. peggs & thongs wt. the bottom hollow from Virginia by Mr. Clerk.' It is important to reflect on which details Sloane recorded: his knowledge of the ethnic group from whom it originated, its physical description, and from where and whom it was purchased. Yet, considering the diverse social and cultural dimensions of musical instruments, as well as the drum's undoubtedly complex journey across the Atlantic from maker to collector, there are many details missing from this account, including information regarding the drum's creator, and its function. Whether this was the result of purposeful omission, reflected the extent of Sloane's knowledge or interest in the object, or just communicated his understanding of what constituted catalogue-worthy details, we can only speculate. However, object-focused and

contextual object researches enable us to tell a different story. Curators now know that this drum, called an *Apentemma*, was made in what is now Ghana in the early 18th century, before making its way to Virginia, likely on the middle passage of a slave-trading voyage, and from there transported into Sloane's London collection. By resituating the drum in its original context, we can begin to explore questions like how such drums were used, by whom, in which occasions, how commonly, how they were made and by whom, and their cultural and social status. Details like that it was played with an open hand, not sticks; that such instruments were commonly brought by slave-ship captains to 'dance the slaves' to preserve their health during the voyage (Delbourgo, 2018); and that drums, unlike many other products and traditions of West African culture that slaves transported with them, were generally not forbidden in the Americas, all help to write absent persons and communities back into the historical record as the drum can no longer be atomized from its human context. Of course, such factual corrections are made possible by scientific developments and historical expertise unavailable to Sloane. However, they also lay bare the deep connections between 18th-century collecting culture and the slave trade, of which he would have been well aware as a collector of varied slave possessions, such as instruments, utensils, and clothing, as well as myriad objects from colonial settings. Moreover, the discovery is testament to the dangers of textual bias. The accepted 'Indian' provenance of the drum, perpetuated by ongoing museum documentation, was only called into question in the early 20th century.

A very different example from Sloane's collection is the Gray's Inn handaxe, a Palaeolithic flint handaxe approximately 350,000 years old, which was excavated alongside an elephant tusk on Gray's Inn Road, London in 1696.¹⁶ Unlike the Akan drum, the handaxe is accompanied by a rich body of textual documentation, having been an item of great curiosity and discussion that passed through many private collections before reaching Sloane. His own catalogue entry describes the object as: '246. A British weapon found wt. Elephants tooth opposite black Mary's near Grayes inn lane. . . It is a black flint shaped into the figure of a Spears point. K' ('K' stands for 'Kempe', the collector from whom Sloane received the axe). Unlike the contemporary Akan drum, Sloane could not have

known who the makers, users, or communities of such a Palaeolithic tool were. Nonetheless, there are striking absences of knowable information in his documentation which write individuals out of the historical record, such as the previous owners of the axe before Kempe, the diverse interpretations of the axe's origins, and the persons involved in its excavation, identification, and exchange. Most curiously, Sloane provides no information regarding its physical state or the conditions in which it was found. This was partially due to the state of early modern knowledge at Sloane's time. However, it was also a purposeful attempt not to disrupt received history. The handaxe, when examined as evidence, presented Sloane and his contemporaries with a possible human antediluvian world, which called biblical time into question. The handaxe's proximity to an elephant's tusk, meant that either both human and elephant had been subsumed during the Flood, thus contradicting beliefs that there were no human beings in Britain before the repopulation of the earth by Noah's sons, or that an elephant brought over by the Roman Emperor Claudius in AD 43 had died and fallen into a river. While the first hypothesis was religiously radical, the latter posed significant empirical problems. Not only was the course and flow of the Fleet river unable to deposit the quantities of sediment under which the objects were buried, but their depth was also inconsistent with that of other Roman finds. The potential ramifications of this discovery were consequently sidestepped by omitting such details from the 'legitimate' representation of the object—its catalogue entry—thus removing Palaeolithic beings from the historical record. Moreover, this was echoed in Sloane's paper 'An account of elephants teeth and bones found under ground' published in the *Philosophical Transactions* of the Royal Society in 1728 (Sloane, 1728). His detailed discussion of the Gray's Inn tusk made no mention of the handaxe, and concluded that such remains could only be proof of a 'Universal Deluge' (Sloane, 1728). We thus see, as with the Akan Drum, a fracture between the object and its textual representation. However, the handaxe indicates a purposeful omission of sensitive information from Sloane's catalogue and suggests that the description of the object carried greater importance than the physical object itself. Sloane and previous collectors did not destroy the axe for its potentially

blasphemous implications, but instead inserted it into existing, accepted narratives, and categories.

This imposed distance between the voices of the object and its representation demonstrates how textual evidence has long been the dominant script. While great energy has been spent across the centuries figuring out how best to use words to describe things, Sloane included these words and the categories they are placed into—be it early modern biblical time, or current day museum thesauri—will always take away some of the 'thing-ness' of objects. Those who use these categories most frequently—archivists, curators, etc.—are the most keenly aware of this, treating catalogues and descriptions as starting points only. However, there is a great headway still to be made in bringing objects and texts together for the wider research community, which, we propose, can be achieved through integrative, networked digital tools. While much digital heritage practice has involved creating digital surrogates of objects—photographs, 3D models, or online database records—which often only encode the existing logic of the archive, as the following two sections will demonstrate, some digital approaches facilitate the subversion or mediation of existing ontologies.

3.3 Contextual Knowledge Systems and Contextualized Ontologies

Much of the growing dissatisfaction with institutional attempts to reinstate marginalized voices into digital spaces stems from concerns that their underlying standardizing and meta-ontological approaches negate the existence of overlapping knowledge systems. This issue primarily originates in the forms of information management used by heritage institutions, such as catalogues, indexes, and thesauri, which are not inherently concerned with issues of access (Boast *et al.*, 2007), and which do not capture important contextual and experiential information, the communication of which is often left to curators and researchers at a secondary stage. Moreover, with the increasing demand for data interoperability, important object details are being further lost in order for classification to comply with standards derived from data, collections, and information management systems rather than from contextual and historical knowledge. Taylor and Gibson have demonstrated the dangers

by an ongoing mechanism involving records creators, administrators (including archivists), and users. The documentation strategy is carried out through the mutual efforts of many institutions and individuals influencing both the creation of the records and the archival retention of a portion of them. The strategy is refined in response to changing conditions and viewpoints (Samuels, 1986, p. 115).

We argue that the necessity of documenting, or at least acknowledging, the hidden contributions that were made to Sloane's catalogues, can be positioned as an ongoing issue. The asymmetries of attribution that are manifested in Sloane's catalogues are of continuing concern, not only because we are today still working through with their consequences,¹⁸ and seeking to redress them, but also because knowledge of these mechanisms directly informs questions about the breadth and depth of the contextual information that is required to ethically read and use (also in a computational sense) Sloane's catalogues. This material is not currently packaged with the hardcopy or digitized catalogues, but must be derived from secondary literature such as the scholarly studies discussed above. Moreover, as is implicit in Samuels' quote, if we define the idea of a 'user' widely, then documentation strategy can accommodate a more inclusive and participatory approach to documenting such an ongoing issue.

The participatory approach that we hope to pursue would include working with later-day communities, for example, the British Jamaican community in London, to explore ways of co-creating contextual ontologies and of visualizing representations of absence and loss in Sloane's catalogues (and those sections of the British Museum and other national collections the catalogues map). The reality is that the identity of some of those people whose knowledge and agency shaped the catalogues and Sloane's collecting practices may never be recovered. However, working with later-day communities in a participatory way, to understand, for example, how they read Sloane's catalogues and the information contained in them, could prove to be a fruitful avenue. By appropriately interfolding their readings of the catalogues, especially with regards to local and vernacular knowledge, with the digital representation of Sloane's catalogues we might find a satisfactory way of acknowledging some

of the absences upon which the catalogues are built.¹⁹ We do not claim that this would redress historical exclusions; however, rather we hope it would not reinforce exclusion by restricting decisions about how the absences that shape the catalogues can be focalized to present-day users of the catalogues to our small, white, and institutionally privileged research team.

This path is, moreover, intended to circumvent some of the issues surrounding participatory approaches in heritage. While participatory measures such as the integration of user responses to digital heritage into existing narratives have been frequently interpreted as the 'co-creation of heritage' (Ciolfi, 2013), such interactions are predominately responsive, taking place after 'profound issues of appropriation and agenda-setting (ideas of what and whose heritage) have already been made' (Taylor and Gibson, 2017). By actively involving users in the research design process, we could open Sloane's catalogues up to discussion as to what should be investigated, and how and why this should be approached. An approach often explicitly taken against the backdrop of systematic exclusion (Jules, 2016), a truly participatory approach has the potential to open a virtuous circle, and foster 'representational belonging' (Caswell, 2014; Caswell *et al.*, 2016) in communities currently marginalized in and from Sloane's catalogues while speaking to aspirations for social justice that are expressed by archivists, information professionals, and digital humanists (Punzalan and Caswell, 2016).

4 Conclusion and Recommendations for Cultural Heritage

In *Potential History: Unlearning Imperialism*, Ariella Aïsha Azoulay lays bare the fundamental misconception surrounding archival silences:

something is wrong with the paradigm of alternative history. The problem is that it proposes some things as 'hidden histories' in need of discovery, but in fact, these aren't hidden things or histories but rather open secrets known far beyond the archive and the grammar invented as guardian of its scholarly uses (Azoulay, 2019, p. 197).

The parallels raised by examining absences in Sloane's catalogues amplify the importance of current-day questions regarding the ethics of digital humanities scholarship and the duty to investigate and 'right' these data absences, especially considering the now well-accepted views on the perpetuation of subjectivities by digital tools and archives. Accordingly, like Sloane, and as Onuoha has warned, we cannot predict the future use of the data we made machine readable in *Enlightenment Architectures*, textual hauntings, and all. Sloane recorded his objects in catalogues for his own and his contemporaries' use. While he anticipated that his collection would be used by the public in bequeathing it to the nation, he could have never imagined the ways in which his catalogues would be interrogated. Similarly, we cannot foresee the uses of our datasets, a reality that requires careful planning to ensure that our intentions for the data, and our choices and rationale in data selection are made explicit. As a step towards this, in addition to the online editions of Sloane's catalogues that are an outcome of *Enlightenment Architectures*, we have also made the XML-encoded versions of the catalogues available through UCL's Research Data Repository, with their corresponding TEI headers that contain important statements related to this, and hard links to the articles that the project has published which treat of these issues at length (see https://rdr.ucl.ac.uk/collections/Enlightenment_Architectures/5231765).

Addressing absences in Sloane's catalogues consequently raises questions regarding the ethics of digital humanities and digital cultural heritage scholarship and whether there is a duty to investigate and at least alert the reader and users of the data to these data absences. Thus, we close with the following recommendations for custodians (like memory organizations) and interpreters (like digital humanists and digital historians) of data collections that have, or may be, digitized:

- (1) Reflect on whether the subjectivities that shape the management and representation of historical data within in your institution are made sufficiently clear to the stakeholders and computational agents that may act upon them.
- (2) Data and information that explain or alert users of analogue and digital historical collections should be bundled with the corresponding resources, especially when such contextual

matter is significant for reading, digitally interrogating, or transforming and interweaving that data in an ethical way.

- (3) Reflect on whether the data absences of collections, at the individual and aggregate collection, are communicated in meaningful ways.
- (4) Narratives of the role of digital technologies in the cultural heritage sector are often techno-futurist and techno-triumphalist in viewpoint. Reflect on how the use of digital technologies in this sector may, in fact, serve to reconstitute existing inequalities or reinforce hegemonies of perspective.
- (5) Investigate digital and blended methodologies that may give rise to digital collections and tools that accommodate, and focalize, a plurality of voices and communities so as to obviate the potential of digital technologies to amplify the subjectivities of the archive in ways that may be as yet unimagined

This article may be seen as a first step towards an implementation of these recommendations in in the context of *Enlightenment Architectures*, given the attention that we have drawn to the subjectivities and absences that shaped Sloane's catalogues, their potential re-inscription, and entrenchment in present and future digital heritage networks and our ongoing work to respond to them. Moreover, these are the issues that we hope to further address in subsequent research. This research should be interdisciplinary and extramural not only in the intellectual sense of crossing disciplinary spaces and boundaries, but also in the practice-based sense of attending to the situatedness of curatorial and museum practice, and collections history, so as to better understand, in full collaboration with memory institutions and participatory communities, the particular difficulties, for example, financial, or in terms of institutional mandate and national context, and indeed opportunities, of actively implementing such recommendations.

Acknowledgements

We wish to thank the anonymous peer reviewers for their corrections and insightful requests for clarification. We thank Dr Kim Sloan, Curator of British Drawings and Watercolours in Prints and Drawings,

- Mak, B.** (2014) Archaeology of a digitization. *Journal of the Association for Information Science & Technology*, **65**(8): 1515–26.
- 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. 63–81.
- Mason R. and Sayner J.** (2019). Bringing museal silence into focus: eight ways of thinking about silence in museums. *International Journal of Heritage Studies*, **25**(1): 5–20.
- McPherson, T.** (2013). Why are the Digital Humanities so white? Or thinking the histories of race and computation. In Gold, M. K. (ed.), *Debates in the Digital Humanities*. Minneapolis: University of Minnesota Press.
- Meyer, M. and Woodthorpe, K.** (2008). The material presence of absence: a dialogue between museums and cemeteries. *Sociological Research Online*, **13**(5): 127–35.
- Nieves, A. D. and Jaksch, M.** (2014). Africa is a country? Digital diasporas, ICTs, and heritage development strategies for social justice. *Journal of Interactive Technology and Pedagogy*, **6**. <https://jitp.commons.gc.cuny.edu/africa-is-a-country-digital-diasporas-icts-and-heritage-development-strategies-for-social-justice/> (accessed 25 July 2020).
- Noble, S.** (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press.
- Oldman, D. and Tanase, D.** (2018). Reshaping the knowledge graph by connecting researchers, data and practices in ResearchSpace. In Vrandečić, D. et al. (eds), *The Semantic Web – ISWC 2018. ISWC 2018. Lecture Notes in Computer Science*, vol. 11137. Springer, Cham, pp. 325–340.
- Onuoha, M.** (2016). *The Point of Collection. Data and Society: Points*. <https://points.datasociety.net/the-point-of-collection-8ee44ad7c2fa> (accessed 25 July 2020).
- Ortolja-Baird, A., Pickering, V., Nyhan, J., Sloan, K. and Fleming, M.** (2019). Digital humanities in the Memory Institution: the challenges of encoding Sir Hans Sloane’s early modern catalogues of his collections. *Open Library of Humanities*, **5**(1): 44.
- Ortolja-Baird, A.** (2020). ‘Chaos naturae et artis’: imitation, innovation, and improvisation in the library of Sir Hans Sloane. Part 1. *Library and Information History*, **36**(3): 155–74.
- Pickering, V. M. R.** (2017). *Putting Nature in a Box: Hans Sloane’s ‘Vegetable Substances’ Collection*. Ph.D. thesis, Queen Mary University of London.
- Powell, A.** (2020). *To Be Seen We must be Measured: Data Visualisation and Inequality: How Data, Bodies and Experience Entwine*. Blog, Ada Lovelace Institute, 23 June 2020. <https://www.adalovelaceinstitute.org/blog/to-be-seen-we-must-be-measured-data-visualisation-and-inequality/> (accessed 4 January 2021).
- Punzalan, R. and Caswell, M.** (2016). Critical directions for archival approaches to social justice. *Library Quarterly*, **86**(1): 25–42.
- Ramirez, M. H.** (2015). Being assumed not to be: a critique of whiteness as an archival imperative. *The American Archivist*, **78**(2): 339–56.
- Robertson, T.** (2018). Not All Information Wants to be Free: The Case Study of on Our Backs. In Fernandez, P. D. and Tilton, K. (eds), *Applying Library Values to Emerging Technology: Decision-Making in the Age of Open Access, Maker Spaces, and the Ever-Changing Library* (Publications in Librarianship #72). Chicago, Illinois: Association of College and Research Libraries, a division of the American Library Association, pp. 225–39.
- Robinson, P.** (2009). What text really is not, and why editors have to learn to swim. *Literary and Linguistic Computing*, **24**(1): 41–52.
- Salmond, A.** (2012). Digital subjects, cultural objects: special issue introduction. *Journal of Material Culture*, **17**(3): 211–28.
- Salmond, A.** (2015). Artefacts of encounter: the Cook-Voyage collections in Cambridge. In Coote, J. (ed.), *Cook-Voyage Collections of ‘Artificial Curiosities’ in Britain and Ireland, 1771–2015* (MEG Occasional Paper, No. 5). Oxford: Museum Ethnographers Group, pp. 33–73.
- Samuels, H.** (1986). Who controls the past. *The American Archivist*, **49**: 109–24.
- Sanders Garcia, A.** (2019). Silent no more: using text mining and social networks to decolonize the history of Algerian women. *2019 Global DH Symposium*. <https://www.youtube.com/watch?v=FuF6-2oZHS4> (accessed 25 July 2020).
- Sherratt, T.** (2015). *Seams and Edges: Dreams of Aggregation, Access & Discovery in a Broken World*. Discontents.com. <http://discontents.com.au/seams-and-edges-dreams-of-aggregation-access-discovery-in-a-broken-world/> (accessed 25 July 2020).
- Smith, L.** (2006). *Uses of Heritage*. London: Routledge.
- Smith, T. A. and Hann, K.** (2012). Sloane, Slavery and Science. In Walker, A., MacGregor, A. and Hunter, M. (eds), *From Books to Bezoars: Sir Hans Sloane and His Collections*. London: British Library, pp. 227–35.

- Sloan, K. and Nyhan J.** (2020). Enlightenment Architectures and the reconstruction of Sir Hans Sloane's Cabinets of Miscellanies. *Journal of the History of Collections*.
- Sloane, H.** (1728) An account of elephants teeth and bones found under ground. *Philosophical Transactions of the Royal Society*, 35: 457–71.
- Sperberg-McQueen, C. M.** (1991). Text in the electronic age: textual study and textual study and text encoding, with examples from medieval texts. *Literary and Linguistic Computing*, 6(1): 34–46.
- Taylor, J. and Gibson, L. K.** (2017). Digitisation, digital interaction and social media: embedded barriers to democratic heritage. *International Journal of Heritage Studies*, 23(5): 408–20.
- Thylstrup, N. B.** (2019). *The Politics of Mass Digitization*. Cambridge, MA: The MIT Press.
- Trouillot, M. R.** (1995). *Silencing the Past: Power and the Production of History*. Boston: Beacon Press.
- Uncomfortable Truths: The Shadow of Slave Trading on Contemporary Art.** (2007). [Exhibition], V&A, London, February–June 2007.
- Walker, A., MacGregor, A. and Hunter, M.** (2012). *From Books to Bezoars: Sir Hans Sloane and His Collections*. London: British Library.
- Weingart, S.** (2011). *Fidelity at Scale: How Time Affects Data*. <http://scottbot.net/wp-content/uploads/2011/11/finland.pdf> (accessed 25 July 2020).
- ## Websites
- A Colony in Crisis: The Saint-Domingue Grain Shortage of 1789.** (2020). <<https://colonyincrisis.lib.umd.edu/>> (accessed 25 July 2020).
- Breve.** (2020). <http://hdlab.stanford.edu/breve/> (accessed 25 July 2020).
- CIDOC-CRM.** (2020). <http://www.cidoc-crm.org/> (accessed 25 July 2020).
- Documenting the Now.** (2020). <https://www.docnow.io> (accessed 25 July 2020).
- Enlightenment Architectures: Sir Hans Sloane's Catalogues of his Collections.** (2020). <<https://reconstructingsloane.org/enlightenmentarchitectures/>> (accessed 25 July 2020).
- Enslaved: The People of the Historic Slave Trade.** (2020). <<http://enslaved.org/>> (accessed 25 July 2020).
- Eyes on the Past Project.** (2020). <<http://eyespast.heroquapp.com/>> (accessed 25 July 2020).
- Freedom on the Move.** (2020). <<https://freedomonthe.move.org/>> (accessed 25 July 2020).
- Invisible Australians: Living under the White Australia Policy.** (2020). <<http://invisibleaustralians.org/>> (accessed 25 July 2020).
- Mount Vernon Slavery Database.** (2020). <<https://www.mountvernon.org/george-washington/slavery/slavery-database/>> (accessed 25 July 2020).
- O Say Can You See: Early Washington D.C, Law & Family.** (2020). <<http://earlywashingtondc.org/>> (accessed 25 July 2020).
- Palladio.** (2020). <<https://hdlab.stanford.edu/palladio/>> (accessed 25 July 2020).
- The Georgetown Slavery Archive.** (2020). <https://slaveryarchive.georgetown.edu/> (accessed 25 July 2020).
- The Nana Project.** (2020). <<https://www.thenanaproject.org>> (accessed 25 July 2020).
- Topotime.** (2020). <<http://dh.stanford.edu/topotime/index.html>> (accessed 25 July 2020).
- VIAF® (Virtual International Authority File).** (2020). <<http://viaf.org/>> (accessed 25 July 2020).

Notes

- 1 This article includes historic language that some people might find distressing.
- 2 One notable exception being [Kidd \(2019\)](#).
- 3 One such example is Documenting the Now (DocNow), which provides a variety of tools to help 'archivists, activists and researchers work with social media data', thus enabling marginalized communities to take control of digital archiving practices of selection ([DocNow, 2020](#)).
- 4 These moments are: when sources are created; when these sources are collected in archives; when the sources are retrieved from the archives by researchers; the moment of retrospective significance (making history) ([Trouillot, 1995](#)).
- 5 Delbourgo's argument that Sloane is better considered as a hub around which his collection was formed by a diverse network of individuals advances Caygill's earlier attention to claims that Sloane can hardly be credited with having single-handedly founded his own collection ([Caygill, 1994](#)).
- 6 The perhaps greatest absence in the catalogues is the partial origins of Sloane's wealth in Jamaican plantations which afforded him the privilege of amassing such an extensive collection.

