Journal of Southern African Studies

The Politics of the Parallel Archive: Digital Imperialism and the Future of Record-Keeping in the Age of Digital Reproduction

Keith Breckenridge[a]

[a] Wits Institute for Social and Economic Research, University of
the Witwatersrand
Published online: 23 May 2014.

PLEASE SCROLL DOWN FOR ARTICLE

Routledge
Taylor & Francis Group

*The Politics of the Parallel Archive: Digital Imperialism and the Future of Record-Keeping in the Age of Digital Reproduction**

KEITH BRECKENRIDGE

(Wits Institute for Social and Economic Research, University of the Witwatersrand)

This paper takes as its subject the fact that digital archival production – of existing materials and born-digital records – has collapsed in contemporary South Africa, and it offers some arguments about why it is important to reverse this process. The current situation can be explained by the fact that digitisation has been widely described as a form of intellectual imperialism, a characterisation that echoes influential strands of postcolonial theory and South African nationalism. The reasons for this unusual understanding lie in the difficult history of the last major digitisation effort, the Mellon-funded collaboration between Aluka and the Digital Imaging Project of South Africa (DISA). The paper reconstructs that project in some detail in an effort to understand what went wrong, arguing that in place of the geopolitical explanation that many participants adopted, most of what went wrong was much more narrowly technological. Yet, the same technological issues have already been great assets to South African researchers, holding out the promise of solutions to some pressing local difficulties of digital preservation and archival assembly. The last section of the paper takes up some of the reasons why scholars need to take digital record-keeping much more seriously than they have to date – chief amongst these being the fertile possibilities of forgery and impersonation.

The South African National Archives has fallen on hard times. Historians complain habitually of the dishevelled state of the provincial repositories, of disorder in the stacks, noise in the reading rooms, and of failures of governance.[1] For years the service, like so many other branches of the state, has been paralysed by the internal disciplinary investigation of senior managers.[2] Some of this lament seems a little overdone, especially for those who can recall having to get past the presidential guard to work in the basement of the Union Buildings. Much more worrying is the apparent collapse of record-keeping in the current bureaucracy, a failure that will make the history of the current era oddly difficult to write from within the state, but, much more importantly, endangers both the state and any possibility of real parliamentary accountability. Ironically, all of this is happening as the Internet, for the first time, makes it possible to conceive of an archival project outside of the state which does

---

1   Shula Marks, 'Do Not Let Our Archives Turn to Dust', *Mail and Guardian*, 29 June 2012. http://mg.co.za/article/2012-06-28-do-not-let-our-archives-turn-to-dust; Chris Saunders, 'National Archives Are a National Disgrace', *Business Day*, 15 August 2011. http://www.businessday.co.za/articles/Content.aspx?id=150654; Brenton Maart, 'Field Brat's Blog from the Bundu: The Mthatha Archives – Opinions – Archival Platform', *The Archival Platform*, 26 March 2012. http://www.archivalplatform.org/blog/entry/field_brats_blog_from_the_bundu_the_mthatha_archives/.

2   Jo-Anne Duggan, 'What's Up at the Department of Arts and Culture', *Archival Platform*, 4 November 2010. http://www.archivalplatform.org/news/entry/whats_up/.

the work of holding the bureaucracy to account and builds the documentary resources required to write good history. In this paper I want to consider some of the implications of this moment and to suggest what it might mean for the National Archives. But I also want to attempt to restore the political reputation of the archive itself.

The paper that follows is composed of four parts. It begins with a brief review of the current state of digitisation projects in South Africa, concluding that the production of online materials has effectively collapsed (especially viewed in global and historical comparison). The paper then examines the reasons for the very rapid decline in digitisation, reconstructing the story of the debacle of the Digital Imaging Project of South Africa (DISA)–Aluka project, a generously funded international collaboration that ran between 2003 and 2009. It was this bad-tempered collaboration, more than anything else, that gave substance to the charges of digital imperialism that have brought the production of electronic archives to a standstill. A close consideration of the conflicts involved in this project offers some useful lessons for the design of future collaborations. The paper then turns briefly to the existing sources that are available online in an effort to show that South Africans have vastly more to gain from the Internet than they could ever lose to it. In fact, South African researchers have every reason seriously to think through the opportunities and the dangers of relying upon online materials. To make this point the paper turns, finally, to some of the new problems of forgery and authenticity that lie in wait for us.

If there was a debate over the political work of the archives in South Africa, it might take the form of a bad-tempered dichotomy – a zero-sum argument between ghostly caricatures emerging from the writing of Edward Said and Jürgen Habermas. Some scholars invoke Said's compelling arguments about the place of an assiduously assembled library of Orientalist scholarship as both cause and product of Anglo-French colonisation. The archive, in this view, has been an essential instrument of global racial and cultural domination.[3] Others (and here I would have to position myself) cling to the liberal (and Habermasian) view that an effusive official archive is a prerequisite of representative democracy and a potential source of good politics.[4] Yet, like all dichotomies, this is probably an unhelpfully simplified description of the problem. No doubt, first, there are other ways of viewing the question of the archive. And, second, it is obvious that Said's stress on the epistemological work of Empire and Habermas' arguments about the institutional foundations of public discourse both have useful things to say about the place of archives in a postcolonial democracy. In this paper I concede that we need to understand and hold what is valuable from both of these positions. But I also want to show that the expansion of the Internet in our lifetimes rubbishes many of the foundational claims about the politics of the documentary archive.

Viewed comparatively, against, for example, the contemporary efforts of the British National Archive (formerly the Public Record Office), the British Library, the Library of Congress or the Australian National Library, the official production of digital archives inside South Africa has simply collapsed. The commercial development of newspapers and other valuable privately held materials seem, also, to be comparatively stillborn. One reason for this, and the focus of the second part of this paper, is the fraught relationship between archives

---

3 Edward W. Said, *Orientalism* (New York, Vintage Books, 1978), e.g., pp. 41–2; Bernard S. Cohn, *Colonialism and Its Forms of Knowledge: The British in India* (Delhi, Oxford University Press; 1997), p. 19; G. Steinmetz, *The Devil's Handwriting: Precoloniality and the German Colonial State in Qingdao, Samoa, and Southwest Africa* (University of Chicago Press, 2007), pp. 24–7; Clifton Crais, *The Politics of Evil: Magic, State Power, and the Political Imagination in South Africa* (Cambridge, Cambridge University Press, 2002); The local protagonist is Premesh Lalu, *The Deaths of Hintsa: Postapartheid South Africa and the Shape of Recurring Pasts* (Cape Town, HSRC Press, 2009).

4 Jurgen Habermas, *The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society* (Cambridge, Polity Press, 1992), especially pp. 57–66, 89–102; V.S. Harris, *Archives and Justice: A South African Perspective* (Chicago, Society of American Archivists, 2007), pp. 269 – 85.

and empire in South Africa. Some scholars have argued, with good grounds for doing so, that digitisation works as a form of heritage theft, making valuable documentary resources more easily available to foreigners than they are to citizens.

As the paper shows, the critics of digitisation as being a manifestation of imperialism have insisted that the most important issues at stake are social and political ones, pitting the ambitions of a resourced North against a labour-providing underdeveloped South. And they have argued that digitisation will sustain long-established imperial structures of knowledge and power. I suggest that this characterisation of digital imperialism has been very unhelpful and that it effectively obscures the fact that the difficulties often experienced in international collaborations are actually technological ones. Most importantly, the argument neglects the enormous (sometimes accidental and often pirated) resources that the Internet already works to provide to South African researchers.

The digital repositories available on the Internet shred some of the long-cherished assumptions about the relationships between the archives, historians and the nation-state. The most obvious of these is that history, and the archives that produce it, can no longer easily be described as the handmaiden of nationalism: both the subjects and the sources of historical writing have come significantly adrift of individual states and their politics. Another caricature of historical research – the fetishisation of the document in the dusty archive as the uncomplicated source of truth – is also at risk: historians must now work in a sea of millions of widely available digital records. But, and this is important, they have also been stripped of even the dubious safety of the original paper document. Gone are the dust, textures, smells and bindings that have provided the clues of documentary authenticity. I show that digital records can be forged and distributed trivially, and it is clear that a new set of methodologies that focus on the collective study of meta-data, dense inter-textual analysis and document registration must still be developed.

In thinking through the effect of the Web, and digital archives, on the project of official record-keeping, it is helpful to return to Walter Benjamin's famous essay on the technological effects of mechanical reproduction on twentieth-century art.[5] Where Benjamin was interested in the effects of massive mechanical production on the subjects and audience of modern cinema, I want to think through the political implications of the many web-based archives – each exploiting the possibilities of the free, rapid and dispersed reproduction of digital archives – for the project of official record-keeping. The key to this new technopolitics is the extraordinary ease and efficiency of copying and distributing massive amounts of digitised material on the Internet. The act of reducing paper to electronic bits makes it possible to reproduce entire libraries without a meaningful cost in time or money. It is true that serious preservation still requires very high-resolution master images, whose storage entails enormous amounts of disk space; this is a technological problem that is only visible in the background, imposing heavy costs on all digital libraries. But for users on the Web, the replication of working copies of digital libraries has become practically limitless. This unrestricted and effortless replication also raises a set of interesting problems – policing ownership is one of them. Another is a growing problem of reputability: if digital objects can be copied and distributed almost without restriction, who is responsible for establishing their provenance? In many respects it is clear that the project of massive online archive construction has escaped the grasp of the National Archives. But the major problem of authenticity will remain, very probably growing more significant over time. And there are opportunities here for a renewed National Archives project that might help to restore the political significance of the archives service as the main instrument of parliamentary accountability.

---

5 Walter Benjamin, 'The Work of Art in the Age of Mechanical Reproduction', in Hannah Arendt (ed.) *Illuminations* (New York, Schocken Books, 1969).

Digitisation in Contemporary South Africa

Over the last five years, the digitisation of archival documents in South Africa has ground almost to a halt. There are, certainly, some important digitising projects currently under way. Sabinet, in my view the most valuable South African online resource, has a large grant from the Carnegie Corporation to scan and publish the contents of 200 African journals, and several related electronic journal publishing projects.[6] The Institute for Contemporary Affairs in Bloemfontein, the Heritage Foundation in Pretoria and the National Library in Cape Town all have plans in place for ongoing digitisation of their materials. But the online results have been desultory and unpromising. The Rock Art Research Institute, supported by the Mellon Foundation, is currently busy with the digitisation of the tens of thousands of images owned by museums around the country.[7] Google has recently supported two trophy projects: the digitisation of the Mandela archives and the records of the Tutu Peace Centre. And most of the universities are building digital repositories of their own unpublished research papers and dissertations. Yet these exceptional projects, important as they are, highlight the rule that South African efforts of digitisation have entered a troubling period of quiescence.[8]

What is particularly noticeable about the contemporary period is the absence of workable plans for the production of digital records and, especially, public records. The electronic version of the South African parliamentary record, Hansard, consists of a set of detached Microsoft Word files, which cannot be viewed online and must be downloaded separately. There are no plans in place to produce a hypertext or searchable version of the current proceedings, nor of the typeset volumes of the paper-based record that date back to 1910. Similarly, the National Archives Service has no plans for the digitisation of its existing paper holdings and, much more worryingly, no capacity for the systematic extraction of the electronic records that have been generated by the state since the early 1990s.[9] Even the universities, and the Research Councils that support them, have only ad hoc plans for the ongoing preservation of digital materials.

This digital hiatus stands in marked contrast to the atmosphere a decade earlier. In 2005, the proliferation of large-scale digitisation projects under way in collaboration with European and American partners prompted some of those closest to these efforts to warn of a 'virtual stampede'.[10] The decline is a surprise, especially in light of the very rapid and productive turn to digitisation in the 1990s. The collapse may reflect ongoing factional disputes (see below) and the churning of senior personnel in the key organs of policy making. In the years since the complaints about digital proliferation were first made, the National Heritage Council and the Department of Arts and Culture have struggled, without a formal result, to develop a national policy on digitisation.[11]

---

6 Sabinet, 'Sabinet Gateway Awarded $1,8m Grant from Carnegie Corporation to Establish African Online Journal Archive', press release, 8 June 2008.
7 Rock Art Research Institute, 'The African Rock Art Digital Archive – About', web page. Retrieved 25 April 2012 from http://www.sarada.co.za/about/.
8 For a comprehensive review of digitisation projects, see Roy Page-Shipp, 'An Audit of Digitisation Initiatives, Ongoing and Planned', in 'South Africa Report for Stakeholders Workshop', 2 March 2009.
9 Harris, Archives and Justice, pp. 160–5.
10 Premesh Lalu, 'The Virtual Stampede for Africa: Digitization, Postcoloniality and Archives of the Liberation Struggles in Southern Africa', *Innovation* 34 (June 2007), p. 30; Michele Pickover, 'Notes Re Challenges of the Aluka Projects' (National Heritage Council, 2007).
11 Thendo Ramagoma, 'Legal Review – Collaboration Agreement Between the University of Kwazulu Natal and Ithaka Harbors, Inc', 24 November 2005; M & R Badiri Consultants, 'Report on the Workshop on the Protection and Promotion of Digital Heritage Resource: Towards a National Startegy [sic]', 12 March 2007, DISA Collection, Historical Papers, Wits University; Page-Shipp, 'An Audit of Digitisation Initiatives, Ongoing and Planned', in 'South Africa Report for Stakeholders Workshop', National Policy on the Digitisation of Heritage Resources (National Department Of Arts And Culture, August 2010).

It is possible, also, that access problems – which are themselves a product of the weakness of the network infrastructure on the African continent – contribute to popular and government disinterest in digitisation. 'The Internet', as Michele Pickover has put it, 'relies on technology that is much less accessible and much more expensive in the South than in the North'.[12] But this idea seems misplaced for several reasons. The first is that network capacity has expanded very dramatically over the last five years, especially in the tertiary sector which hosts much of the digitisation effort. Ironically, the decline in digitisation coincided with a renewal of the popular web in South Africa. The opening of new undersea cables in 2007 prompted very significant drops in the cost of domestic data and thickening of the network by the national telecoms monopoly.[13] At the same time the introduction of smart phones and cost-effective data packages tailored to the southern African market saw, for the first time, significant growth in Internet usage amongst the general population.[14] Ironically, the decline of digitisation has coincided with the moment at which Internet access, especially among young people, is becoming a normal part of South African life.

The problem seems not to be a resource issue. The cost, in South Africa, of the tools of digitisation and publishing fell dramatically in this period – especially the cost of high-definition cameras and cloud-based servers. (In 2012, cameras that can produce high-resolution images at very high speed could be purchased for around R300, and the total annual cost of an Internet-based server was around R1,200; these amounts represent small fractions of the costs for the same tools a decade ago.) At the same time, a host of powerful, user-friendly and free open-source tools – like Ubuntu Linux, Drupal and Omeka – reached maturity, offering standards-based, reliable and easily learned tools to anyone interested in using them. And both software and the hardware tools were much more easily available in South Africa than was the case a decade earlier.

Nor was there any shortage of government funding in general: in 2013, the budget of the Archives (harnessed unhappily to the Library Services) was nearly R1 billion. The decline has coincided with a period of elaborate investment by the National Department of Arts and Culture into the development of a policy on digitisation culminating, in August 2010, in the publication of 30 key policy objectives.[15] It is likely, of course, that the decline in the production of digital archives can be attributed to the same deficiencies that constrain state action in other areas. Shortages of skills, of funds not already earmarked for salaries, of competent and tenacious managers and, perhaps most importantly, of trust have battered the provincial and national archives services over the last decade.[16] The public activities of the National Department of Arts and Culture also suggest that the leaders of the African National Congress much prefer the mute testimony of the monuments and artefacts of a moribund nationalism to the labyrinthine, and sometimes racially derogatory, testimony of South Africa's archival record. The fostering of archival capacity, especially of the sort that might

---

12 Michele Pickover, 'Negotiations, Contestations and Fabrications: The Politics of Archives in South Africa Ten Years after Democracy', *Innovation* 40 (June 2005), pp. 1–11.

13 Tertiary Education and Research Network of South Africa (TENET), Annual Report 2007–8. Retrieved 25 April 2012 from http://www.tenet.ac.za/about-us/tenet-documents-and-publications/annual-reports/.

14 Google Public Data Explorer, 'World Development Indicators and Global Development Finance', accessed 30 March 2012. http://www.google.com/publicdata/explore?ds=d5bncppjof8f9_&met_y=it_net_user_p2&idim=country:ZAF&dl=en&hl=en&q=south+african+internet+usage.

15 National Department of Arts and Culture, *National Policy on the Digitisation of Heritage Resources*.

16 Page-Shipp, 'An Audit of Digitisation Initiatives, Ongoing and Planned', in 'South Africa Report for Stakeholders Workshop'; Chris Saunders, 'National Archives Are a National Disgrace'; Verne Harris, 'The Archives Advisory Council: A Chequered Past, with Hope for the Future?', *The Archival Platform*, 18 August 2009. http://www.archivalplatform.org/blog/entry/the_importance_of_making_nominations_to_the_archives_advisory_council/; Maart, 'Field Brat's Blog from the Bundu: The Mthatha Archives – Opinions – Archival Platform'.

improve the citizens' ability to hold the bureaucracy to account, does not appear on the government's public agenda.

This disinterest in the new tools of digital government (as applied to official records) also stands in contrast to the ANC government's otherwise enthusiastic embrace of computerised administrative solutions.[17] It would be comfortably within the general pattern of governance over the last decade – like the provision of welfare or the building of roads – for the archive service to outsource the digitisation of all state records to one of the large international information technology corporations. The fact that that has not happened can be traced, in significant part, to the ignominious account that South African scholars and archivists have offered of archival work in general – and digital archives in particular – over the last decade.

What is particularly notable about this period is that both the state and private record owners seem to have little interest in digitising and classifying the most important sources in their own archives. This lethargy is marked in comparison with the full-throated transformation of archival materials that is under way around the world. Many of these global efforts have been undertaken to reduce costs and reach a new and very large online audience of users.[18] An excellent example is the complete digitisation of the British Hansard records from 1803 to the present by a small, unfunded team at the UK Parliament[19] which provides an extraordinarily powerful, universally available and free research tool, while at the same time significantly altering the public work of the legislature. The stillbirth of similar projects is an important but mostly hidden opportunity cost of the current malaise in South Africa. Interestingly, the current situation in South Africa also compares badly with the relatively recent history of archival practice in the country. At its apex the apartheid state was a precocious early adopter of digital archival technologies. In the early 1970s, the South African state was one of the first customers for the IBM STAIRS system. And it was that system of centralised information control that provided the architecture of the current very useful online catalogue of archival holdings.[20] One of my objects in this paper is to show that archivists and historians can exploit those efforts much more effectively – and easily – than they are currently doing.

While weaknesses of policy, funding, and administrative capacity have all had some influence, the current crisis in the digitisation of archival records is primarily the result of the collapse of a single important project. For ten years after 1997, DISA was a joint project between the University of the Witwatersrand (Wits), University of KwaZulu-Natal (UKZN) and University of Cape Town (UCT), funded by the Mellon Foundation and based in Durban. By 2003 DISA had begun to function as the primary source of digitisation expertise and capacity, setting standards and servicing a scanning infrastructure based at the universities, including Fort Hare and the University of the Western Cape (UWC), and the National Library, and with strong policy-making relationships with the National Archives and the

---

17 Department of Arts and Culture, South Africa, 'Media Releases', Departmental website, 15 November 2011. http://www.dac.gov.za/media_releases08_06.htm. See, for example, Keith Breckenridge, 'The Elusive Panopticon: The HANIS Project and the Politics of Standards in South Africa', in Colin Bennett and David Lyon (eds), *Playing the ID Card: Surveillance, Security and Identity in Global Perspective* (London, Routledge, 2008), pp. 39–56.

18 This was the primary justification for JSTOR: see Roger C. Schonfeld, 'JSTOR: A Case Study in the Recent History of Scholarly Communications', *Program: Electronic Library and Information Systems* 39, 4 (1 December 2005), pp. 337–44, doi: 10.1108/00330330510627953.

19 See Hansard web page, undated. http://hansard.millbanksystems.com

20 S.A. Argiefblad, 'Die Rekenaar: Ontsluiter van 60 Duisend Meter Argiewe', *South African Archives Journal* (1974), pp. 12–20; Charles P. Bourne and Trudi Bellardo Hahn, *A History of Online Information Services, 1963–1976* (MIT Press, 2003), pp. 130–3.

National Heritage Council.[21] Five years later the project lay in ruins, its relationship with its primary funder wrecked, the staff dispersed and its technical capacity – including the hardware required to preserve the digital holdings – decaying.

The problems that undermined DISA were viewed by many of the protagonists on the South African side as issues of control and dependency, and they were characterised quite quickly in the rhetoric of neo-colonialism. These claims resonated with an important scholarly account of the official archive itself. The idea that the official archive was a strategic instrument of imperial and colonial control (which can usefully be described as the Chicago view – following the work of Cohn, Dirks, Appadurai and Chakrabarty) has strongly influenced Lalu's work in particular.[22] And these accusations of renewed archival imperialism found a very sympathetic audience in the state institutions designed to regulate (and foster) digitisation and access. In these accounts, DISA's problems after 2002 were described as ideological and political, and not technological. Yet, as I will show, the opposite seems to have been the case.

Digitisation as Imperialism Reinvented

The Digital Imaging Project of South Africa (DISA) was first funded in 1997 to undertake the digitisation and publishing of important documents from the struggle era. The project was based at the Campbell Collections at the old University of Natal, but it included materials and labour from many other archival repositories, assembling an invaluable collection of documents and images. From the outset DISA was conceived as a project that combined digital innovation with redress, producing new online archives of Struggle media. The first phase of the project received a grant of US$375,000 from the Andrew W. Mellon Foundation.[23] Between 1998 and 2003 DISA scanned some 70,000 pages from 40 serials in five libraries producing an invaluable archive of materials which is unreliably available at http://www.disa.ukzn.ac.za. In the process of developing these extensive materials, DISA fostered technical capacity at each centre and a growing institutional awareness of the possibilities of digital preservation and distribution.[24]

The first phase of the DISA project coincided with the establishment of JSTOR, which began in earnest at the University of Michigan in 1995. Both projects were funded by Mellon, but JSTOR quickly became the foundation's most important project. Indeed, shortly after it began the Foundation removed JSTOR from Michigan and created its own institutional architecture. The results have had genuinely revolutionary effects on the production of

---

21 Heather Edwards, 'Report to the Andrew W. Mellon Foundation on the DISA 2 (Phase 2) Grant for the Year 2005', 29 March 29 2006, DISA Collection, Historical Papers, Wits.

22 The genealogy of this intellectual movement can be sketched like this: Said, *Orientalism*; R. Guha, 'The Prose of Counter-Insurgency', in *Selected Subaltern Studies* (New York, Oxford University Press, 1988), pp. 45–84; Bernard S. Cohn and Nicholas B. Dirks, 'Beyond The Fringe: The Nation State, Colonialism, and The Technologies of Power', *Journal of Historical Sociology* 1, 2 (1988), pp. 224–9, doi: 10.1111/j.1467-6443.1988. tb00011.x; Cohn, *Colonialism and Its Forms of Knowledge: The British in India*; Nicholas B. Dirks, 'Annals of the Archive: Ethnographic Notes on the Sources of History', in Brian Keith Axel (ed.), *From the Margins: Historical Anthropology and Its Futures* (Durham, Duke University Press, 2002), pp. 47–63; Crais, *The Politics of Evil: Magic, State Power, and the Political Imagination in South Africa.* This genealogy, and others, is explored in Ann Stoler, 'Colonial Archives and the Arts of Governance: On the Content in the Form', in Carolyn Hamilton *et al.* (eds), *Refiguring the Archive* (Cape Town, David Philip, 2002), pp. 83–102. See, especially, P. Lalu, *The Deaths of Hintsa*; Lalu's 'Virtual Stampede' cites Stoler and Crais as key figures in the South African debate.

23 Report from 1 January 1998 through 31 December 1998 (New York, The Andrew W. Mellon Foundation, 1999), n. 3.

24 Dale Peters and Heather Edwards, 'DISA's Relationship with Mellon Fund and Aluka', 20 December 2007, DISA Collection, Historical Papers, Wits.

knowledge globally. The new deep and wide search capacity offered by JSTOR, and the potential cost savings of digital archiving, very quickly attracted commercial subscriptions from research libraries around the world.[25] By 2004, JSTOR was drawing subscriptions from some 2,000 libraries (paying an average US$25,000 per annum).[26] JSTOR and the first phase of the DISA project were both examples of the tremendous research value that can be gained from the digitisation of long runs of typeset serials.

The second phase of DISA was strongly influenced by the Mellon Foundation's establishment of Ithaka, a co-ordinating body for digital scholarship. This new organisation took control of JSTOR in 2002. From the outset, South Africans, and their materials, were surprisingly important to Ithaka's mission. Mamphela Ramphele was appointed to the organisation's board and the digitisation of Struggle materials to form 'JSTOR-like electronic collections' formed an important part of the organisation's founding mission.[27] This project – of including South African digital content in a globally distributed and properly sustained electronic archive on the JSTOR model – had been mooted as early as 1998, and more than anything else it reflects the relatively heavy investments that the Mellon Foundation has long made in South African higher education.

When the people involved with the DISA project began to look for a second round of funding, they quickly learned that Mellon was interested in them collaborating with a new digitisation project being started by Ithaka. The new project was called Aluka; it was fostered by Allan Isaacman, Gail Gerhart, and Premesh Lalu, and intended to 'develop electronic content concerning Africa and other parts of the global south that could be made available online'.[28] Their initial proposal for an extension was temporarily granted, and the DISA committee began negotiations to provide content to the Aluka project. Unlike the earlier focus on easily standardised serials in both JSTOR and DISA, Aluka had in mind an archive of high-value African objects – museum representations of landscapes, biological accounts of plants, and an archive of carefully selected documents to be produced by DISA. After the initial enthusiasm for collaboration – fostered by the prospect of a new round of very generous funding (the project was granted US$1.15 million) – the marriage of DISA and Aluka almost immediately became dysfunctional.[29]

From the outset Aluka made many promises that proved unachievable. Drawing on the arguments from *Refiguring the Archive*, an important anthology of work sponsored by the Wits Graduate School in the Humanities and published in 2002, Premesh Lalu urged the scholars gathered for the first Aluka content selection meetings in Durban to assemble a 'new form of archive', one that would, as he later put it, 'unsettle the seamless narrative of the liberation struggle'.[30] Key to this new archive, and the new narrative it would support, was a regional collection effort that would emphasise the 'transnational struggle against Apartheid'. This project of regional documentation was undermined almost immediately by a second aim – also unrealisable – to employ scholars to find rare artefacts of 'high socio-political interest'. These two objectives – the broad geographical scope and the use of specialist political historians to source rare and valuable artefacts – effectively undermined each other. But it was the failure of the third goal of the project – to 'develop knowledge and expertise in digital imaging amongst librarians and archivists in SA' – that contributed most to the

25 Report from 1 January 1998 through 31 December 1998.
26 Schonfeld, 'JSTOR'.
27 'Ithaka Background Document', 9 January 2003, DISA Collection, Historical Papers, Wits.
28 Allen F. Isaacman, Premesh Lalu, and Thomas I. Nygren, 'Digitization, History, and the Making of a Postcolonial Archive of Southern African Liberation Struggles: The Aluka Project', *Africa Today* 52, 2 (2005), pp. 55–77.
29 Stuart Saunders to Heather Edwards, Chris Saunders, and Michele Pickover, 'Disa', 11 May 2007.
30 Isaacman, Lalu, and Nygren, 'Digitization, History'.

accusations of neo-colonialism and, in the long run, weakened the production of digital archives locally.[31]

Anxieties about the embedding of neo-colonial economic relationships into the structures of digital archives were loudly expressed from the outset. Indeed, drawing on the notorious private sale of the Mandela trial prosecutor's records, Michele Pickover had warned of the dangers to public archives of the new markets in digital records even before Aluka began.[32] After the first Durban gathering of archivists and historians, the project leaders reported that 'numerous participants expressed suspicion that this digitising initiative would be yet another North American project designed to appropriate Africa's patrimony and subvert intellectual property rights and national heritage'.[33] The South African participants worried that the new online records would remove the local resource advantages available to local scholars, freeing northern researchers of the requirement to use archives on the continent, and that the 'international standing of their repositories' would be diminished.

These criticisms about the global politics of the digital divide resonated with postcolonial anxieties about imperialism renewed, but the issues were really technopolitical.[34] The conflicts between Aluka and DISA were directly related to the hardware expectations of each project and, especially, to the technical problems of sustaining disk storage for high-resolution masters in perpetuity. From the early versions of the funding proposal for Phase 2 of the DISA project, the Mellon Foundation had been intent on feeding DISA's data in to the Aluka project. The collaboration agreement that University of KwaZulu-Natal and Ithaka signed early in 2004 required DISA to grant Aluka a 'worldwide, perpetual, nonexclusive, royalty-free license' to all of its content – both the new, carefully chosen artefacts still to be captured and the long series of struggle periodicals that had been done during the first phase.[35] The breathtaking implications of this requirement – and the destructive effects of digitisation on the sentiments of ownership and the value of labour – became clear when one of the archivists at Aluka asked that all of the data from the first five years of the DISA project be copied to a hard disk and shipped to Princeton.[36] The realisation that Aluka had the right not only to publish content that was available on the DISA website but also had unrestricted control of the masters of the data itself left the South African archivists in a state of astonishment. Adding to this disheartening recognition of dispossession was Mellon's insistence that the artefacts that Aluka was able to gather from research universities in the North (from, in theory, the substantial South African collections at Oxford or Northwestern) would require individual licensing agreements between DISA and the providing institution for each item.[37] Under the design of DISA 2, the flow of data from South Africa was to be immediate and unrestricted, while the movement of records from Princeton to Durban was hedged about with careful licensing requirements designed to protect the value of the holdings in the US and Britain. This concern with the transfer of rights to the DISA content was taken up by the statutory body charged to 'co-ordinate heritage management', the

---

31 Aluka Content Committee, 'Meeting Held at the Campbell Collections, UKZN', 5 February 2004, DISA Collection, Historical Papers, Wits.
32 Pickover, 'Negotiations, Contestations and Fabrications'.
33 Isaacman, Lalu, and Nygren, 'Digitization, History'; Aluka Content Committee, 'Meeting Held at the Campbell Collections, UKZN'. The minutes suggest that these doubts were expressed by Michele Pickover from Wits and Sifiso Ndlovu from SADET.
34 For a discussion of the material politics of technological processes, see Gabrielle Hecht, *The Radiance of France: Nuclear Power and National Identity after World War II* (Cambridge, MA, MIT Press, 1998), pp. 15–17.
35 Peters and Edwards, 'DISAs Relationship with Mellon Fund and Aluka'.
36 DISA Governing Committee, 'Minutes of the DISA Governing Committee Meeting Held at the Campbell Collections, University of Kwa-Zulu Natal, Durban', 4 April 2004, DISA Collection, Historical Papers, Wits.
37 Heather Edwards and Claire Wright, 'Principles to Be Included in the Agreement Between DISA and Aluka', 3 September 2003, DISA Collection, Historical Papers, Wits.

National Heritage Council (NHC), late in 2005: 'what is the wisdom', the council's secretary asked, 'of granting Aluka irrevocable rights and for that matter in perpetuity?' And the NHC formally asked whether 'ordinary South Africans, students in particular, would be able to freely access the digitised material?'[38] Adding to this growing sense of subordination and exclusion through licensing, after the collaboration had been under way for a year it became clear that – notwithstanding the noble developmental commitments in the original proposal – that no technical support would be forthcoming from Aluka.[39]

The clumsy arrangement of the licensing rights to material produced by DISA was a serious problem – one that certainly informed the sense of neo-colonial subordination – but it was not the most important difficulty that the project faced. That was a much simpler failure. At a workshop in New York in February 2005, the Aluka managers reminded the South Africans that it is 'critically important to achieve mutually agreed production targets'.[40] In fact almost nothing was being done, and scanning was taking place at a glacial pace.

The reason for this standstill was double-sided, reflecting a tension built into the design of the project between careful selection and broad scale. Aluka was conceived as an African JSTOR, a massive archival repository of research materials searchable from a single point. Reputability was to be one of the distinguishing features of the content, but scale was another.[41] When Tom Nygren, the Executive Director of Aluka, was briefing the South African historians charged to select the new content, he sketched out an ambitious programme of acquisition, where a medium-sized collection would run to 10,000 pages.[42] Archival collections of that size outside of the state and mining industry are rare in South Africa (if they exist at all), and it is easy to see how archivists might worry that they would be out of employment altogether.[43]

But this desire for comprehensive scope was undone by another set of intractable requirements. The first of these was Lalu's call for a new kind of archive, one that stood outside of the state, outside of the liberation movements, and astride the region's political borders. In practice his appeal was almost immediately undermined by the appointment of the national selection committees.[44] The detailed taxonomy of Gail Gerhart's original plan for collection – which stressed the recovery of the documents of the political organisations, carefully specifying the exact number of pages for each category of the struggle – left little space for experiment or the harvesting of long-running serials.[45] And then, finally, the commissioning of a large group of political historians to ferret out high-value artefacts for each of the categories of Gerhart's taxonomy proved disastrous. Only a few among them were able to generate the targets specified in their contracts, and almost all of them quickly fell far behind their deadlines.[46] At the beginning of 2006 the Mellon Foundation – after noticing

38 Ramagoma, 'Legal Review – Collaboration Agreement Between the University of Kwazulu Natal and Ithaka Harbors, Inc'.

39 Heather Edwards, 'DISA's Relationship with Aluka', 6 May 2004, DISA Collection, Historical Papers, Wits.

40 Heather Edwards, 'Aluka-DISA Collaboration: Principles', 25 February 2005, DISA Collection, Historical Papers, Wits University.

41 Kevin Guthrie and Tom Nygren, 'Aluka: Building A Digital Library Of Scholarly Resources From Africa', January 2007. http://ts-den.aluka.org/fsi/img/misc/pdf/Background.pdf.

42 Tom Nygren to DISA Committee, 'Selection Process', 20 May 2004, DISA Collection, Historical Papers, Wits University.

43 DISA/Aluka Content Committee, 'Report to the Aluka Meeting, Maputo', 12 March 2006, DISA Collection, Historical Papers, Wits University.

44 Aluka Content Committee, 'Meeting Held at the Campbell Collections, UKZN'.

45 Gail Gerhart, 'Aluka Architecture', 21 November 2004, DISA Collection, Historical Papers, Wits University.

46 DISA/Aluka Content Committee, 'Minutes of Meeting', 5 December 2006, DISA Collection, Historical Papers, Wits University. Sekiba Lekgoathi and Noor Nieftagodien were among the very few to complete their targets on the UDF (7000pp), Township Revolts (3000pp) and Urban Community Struggles (3000pp).

that the grant was not being spent because almost no documents had been identified and licensed for scanning – began to lean on the DISA managers to speed up production.[47]

It was in this atmosphere – as the contradictions between Aluka's expectations of mass collection, the project's rigid subject architecture, the slow pace of the historians' selections and Mellon's growing impatience – that key figures in the DISA project began to articulate loudly a critique of digitisation as neo-imperialism. The protests were triggered by another project between King's College London (KCL), Wits and UWC into the digitisation of struggle papers. The DISA project's managers wrote to Higher Education South Africa (the coordinating body of university vice chancellors) to warn of the dangers of international collaboration. While the letter was aimed at the KCL project (a DISA competitor), their complaints could unmistakably apply to Aluka (an assumption supported by the fact that the Mellon project was not mentioned by name in the letter). South African universities' 'archival resources, comprising one or more local collections, are being digitised and held outside the country, often under unilateral subscription license in conjunction with commercial publishers', they warned. The international partners had no interest in 'long-term sustainability and growth of electronic resources in South African libraries and archives'. And the result was that 'South African universities are at risk of relinquishing their unique holdings, often representing important indigenous knowledge systems, without reciprocal or equitable exchange'.[48] In Thabo Mbeki's country, this was fighting talk. When the presidents of the Mellon Foundation and Ithaka were given sight of the letter they were both furious. Heather Edwards, the Wits librarian, attempted to undo the insult to their only funder, warning of the dire consequences of alienating the Mellon Foundation, but with little success.[49]

Shortly afterwards, in September of 2006, Premesh Lalu – another key figure in the DISA/Aluka collaboration – presented a paper entitled the 'Virtual Stampede for Africa' to the History Seminar at UWC. Drawing on the disappointing work of the selection committees, the paper complained that as Aluka progressed it was growing increasingly oblivious to the debates about the politics of the archive in South Africa.[50] And it warned that digitisation 'will probably perpetuate the unequal relations between the global North and South'.[51] These two presentations expressed the breakdown between Aluka and DISA clearly, but both could still be presented as indirect criticisms. That was not true of the paper that Michele Pickover presented to the National Heritage Council in March 2007.

For some years the NHC had been fussing (with encouragement from DISA) about setting national policy for the digitisation of heritage artefacts. Pickover was invited to address the council about the problems that were being generated by the Aluka project.[52] She did not hold back. Describing the project as being driven by haste, commercialism and an effort to secure intellectual resources in the North, Pickover concluded her presentation by presenting a string of questions. How might better management of digitisation address 'problems of cultural pillaging'? 'Is the temptation of financial aid', she asked, 'producing a new form of

47 Stuart Saunders, 'Disa', 25 May 2006, DISA Collection, Historical Papers, Wits University.

48 Heather Edwards and Dale Peters, 'Brief to HESA', 30 June 2006, DISA Collection, Historical Papers, Wits University.

49 DISA Governing Committee, 'Minutes of the DISA Governing Committee Meeting Held at Wits Club, University of the Witwatersrand, Johannesburg', 5 October 2006, DISA Collection, Historical Papers, Wits University; Edwards and Peters, 'Brief to HESA'.

50 Lalu cited in Carolyn Hamilton *et al.*, (eds), *Refiguring the Archive* (Cape Town, David Philip, 2002) as the arena of this debate.

51 Lalu, 'Virtual Stampede'.

52 DISA Governing Committee, 'Minutes of the DISA Governing Committee Meeting Held at Disa, EG Malherbe Library, UKZN', 5 June 2007, DISA Collection, Historical Papers, Wits University.

imperialism reinforcing the digital divide?'[53] In the Mbeki state this rhetoric of imperialism renewed resonated powerfully, and it neatly served the NHC's own aggrandising concern to police the 'north and south hegemony'.[54]

The presentation was not well received at the Aluka headquarters. The managers there were, to put it mildly, outraged. They rejected the accusation of neo-colonial dominance, insisting that DISA was responsible for its own activities and that the transfer of all the South African content to the US was an open access policy enforced by Mellon. And they expressed deep offence at the 'inflammatory language' that was used to describe their project which, they claimed, was 'aimed at damaging Aluka and excusing DISA from any responsibility for criticisms of the project to date'.[55]

In the year that followed, the DISA/Aluka collaboration ground to an ignominious halt amidst increasingly bitter disagreements among the South Africans. As ever, the historians proved slow and unreliable. The key archival collections at Wits, UCT, Fort Hare and UWC refused to provide the kind of licensing that would allow their most obvious holdings to be used.[56] Adopting the strategies widely used at this time in Durban, the staff at UKZN attempted to impose a code of conduct on the other members of the DISA Governing Committee that would prevent any further embarrassment.[57] Mellon, impatient at the lack of progress – which left almost the entire grant of US\$1.15 million unspent after four years – allowed DISA to retain about half of the funding for a final year of work. In that last year DISA hastily reverted to the scanning of long-running serials without copyright – thus it was that the very useful digital holdings of the Annual Reports of the South African Institute of Race Relations, the *South African Labour Bulletin* and *African Communist* all found their way into http://www.aluka.org (but not, even when the server is available, to http://www.disa.ukzn.ac.za).[58] As the Mellon funding cycle ended in 2009, the DISA offices were closed at UKZN, and the server hosting the materials was left running with only unpaid support. In the early days of the project the multi-institutional arrangement seemed to be a strength, but as the funds dried up the project was left without an institutional champion of last resort (or perhaps, more accurately, with a predictably unreliable champion). The archive's current status at UKZN – not to speak of its future – is in real jeopardy.

The transfer of the Aluka collection in 2008 in to the very secure and capable hands of the JSTOR consortium eliminated the concerns that many have for the DISA records. The risks faced by the valuable DISA records sitting on an unmanaged, aging server in the UKZN server room compared with those faced by Aluka in the custody of JSTOR – with its 7,000 fee-paying institutional members – capture the full range of possibilities allowed by the idea that digital preservationists call future-proofing.

But JSTOR's adoption of the Aluka materials also stoked the old concern about access and the theft of intellectual property. JSTOR is an unparalleled online resource and, not least because of this story and Aluka's efforts, it is widely (and, much more importantly, freely) available to Internet users on the networks of not-for-profit institutions on the African continent. But it is not available to the general public (see Figure 1). This arrangement – despite its undeniably generous and humane intentions – has the unfortunate (and often

53 Pickover, 'Notes Re Challenges of the Aluka Projects'.
54 Sonwabile Mancotywa, Chair of the NHC, cited in M & R Badiri Consultants, 'Report on the Workshop on the Protection and Promotion of Digital Heritage Resource: Towards a National Startegy [sic]'.
55 Tom Nygren to DISA Executive Committee, 8 May 2007.
56 Heather Edwards, 'Report to the Andrew W. Mellon Foundation on the DISA 2 (Phase 2) Project for the Year 2007', 24 March 2008, DISA Collection, Historical Papers, Wits University.
57 Saunders to Heather Edwards, Chris Saunders, and Michele Pickover, 'Disa'; DISA Governing Committee, 'Minutes of the DISA Governing Committee Meeting Held at Disa, EG Malherbe Library, UKZN'; DISA, 'Draft Code of Conduct', 2007, DISA Collection, Historical Papers, Wits University.
58 Edwards, 'Report to the Andrew W. Mellon Foundation on the DISA 2 (Phase 2) Project for the Year 2007'.
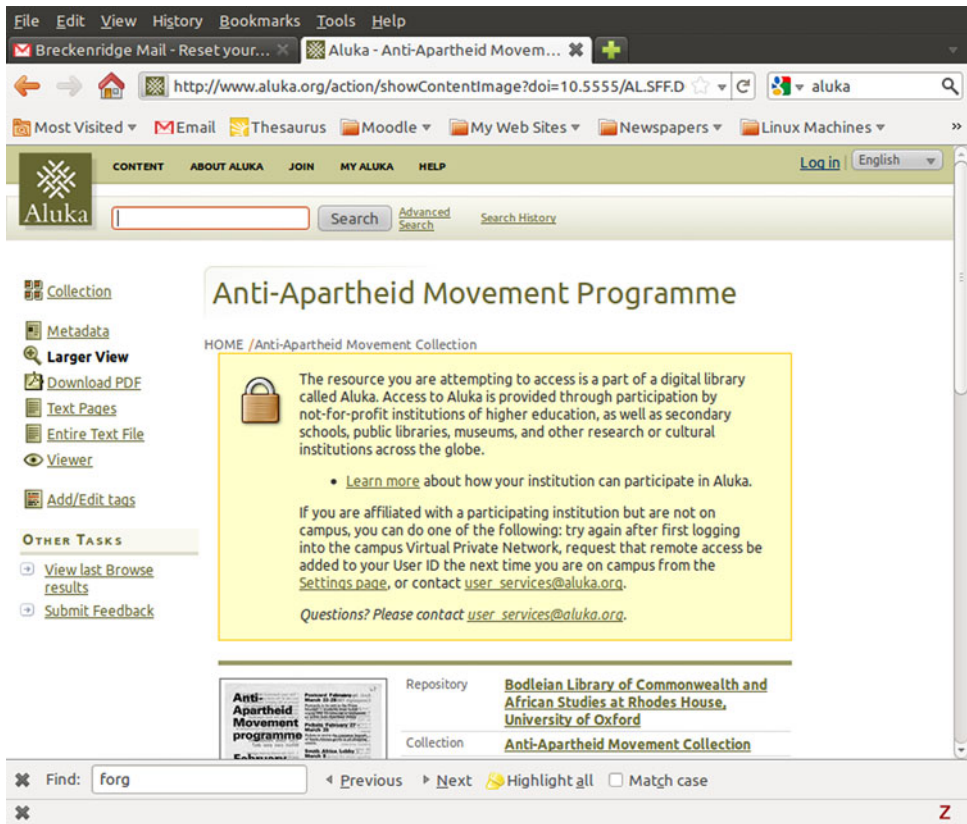
Figure 1. Access Denied to Aluka.

predicted) result that most people in southern Africa are denied access to the digital archive of their own struggle history. It is this resource paradox, perhaps more than any other issue, that has given substance to the charges of heritage theft that have long bedevilled digitisation projects in this region.

The very different prospects of the DISA and the Aluka arrangements seem to provide tangible proof of the arguments of neo-colonialist politics inherent to digitisation that have been expressed in South Africa for a decade. In the years since the Aluka records moved into JSTOR's walled garden, these worries about the politics of digitisation projects emerging from the wealthy North have taken on the form of established convictions. Yet, in the space that remains I want to suggest that this is an incorrect and especially an unhelpful view of the South African interest in digitisation. Considered globally, the Internet provides researchers in South Africa with access to digital resources that vastly outweigh the value of the holdings they currently present online. Much of this content is driven by the effort in the North to escape preservation costs by scanning nineteenth- and early twentieth-century volumes, but it functions, in effect, as a massive gift economy with dozens of libraries making available very valuable resources to an unrestricted audience. South Africans have every reason to join this economy, and to attempt to shape its value by scanning important sources of long-running series that are stored locally – examples might include the South African Hansard or the *Rand Daily Mail*. Freed of the cherry-picking objectives that bedevilled the Aluka project,

South African archives have every reason to return to the kinds of industrial digitisation projects that motivated the original success of the DISA project.

The Work of Public Archives on the Internet

For many years historians in South Africa have been using JSTOR, the journal repository funded initially by the Mellon Foundation in 1995 that is very widely distributed across the world and supported by some 7,000 institutions in 158 countries.[59] JSTOR demonstrates some of the compelling advantages of rapid single-entry-point full-text search, and it makes available to many scholars a range and depth of academic sources that was unimaginable as recently as 1995. The scope of this project is genuinely remarkable, covering the entire publication runs of 1,400 journals in 50 disciplines.[60] It is also, as noted earlier, a contradictory case of open and closed access – providing free membership to not-for-profit institutions on the African continent but, otherwise, requiring users to be registered at one of the fee-paying libraries (including, importantly, many public and municipal institutions).[61] JSTOR has also recently been attacked by one of the outstanding international advocates of open online publishing, and the controversy highlights several interesting and important features of the technopolitics of digital archival repositories.

In October 2010, Aaron Swartz, a 24-year-old with a singular reputation as an adroit open access hacker and militant, used the high-speed network at Massachusetts Institute of Technology (MIT) to download five million articles from the JSTOR repository. Swartz, who was a visiting fellow at Harvard's Safra Center for Ethics, was quite open about the goal he had in mind in extracting almost the entire contents of the JSTOR repository. He was enraged by JSTOR's assertion of copyright over the scholarship published in the online journals. In 2008 he announced a programme to oppose 'the private theft of public culture' by publishing copyrighted content on the world's file-sharing networks.[62] For JSTOR, and for Ithaka, the Swartz attack was clearly an awkward moment, echoing many of the political charges that were levelled against Aluka. It was difficult to oppose Swartz's central claim that knowledge should be free and available, especially to the poor, those who need it most. Yet, at the same time, without some mechanism to protect the property rights of the collaborating academic publishers and to support the ongoing heavy costs of digital preservation, the JSTOR project was clearly doomed. When the US Department of Justice indicated its intention to prosecute Swartz for computer fraud, JSTOR announced that it was 'fully cooperating'.[63]

The reaction of the open access movement was instructive, highlighting the technical difficulties that copyright owners and states face when confronted with digital archives. Two days after Swartz's indictment was announced, another activist, Gregory Maxwell, posted a very large archive of the JSTOR material on to the torrent-sharing website, The Piratebay. That file contained the entire out-of-copyright publication run of the oldest academic journal in the world, the *Philosophical Transactions of the Royal Society*, dating from 1665. Maxwell explained himself at some length, concluding thus:

59　Schonfeld, 'JSTOR'.
60　JSTOR, 'JSTOR at a Glance', Academic Publishing, JSTOR, 13 February 2012. http://about.jstor.org/sites/default/files/misc/jstor-factsheet-20120213.pdf.
61　JSTOR, 'African Access Initiative', Academic, JSTOR. Retrieved 25 April 2012 from http://about.jstor.org/participate-jstor/libraries/african-access-initiative-0.
62　Aaron Swartz, 'Guerilla Open Access Manifesto'. Retrieved 19 November 2012 from http://archive.org/details/GuerillaOpenAccessManifesto.
63　JSTOR, 'JSTOR Statement: Misuse Incident and Criminal Case', 19 July 2011. Retrieved 19 November 2012 from http://about.jstor.org/news/jstor-statement-misuse-incident-and-criminal-case.

> The portion of the collection included in this archive, ones published prior to 1923 and therefore obviously in the public domain, total some 18,592 papers and 33 gigabytes of data. The documents are part of the shared heritage of all mankind, and are rightfully in the public domain, but they are not available freely. Instead the articles are available at $19 each – for one month's viewing, by one person, on one computer. It's a steal. From you.[64]

At the time of writing, the torrent was being hosted on 50 computers around the world. When Aaron Swartz committed suicide in January 2013, his death was presented by many activists as a laudable example of the sacrifice demanded by the open access movement.[65]

The Internet, despite its origins in the US Defense Advanced Research Projects Agency, acts for the moment at least as a stateless archive, and it is home to thousands of file-sharing systems and repositories. Some of these repositories, like http://libgen.org, are clearly illegal, hosted in countries that are difficult to police, by clever and resourceful students who understand that a digital text (and the databases needed to host and find them) can be reproduced and distributed instantly, without meaningful costs, and that the enormous use value of instant access will trump the threats of copyright fines. Hundreds of thousands of commercially published e-books are also being shared by torrent users. In both cases, what makes these archives function efficiently is their highly dispersed structure and intensive duplication. These are also the features that make them very difficult to police.

The publication in October 2010 of some 250,000 classified US State Department briefing cables on http://www.wikileaks.org by the *Guardian*, and the *New York Times* announced a fact that had been true for some time[66]: the project of assembling some of the very largest, and politically most sensitive, archives had escaped from the control of states. The Wikileaks story has still to run its course, and the questions of whether Julian Assange will remain out of prison and the Wikileaks web repository will be freely (and legally) available on the Internet have yet to be answered. But – despite the appointment of a US Federal Grand Jury, attacks on the website's domain name and banking assets – the technical and political obstacles to the control of stateless online publishing have proven remarkably difficult for the US government to overcome.[67] The ongoing publicity about a forthcoming future cyberwar is evidently part of an effort to strengthen the state's capacity to act online. The masses of evidence released recently by Edward Snowden demonstrate several points very powerfully: the technopolitics of massive digital replication, the difficulty of policing content on the Internet and the enormous investments that states are making to intervene online. The US state has clearly moved faster to build up its capacity in this area than all but the most committed conspiracy theorists had suggested. But the individuals behind Wikileaks, and the broader project of online anonymity, open publishing and state and corporate transparency, also have well worked out ideas and capacity. They are technically adept at the use of the

---

64 Greg Maxwell, 'Papers from Philosophical Transactions of the Royal Society', Torrent Sharing, The Piratebay, 21 July 2011. Retrieved http://thepiratebay.se/torrent/6554331/Papers_from_Philosophical_Transactions_of_the_Royal_Society__fro.

65 Noam Cohen, 'Aaron Swartz, a Data Crusader and Now, a Cause', *The New York Times*, 13 January 2013, 'Technology' section. Retrieved from http://www.nytimes.com/2013/01/14/technology/aaron-swartz-a-data-crusader-and-now-a-cause.html.

66 Scott Shane and Andrew W. Lehren, 'WikiLeaks Archive — Cables Uncloak U.S. Diplomacy', *The New York Times*, 28 November 2010, 'World' section. Retrieved from http://www.nytimes.com/2010/11/29/world/29cables.html; Simon Rogers, 'WikiLeaks Embassy Cables: Download the Key Data and See How It Breaks down', *The Guardian*, 28 November 2010, 'Media' section. Retrieved from http://www.guardian.co.uk/news/datablog/2010/nov/29/wikileaks-cables-data#; Perry Link, 'Waiting for WikiLeaks: Beijing's Seven Secrets', *NYR Blog*, 19 August 2010. Retrieved from http://www.nybooks.com/blogs/nyrblog/2010/aug/19/waiting-wikileaks-beijings-seven-secrets/; Simon Shuster, 'WikiLeaks: Is Russia the Next Target?', *Time*, 1 November 2010. Retrieved from http://www.time.com/time/world/article/0,8599,2028283,00.html.

67 Dominic Rushe, 'WikiLeaks Cements Its Role as Thorn in the Side of US and UK Governments', *Guardian*, 23 June 2013. Retrieved from http://www.theguardian.com/media/2013/jun/23/wikileaks-snowden-governments-thorn-in-side.

tools of concealment, encryption, peer-to-peer content distribution and hosting required to maintain a large and controversial repository of documents. The result is still uncertain.

There are also perfectly legal archival repositories whose inventories are even more valuable than the vast holdings of the illegal e-book libraries or Wikileaks. The best single repository is Brewster Kahle's Internet Archive (http://www.archive.org), which, in addition to preserving old web content and old physical books, hosts scanned copies of the out-of-copyright works of dozens of the largest libraries in the world. At the time of writing, the Internet Archive was processing 20,000 books per week.[68] These books – most of what was written in the English language before 1930 – are freely available as full-text searchable records on an online database and, as many historians know, they are changing the way that researchers investigate nineteenth-century problems, making available for the first time a host of related sources that can be searched systematically. These files are hosted in special shipping containers on the campus of Sun Microsystems in Santa Clara, California, but many of them are also being published using torrent files that distribute both the data and its hosting across the Internet.[69]

Nor is this is simply a matter of the changing repositories that researchers can use to find documents on the Internet; these new repositories have also begun to alter the way that we participate in the publication and distribution of these materials. Zotero, the research managing software produced by the Roy Rosenzweig Centre for History and New Media at George Mason University, provides researchers with a tool for submitting primary documents directly to the servers of the Internet Archive. The Zotero Commons offers scholars very advanced text recognition in exchange for making primary documents available publicly (see Figure 2). Publishing the content is a simple matter of dragging a document record onto an icon. In the spirit of the open access movement, the software presumes that the problems of copyright, classification and, most importantly, authenticity are simply irrelevant, or someone else's business. With a few mouse movements users can transfer primary materials bearing their own meta-data to the Internet Archives in perpetuity.

The Parallel Archive (http://www.parallelarchive.org), based in Budapest and supported by the Soros Foundation and the Open Society, is an example of a self-consciously stateless archival repository designed to address problems that may seem familiar to South Africans: the first of these were those that dealt with the crimes committed by existing states and their officials. The archive emerged from the involvement of the Open Society Archives in the documentation of human rights violations in the Balkans and, in particular, with the involvement of non-governmental actors in the gathering of evidence.[70] But very quickly the archivists became aware of problems that are not, yet, being much discussed here. Working with documentation that was often born electronically raised problems detecting and correcting forms of manipulation and forgery. 'The fragility of authenticity is a dramatic threat for the archives', they write, 'especially but not exclusively, in the case of legal and forensic documents'.[71] To address this threat the Parallel Archive offers researchers a suite of forensic tools: text recognition, tagging, sharing and annotation facilities, and, in the effort

---

68 David Streitfeld, 'Internet Archive's Repository Collects Thousands of Books', *The New York Times*, 'Technology' section, 3 March 2012. Retrieved from http://www.nytimes.com/2012/03/04/technology/internet-archives-repository-collects-thousands-of-books.html.

69 Lucas Mearian, 'The Internet Archive's Wayback Machine Gets a New Data Center', *Computerworld*, 25 March 2009. Retrieved from http://www.computerworld.com/s/article/9130499/The_Internet_Archive_s_Wayback_Machine_gets_a_new_data_center; Lucas Mearian, 'Internet Archive Unleashes 1PB of Data through BitTorrent', *Computerworld*, 8 August 2012. Retrieved from http://www.computerworld.com/s/article/9230098/Internet_Archive_unleashes_1PB_of_data_through_BitTorrent.

70 'What is OSA?', web page. Retrieved 20 November 2012 from http://www.osaarchivum.org/index.php?option=com_content&view=article&id=48&Itemid=133&lang=en.

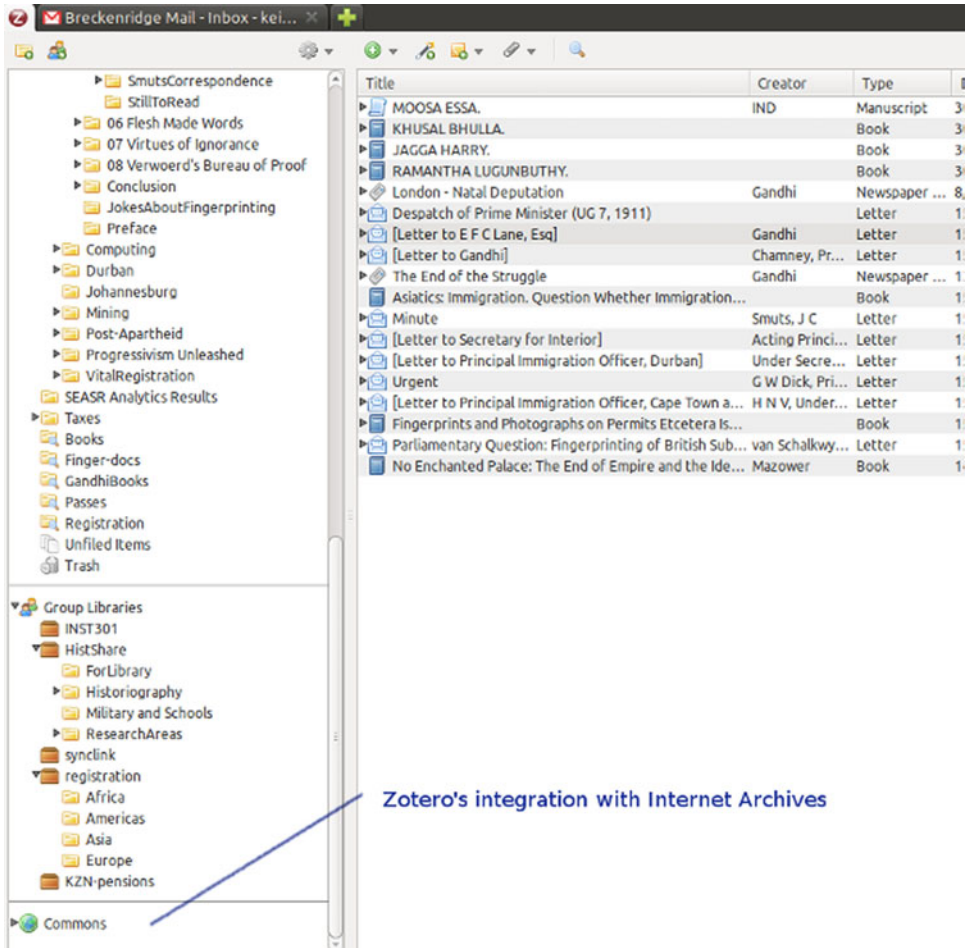71 'Parallel Archive: Vision'. Retrieved 24 June 2011 from http://www.parallelarchive.org/content/vision.

Figure 2. Integration of the Internet Archives and the Zotero Commons.

stabilise and authenticate contentious sources, a permanent and cryptographically certified link. To date the project appears to be something less than an obvious success, but it raises questions about digital authenticity that are important now and likely only to be more so in the future.

Figure 3. @AchilleMbembe impersonation on the Twitter website.

Digital forgery is already common in South Africa and beyond. My colleague Achille Mbembe has an annoying impersonator on Twitter, who uses an automating tool to post arbitrary sentences out of his published works (see Figure 3). The impersonator – who evidently imagines this is an act of homage – had nearly 1,000 followers when Mbembe tried to get Twitter to act against him. After some correspondence with the Twitter administrators involving a brief contest over who was the real Mbembe, the impersonating account was briefly shut down, only to reopen with a new account called MbembeQuotes (perhaps signalling its relationship with the author a little more accurately).

Dozens of similar impersonations of South African political figures are currently at work on Twitter and Facebook. But there are also more serious and more interesting works of web-based forgery.

The Arpanet Dialogues provides a useful example of the potential for these kinds of forged archives. The Arpanet, as many people know, was the original inter-network funded by the US Department of Defense.[72] It provided university-based researchers with access to email and content from the Usenet discussion group archives from the late 1960s. The conceit behind the Dialogues is that conversations that took place on the Arpanet – involving famous individuals who would otherwise never have encountered each other – have suddenly been unearthed and made available on the web. The conversation that is most interesting for South Africans is Volume 2, which took place in June 1976 and involved Samir Amin, Francis Fukuyama, Minoru Yamasaki and Steve Biko (see Figure 4).

These dialogues have been very carefully produced. Key individuals who were deeply familiar with the views of each of the protagonists participated in an online conversation, which was then recorded (as it would have been in 1976) and presented in a plain text format.[73] They are deeply intriguing and persuasive, and announce very loudly that, as the cliché reminds us, anything is possible on the Internet.

Understanding Archival Registration and Meta-Data

Many historians are likely to react to these examples of fertile and feckless invention by retreating to the safe ground of the physical document. Over the last decade many scholars have pointed to the virtues of non-textual meta-data – qualities of paper documents that cannot easily be captured in writing, but which establish their biographical bona fides. This is, in fact, a very old idea. 'The presence of the original is the prerequisite to the concept of authenticity', as Walter Benjamin has reminded us of the fine-grained details required for the authentication of art works prior to the period of mechanical reproduction.[74] The most well-known example is probably the account by Brown and Duguid of the epidemiological eloquence of the smell of vinegar in eighteenth-century letters.[75] But almost all archival researchers have their own sense of the reliability of documents, and bindings, and the methods of those who read them. In this sense there is some truth to the stupid cliché that historians place naïve trust in the document in the archive.

---

72 The story was famously presented in Manuel Castells, *The Rise of the Network Society: The Information Age – Economy, Society, and Culture* (Oxford, Blackwell, 1996); and then qualified by Janet Abbate, *Inventing the Internet: Inside Technology* (Cambridge, MA, MIT Press, 1999).

73 Richard Fisher, 'CultureLab: Ronald Reagan Has Joined the Chatroom', *New Scientist*, 7 April 2011. Retrieved 12 May 2012 from http://www.newscientist.com/blogs/culturelab/2011/04/so-reagan-signs-into-this-chatroom.html.

74 Benjamin, 'The Work of Art in the Age of Mechanical Reproduction', p. 274.

75 John Seely Brown and Paul Duguid, *The Social Life of Information* (Cambridge, MA, Harvard University Press, 2000), pp. 173–4.

Figure 4. http://www.arpanetdialogues.net/vol-ii/.

But this confidence in the authenticity of the non-textual qualities of physical paper is utterly misplaced. The problem of unconstrained forgery is almost as old as the paper archive itself. In his brilliant history of the rise of the documentary state in England in the late Middle Ages, Clanchy shows that forgery and archive building were mutually constituted. In fact, the evidence suggests that the turn to writing coincided with a much more dramatic and extensive falsification of claims about property and power than may actually be possible in the densely relational world of the digital medium. 'Forgers recreated the past in an acceptable literate form', Clanchy observed; 'they are best understood not as occasional deviants on the peripheries of legal practice, but as experts entrenched at the centre of literary and intellectual culture in the twelfth century'. And it was, importantly, the Christian monks themselves – custodians of the archives and 'traditional experts in writing, who were the greatest forgers'.[76] This brings me to my final point, for the remedies that medieval canon law developed to address the problem of forgery point to an important new role for official archives in the world of the digital record.

Lawyers in the twelfth century developed a set of tests of authenticity that remain interesting and useful today. They argued (as anyone who manages email will confirm) that

---

76 M.T. Clanchy, *From Memory to Written Record, England 1066–1307* (London, Edward Arnold, 1979), pp. 129, 234, 249.

the 'style and substance' of the prose should be examined first, and then that the 'physical qualities of the parchment' and its bindings should be tested. But they also insisted on the development of a set of documentary registers that would list the key records by virtue of their meta-data and claims and create a relational mechanism to test and confirm the reliability of records in an environment of almost open-ended forgery and manipulation.[77] This is clearly what the Parallel Archive has in mind for non-state records in central Europe. It is also something that can be easily achieved in South Africa because of the development by IBM, at the high moment of the apartheid state, of a master-register of all manuscripts.[78] Clearly something like a meta-register of archives (or, more likely, a matrix of linked registers) that would allow historians to share, debate and comment on existing and missing records is required now. Whether the state, or some other institution, builds it remains to be seen.

Conclusion

I hope that I have made the point that linking archival assembly, and especially digitisation, to imperialism is, at best, unhelpful. It draws, wrongly, on the characterisation of the place of record-keeping in the African empire, which was, in the first place, an instrument for the economic and ideological regulation of imperial officials. Importantly the preoccupation with heritage theft systematically ignores the extraordinary cultural and intellectual riches which are now being made available to South Africans from the largest web-based archives. A second point is that the debate to date confuses much about the place of the archive under apartheid. The current neglect of record-keeping, and, especially, the contemporary failures of administrative accountability, clearly began decades ago. In the effort to remedy the damage done during the last decades of apartheid we should keep in mind that official secrecy and a massive project of official censorship were the key tools of bureaucratic despotism. A properly working national archives service, and careful parliamentary oversight, must be part of any effort to create a more responsive and efficient democratic civil service. And web-based archival assembly can assist with this project.

The Internet already supports thousands of new sites of documentary disclosure, creating mechanisms for official and commercial transparency that were unimaginable a little more than a decade ago. There is certainly a struggle under way between the bureaucracies of the most powerful nation-states and the advocates of stateless publishing. But the outcome is anything but obvious. The current efforts on the part of governments and copyright holders to police these repositories, or to enforce a currently non-existing copyright in official materials, will face formidable obstacles mostly prompted by the technologies themselves. It is, in short, much too easy to duplicate, distribute and publish very large quantities of data publicly or secretly for policing and copyright to apply very easily.

But these same qualities should encourage historians (and their critics) to think very carefully about the problems of digital authenticity, and the tools we currently have for assessing it. This is partly a matter of being more explicit about the scope of archival work, of the gaps and coverage in our research. But it also requires more carefully worked out tools for sharing and evaluating sources and a commitment to intensive projects of intertextual analysis. New software and search tools can and will help with this. The National Archives service and the NHC are especially well placed to do this work, but that will require a fairly

---

77 *Ibid.*, pp. 254–5.
78 Editor, 'At Long Last a National Register?', *South African Archives Journal* (1970), p. 8; D.J.J. Smith, 'National Register of Manuscripts', *South African Archives Journal* (1978), pp. 27–9.

dramatic change in their attitude to digitisation and our documentary heritage. It also clear that many of the changes will happen whether the state fosters them or not.

What then are the effects of digital records on the work of the public archives? The first among many effects is, I think, on the way we conceive of the relationship between the archives and the nation-state. Nicholas Dirks puts the conventional argument in starkly Hegelian terms: 'The archive is the instantiation of the state's interest in history'.[79] And perhaps from the moment that Hegel first noticed the rising intellectual authority of the state, that claim has been true. National and colonial archives unmistakably work to support a monopoly over the narrative of the past that placed the role of the imperial or local state at the centre of historical writing. That monopoly is broken.

KEITH BRECKENRIDGE

Wits Institute for Social and Economic Research (WISER), University of the Witwatersrand, Postbag 3, P.O. Wits, Johannesburg, South Africa, 2050. E-mail: keith@breckenridge.org.za

---

79 Dirks, 'Annals of the Archive: Ethnographic Notes on the Sources of History', p. 63.