PROJECT MUSE®

Crime, Penal Transportation, and Digital Methodologies

Barry Godfrey, Caroline Homer, Kris Inwood, Hamish Maxwell-Stewart,
Rebecca Reed, Richard Tuffin

Journal of World History, Volume 32, Number 2, June 2021, pp. 241-260
(Article)

Published by University of Hawai'i Press

DOI: <https://doi.org/10.1353/jwh.2021.0023>

➔ *For additional information about this article*

<https://muse.jhu.edu/article/794329>

Crime, Penal Transportation, and Digital Methodologies

BARRY GODFREY, CAROLINE HOMER,
KRIS INWOOD, HAMISH MAXWELL-STEWART,
REBECCA REED, AND RICHARD TUFFIN

This article argues that the ability to systematically analyze hundreds of thousands of life course events provides an opportunity to explore the ways in which an Australian convict archive was originally intended to be used, as well as a means of placing information supplied by subalterns within context. We also show how the digital reconstruction of the bureaucratic instruments of colonial labor management can be used to shed light on state actions. Using a combination of longitudinal and cross-sectional techniques, we place the experience of transported men and women within the colonial context of evolving labor markets, policing, and criminal justice systems, exploring questions of colonial class formation, gender, and labor mobility in the process. We end by pointing to how such datasets might be used in future undergraduate teaching and digitization initiatives.

KEYWORDS: digital history, crime history, convict transportation, life course history, spatial analysis, archival ecologies.

MASS digitization of historical sources, and the ease with which these can be mined, has enriched the research process while simultaneously creating new pitfalls and problems. This is a particular issue for researchers working on the history of colonialism. Digitization may have lowered the costs of conducting transnational research, enabling historians to conveniently access information derived from geographically dispersed archives, but these processes also encourage the indiscriminate targeting of data. The aggregation of disparate

sources into a unified digital “product” has been seen as particularly problematic.¹

In this article, we start by exploring the knotty problem of data sovereignty before moving on to describe ways in which digital methodologies can be used to create improved archival interfaces.² We argue that the ability to systematically analyze and map multiple life course events in parallel provides an opportunity to explore colonial archives “along the grain.” We outline a series of longitudinal, cross-sectional, and spatial techniques that can be used to place the experience of British and Irish transported convict men and women sent to the Australian penal colony of Van Diemen’s Land within the context of evolving labor markets, policing, and criminal justice systems. In the process, we show how the digital reconstruction of the bureaucratic instruments of colonial labor management can be used to shed light on state actions. Finally, we conclude by pointing out that many of the problems alluded to in the literature predate digitization. While access to increasing quantities of digitized data has the capacity to exacerbate these preexisting issues, it also offers the opportunities for the discipline to shore up its foundations, including the way history is taught at an undergraduate level.

THE CHALLENGE OF DIGITIZING THE COLONIAL ARCHIVE

The rate at which digital data has disappeared behind the pay walls of commercial providers is a particular concern in two ways; the first is access. Faculty and students, at least those from well-funded universities, may expect their subscriptions to commercial sites to be paid by their institutional libraries, but this is not universal. The cost of accessing digital data will impact most on nonprofessional historians and staff and students from impoverished nations and institutions. Second, ownership of the means of digital discovery has important further ramifications, precisely because retrieval systems mediate the ability of users to locate and interact with data. Search engines that enable the retrieval of information devoid of context fail to render data

¹ Lara Putman, “The Transnational and the Text-Searchable: Digitized Sources and the Shadows They Cast,” *American Historical Review* 121, no. 2 (2016): 377–402 and A. Gailey, “Some Big Problems with Big Data,” *American Periodicals: A Journal of History and Criticism* 26, no. 1 (2016).

² Timothy Hitchcock, “Confronting the Digital or How Academic History Writing Lost the Plot,” *Cultural and Social History* 10, no. 1 (2013): 14–18.

discoverable in the true sense of the term. Keyword searches, for example, are of little help in assisting users to explore the extent to which different populations might be under- or over-represented in an archive, let alone the factors that shape the decision to document target populations or describe them in particular ways. While databases exist as a back end to search windows, academic researchers typically are denied the opportunity to design their own queries or run their own programs. This is not just a private sector issue. Public sector holders of birth, marriage, and death registration data (and in some countries census data) are available only for one-off searches, often on a pay as you go basis. A further concern is that the prejudices of the past will be inserted into the finding aids of the present as the result of the unproblematic populating of search engines.³ These issues deserve the attention of historians. As Andrea Davis eloquently argues, historical skills are needed to both represent the complexity of source materials and to ensure that data is ethically organized and “openly formatted to promote data reuse and sharing.”⁴

The risks associated with digitization are particularly pertinent to colonial research. In part, this is because the bureaucratic infrastructure that facilitated the alienation, colonization, and administration of overseas territories by the Western empires are often located in dispersed archival collections. It is not unusual for two halves of a correspondence series to be archived in different parts of the world, or for multiple drafts of a record to be housed in different collections. While the digitization of content has dramatically reduced the costs of accessing dispersed collections, this has come at the price of increased risk of selection bias. That the contents of those record collections were highly selected in the first place only adds to the potential danger. No archive is neutral. Each was formed with a certain bureaucratic purpose in mind. If knowledge is power, then this applies with particular force to the archives generated and curated by colonial bureaucracies.⁵

Yet, while we agree that digitization can encourage “disintermediated discovery,” we argue that imaging, transcribing, coding, and linking can enable researchers to place data within a wider frame of reference.⁶ The systematic analysis of purposely assembled digital

³ Megan Garcia, “Racist in the Machine: The Disturbing Implications of Algorithmic Bias,” *World Policy Journal* 33, no. 4 (2016): 111–117.

⁴ Andrea Davis, “Introduction,” *Bulletin for Spanish and Portuguese Historical Studies* 43, no. 1 (2018): 2.

⁵ Ann Laura Stoler, *Along the Archival Grain: Epistemic, Anxieties and Colonial Common Sense* (Princeton: Princeton University Press, 2009), 23–54.

⁶ Putman, “The Transnational and the Text-Searchable,” 377.

research collections presents particular opportunities for colonial historians, although this mode of working has implications for the way historical research is conducted.⁷ Digitization is often associated with an increase in the pace of the research process. We argue, however, that to take full advantage of digital technologies it is necessary for historians to slow down, think broadly, and act collaboratively.

The barriers to this type of digital engagement are often social, intellectual, and institutional, rather than technological.⁸ In a memorable phrase, Roy Rosenzweig and Dan Cohen argue that “historians planning a digital project should think like architects not plumbers.”⁹ Rather than designing a research program to take advantage of the capabilities of particular digital tools, some of the best known digital collaborations have evolved from a collective desire to solve long-running historical issues. The Trans-Atlantic Slave Database provides a good example.

This digital collaboration gave historians of slavery the means to pool several existing resources to create a comprehensive transnational database. A single shared data source provided the opportunity to input new data in systematic ways, check for duplicates, and create the necessary standards to enable researchers to interrogate observations gleaned from multiple records. Consequently, the collaborators were able to produce reliable estimates for both the volume of the Atlantic slave trade, and slave flows and mortality rates over time. In addition to its collective nature, the project differed from previous historical endeavors in that this dataset has become an important evolving output in its own right supporting a new generation of research questions beyond those originally envisioned.¹⁰

What is particularly enticing about this kind of work is the way in which it can lead to the development of “scholarly ecosystems.”¹¹ Such systems are not built overnight. Digital initiatives require considerable commitments in time and resources. Researchers who work on their own are poorly equipped to meet the challenge of creating, analyzing,

⁷ Hamish Maxwell-Stewart, “Big Data and Australian History,” *Australian Historical Studies* 47, no. 3 (2016): 361.

⁸ Kathleen Fitzpatrick, *Planned Obsolescence: Publishing, Technology, and the Future of the Academy* (New York: NYU Press, 2011), 18.

⁹ Dan Cohen and Roy Rosenzweig, *Digital History: A Guide to Gathering, Preserving, and Presenting the Past on the Web* (Philadelphia, 2006), 56, <http://chnm.gmu.edu/digitalhistory/>.

¹⁰ David Eltis and David Richardson, eds., *Extending the Frontiers: Essays on the New Transatlantic Slave Trade Database* (New Haven, CT: Yale University Press, 2008).

¹¹ Andrea Davis, “Introduction,” *Bulletin for Spanish and Portuguese Historical Studies* 43, no. 1 (2018): 5.

and curating digital data.¹² This is not just a matter of many hands making light work, it is also a skills issue. If they wish to take advantage of digital technologies, historians need to reach across the corridors and collaborate with those located in other disciplines. While collaboration is required to get the best out of the digitization of the archive, networking has never been easier thanks to the ways digital technologies have lowered the transaction costs associated with coordinating geographically dispersed teams.¹³

DIGITIZING A COLONIAL PENAL ARCHIVE

In our case, collaboration was prompted by the realization that we shared a set of broadly similar research goals. Collectively, we have an interest in exploring the impact of forced migration and unfree labor on life course outcomes.¹⁴ The ways in which the supply of available labor and management practices shaped colonial environments formed an important parallel research agenda.¹⁵ Finally, we are motivated by a desire to better understand the way the past can influence the lives of subsequent generations. These agendas are marked by significant degrees of complexity. There are multiple pathways, for example, that might explain the transmission of social inequality or prosecution risk from one generation to another.¹⁶ It was clear from the start that none of our key research questions could be studied without assembling large datasets. Thus, the decision to pool resources was not difficult, since it was clear that this was a prerequisite if we were ever going to untangle the legacies of colonialism.

¹² Richard H. Steckel, "Big Social Science History," *Social Science History* 31, no. 1 (2007): 13.

¹³ Daniel J. Cohen, Michael Frisch, Patrick Gallagher, Steven Mintz, Kirsten Sword, Amy Murrell Taylor, William G. Thomas, III, and William J. Turkel, "The Promise of Digital History," *The Journal of American History* 95, no. 2 (2008): 454–455.

¹⁴ James Bradley et al., "Research Note: The Founders and Survivors Project," *The History of the Family* 15, no. 4 (2010): 467–477 and Barry Godfrey, David Cox, and Stephen Farrall, *Criminal Lives: Family Life, Employment and Offending* (Oxford, Clarendon, 2007).

¹⁵ Richard Tuffin, Martin Gibbs, Hamish Maxwell-Stewart, David Roberts, David Roe, Jody Steele, Susan Hood, and Barry Godfrey, "Landscapes of Production and Punishment: Convict Labour in the Australian Context," *Journal of Social Archaeology* 18, no. 1 (2018): 50–76.

¹⁶ Barry Godfrey, Kris Inwood, and Hamish Maxwell-Stewart, "Exploring the Life Course and Intergenerational Impact of Convict Transportation," in *Intergenerational Continuity of Criminal or Antisocial Behaviour*, ed. C. Bijleveld and S. Weijer (London: Routledge, 2018), 61–75.

A realization that we could not achieve these aims without working closely with data custodians prompted a final set of motivations. To build and maintain a digital archive, it was necessary to achieve a better understanding of how the sources of information from which that virtual archive was derived were structured. We quickly became interested in the extent to which the information we had access to had been impacted by record loss, as well as the motivations that had driven the colonial and metropolitan bureaucrats, who made the decisions to capture particular pieces of information in the first place. How digital tools could be used to explore what we might term “the archaeology of the archive” rapidly emerged as a significant research question in its own right.

The principal sources of information that we have used to explore these questions are the records created to manage 13,500 female and 59,000 male convicts dispatched in the years 1803 to 1853 to the British penal colony of Van Diemen’s Land (later renamed Tasmania). Criminals were among the first citizens of the Western empires to be systematically described. By the early nineteenth century, those who broke the law were routinely interrogated, measured, and classified. Details collected for prisoners convicted in Britain and Ireland included information about next of kin, place of birth, literacy, religion, occupation, height, and other distinguishing marks such as tattoos.¹⁷ Much of this data was rerecorded on arrival in Australia, providing researchers with the opportunity to compare details collected at different stages in the sentencing process.¹⁸ Other records were created on the voyage to the Australian colonies, including information about the health and conduct of the prisoners.¹⁹ Postarrival summaries of all subsequent offending were entered into conduct registers, while information about convicts who died under sentence or were granted indulgences, such as permission to marry, was filed in separate registers. Details of pardons and dates of release, appointments to the colony’s police force (largely staffed by serving convicts), and prisoners who had deserted their place of work were circulated in the *Government Gazette*.

¹⁷ Robert Shoemaker and Richard Ward, “Understanding the Criminal: Record-Keeping, Statistics and the Early History of the Criminology of England,” *The British Journal of Criminology* 57, no. 6 (2017): 1442–1461.

¹⁸ Hamish Maxwell-Stewart, “The State, Convicts and Longitudinal Analysis,” *Australian Historical Studies* 47, no. 3 (2016): 414–429.

¹⁹ Rebecca Kippen and Hamish Maxwell-Stewart, “Sickness and Death on Convict Voyages to Australia,” in *Lives in Transition*, ed. Peter Baskerville and Kris Inwood (Toronto: McGill-Queens University Press, 2015), 43–70.

We linked digitized versions of these records to other administrative series including births, deaths and marriages, bank account data, hospital and prison records, and returns contained in official correspondence files and parliamentary reports. The resultant digitized archive numbers well over a million records drawn from thirty-four record series spanning the years 1803–1924. While the bulk of the original records from which our digitized versions have been derived are held by the Tasmanian Archives, important subcollections reside in other institutions including the Port Arthur Historic Site Management Authority, Queen Victoria Museum and Art Gallery (Launceston), the University of Tasmania Archive, the State Records of New South Wales, the New South Wales State Library, the National Archives (UK), and the National Archives of Ireland, as well as regional British and Irish archives. In the process of linking digitized datasets, we have made considerable headway in reuniting these dispersed collections.

In the process, we have come to appreciate the strengths and weaknesses of digital transcripts. It can be difficult, for example, to represent variations in handwriting in a dataset entry, although these will affect the way an original record is read. As a safeguard, we have linked transcripts to the digital images from which they were derived. We have also avoided “coding at source”—that is, to represent data as a numeric value, rather than text strings. While this has slowed the process of converting digitized images into analyzable form, it has resulted in the capture of content that can support a wide range of research questions. In effect, a single record in any of our datasets consists of many parts. These include: the physical record linked by its archive item number to an electronic image; a digital transcript; a cleaned and standardized version of that transcript that includes a series of identifiers, and spatial references including geo-coordinates; and finally, links to various coding dictionaries that contain numeric values that represent attributes such as skill, industry of employment, transportable offence, and levels of literacy.

There is a technical difference between a record and a document, in that a record has a known context and transactional history.²⁰ We are interested in how the creation of archival networks can transform a loose assembly of documents into a collection of records. In a digital research environment, this is not a static process. Rather than disassociating information from its archival context, each systematic

²⁰ Joshua Sternfeld, “Archival Theory and Digital Historiography: Selection, Search and Metadata as Archival Processes for Assessing Historical Contextualisation,” *The American Archivist* 74 (2011): 544–575.

interrogation of a digital archive can add to a record's transactional history. Thus, cross-interrogation of similar series can assist in reconstructing the sequence of individual record production.

For example, when each convict vessel arrived in Hobart Town, "appropriation lists" were created to facilitate the allocation of convict labor to the private and public sector. These list the age and work skills of each convict and name a settler or government department to which that prisoner might be assigned. Multiple appropriation lists survive for most convict voyages, each containing subtle variations. By comparing the lists drawn up for each arrival, it is possible to use the sequence of deletions and additions to place these records in the order of their original creation. Turning static records into a dynamic series enables researchers to reconstruct the process of convict labor allocation. Each research project has the potential to create new sets of links between record sets, adding to what we know of the context and transactional history of the archive as a whole. As the implications of this have become apparent, the division between archival partner and academic researchers has blurred. The research process has effectively become an exercise in record curation as well historical enquiry.

Engagement with the process of digitization disrupts traditional academic hierarchies. Our research endeavors are highly dependent on volunteered time. The labor and skills of organizations like the Female Convict Research Centre and a growing team of "citizen historians" who have transcribed data have been indispensable. As Robin Wharton points out, these kinds of digital collaborations challenge the "institutionalized hierarchies of labor" typical of many university settings.²¹ The same tools that allow academics to share data remotely provide opportunities for volunteers and other noninstitutional-based collaborators to both participate in the research process and have their endeavors appropriately acknowledged. One way of facilitating this process is to treat each dataset and its associated description files as a publication with a DOI and a set of authors that names all of those that assisted with its production.

Our research techniques rely on linking information from many records to form chains of events that can be used to analyze the impact of experiences like solitary confinement exposure on life expectancy. This methodology is hardly new. Many historians have used multiple

²¹ Robin Wharton, "Bend Until it Breaks: Digital Humanities and Resistance," in *Disrupting the Digital Humanities*, ed Dorothy Kim and Jesse Stommel (Santa Barbara, CA: Punctum Books, 2018), 385.

record groups to piece together the lives of individuals.²² It is unusual in the humanities, however, to use such techniques to study thousands of individual lives. While social science historians regularly link information across large record series to explore change over time, the number and diversity of sources of information employed is usually limited. A common approach, for example, is to follow individuals across successive censuses.²³ Our research strategy combines both techniques.

We are fortunate to have access to several complete runs of data including all registered births, deaths, and marriages in the period 1838–99. Access to complete count administrative data is useful. While record series frequently contain gaps, duplicate entries, and other anomalies, because they cover a broad swathe of the colonial population, they can help to locate any individual record (or subset of records) within the distribution of all observations. Triangulation is a powerful and often underestimated research technique that can be employed as an effective check against selection bias and record omissions. It can also be used to analyze variations in the way information is recorded across record groups. By exploring who is enumerated in different administrative datasets, and comparing differences in the way they are described, it is often possible to reconstruct much about the process of record formation.

Most records are generated by encounters with the state. Individuals commonly enter the archive when they provide answers to questions directed at them by a public intermediary. Such self-reported data are often viewed with suspicion. When asked to provide their height and weight, for example, individuals tend to exaggerate how tall they are and underestimate how much they weigh. To make matters more difficult, the extent of misreporting varies across sex, age, and occupational and ethnic categories.²⁴ Yet, while each record is shaped by the circumstances that characterized the moment of its creation, examination of multiple record series in parallel can reveal much about such processes.

²² See for example Clare Anderson, *Subaltern Lives: Biographies of Colonialism in the Indian Ocean World 1790–1920* (Cambridge: Cambridge University Press, 2012).

²³ Luiza Antonie, Kris Inwood, Daniel Lizotte, and J. Andrew Ross, “Tracking People Over Time in 19th Century Canada for Longitudinal Analysis,” *Machine Learning* 95, no. 1 (2014): 129–146.

²⁴ Alla Chernenko, Huong Meeks, and Ken R. Smith, “Examining Validity of Body Mass Index Calculated Using Height and Weight Data from the US Driver License,” *BMC Public Health* 19, no. 100 (2019), <https://doi.org/10.1186/s12889-019-6391-3>.

By way of an illustration, the interview that newly arrived convicts were subjected to can be described as a cat and mouse game. While the colonial state attempted to glean information that would enable it to effectively control and exploit its charges, it was often in the convict's interests to conceal and obfuscate details of their past lives. The state attempted to minimize the extent to which information was withheld or misleading answers were supplied by interrogating new initiates to the penal colony on board the convict vessel before they could rub shoulders with old hands for example. In addition, each convict was informed that the colonial administration had access to details of their former lives gleaned from British and Irish gaol reports.²⁵

The sheer number of convicts that were subjected to these processes provides an opportunity to check the consistency of the information divulged. One of the questions that fresh arrivals were asked was to list their previous convictions. In all, 61 per cent of male convicts and 70 per cent of female convicts admitted to having been in court before. Comparison with gaol reports forwarded from Britain and Ireland reveal that, on balance, convicts confessed to more offences than were officially recorded against their name. Over 20 per cent of male prisoners with gaol reports describing them as "unknown" admitted to previous convictions.²⁶

The rate at which convicts confessed details of former convictions was consistent with other details they provided. Those who claimed to have superior skills and to be able to read and write were less likely to disclose information about previous offences. The answers provided correspond with data gathered in ways other than interrogation. Stature is influenced by early life experiences. Poor nutrition, childhood diseases, and stress can all lead to stunting. Adult height is closely correlated with socio-economic class because those who escape deprivations in childhood are more likely to attain their biologically programmed stature.²⁷ Each convict was measured on arrival, and it is perhaps not surprising that taller convicts confessed to fewer former convictions. The same was also true for those who brought money to Australia. Examination of deposits made into the Convict Savings Bank reveal that the more cash a convict had, the less likely he

²⁵ P. R. Eldershaw, *Public Records of Tasmania. Section Three. Convict Department* (Hobart: Archives Office of Tasmania, 2003), 7.

²⁶ Tasmanian Archives, Conduct Records, CON 31, 33, 40, and 41.

²⁷ Hamish Maxwell-Stewart, Kris Inwood, and Jim Stankovich, "Prison and the Colonial Family," *The History of the Family* 20, no. 2 (2015): 231–248.

or she was to yield information about former convictions.²⁸ The ability to explore multiple datasets jointly enables deep archival reading. This is a particular quality of time series data.

SPATIAL ANALYSIS AND PENAL LABOR

As well as placing events into temporal sequences, we are interested in the spatial distribution of observations. With the advent of spatial mapping software, there has been a proliferation of digital humanities projects that link traditional forms of data to space and place.²⁹ *Digital Harlem*'s melded use of database and mapping tools to craft spatial linkages that illustrate the social, cultural, and political landscape of Harlem, New York City, is a case in point.³⁰ Others use the realization of geographical and social spaces and places in a multilayered approach known as "Deep Mapping."³¹

Many archival records contain references to locations that can be converted to multiscalar and multitemporal objects. These can be linked to form networks that can be used to explore how people, products, information, and ideas moved between locations. Other information contained in a record may be spatially null, but nevertheless rich in contextual detail. If any single observation in a record chain can be linked to a location, it provides an opportunity to considerably enrich the value of spatial analysis.

The application of these techniques to the study of Australia's convict past is currently in its formative stages. Social science historians have tended to focus on the statistical power obtained as a result of linking temporal data, while historical archaeologists have

²⁸ Tasmanian Archives, Convict Savings Band and Derwent Bank records, CON 147, CSO 71, GO 33-1-14, DB Box 15-18 and State Library New South Wales, Tas Papers 21-4.

²⁹ Ian N. Gregory and Alistair Geddes, "Introduction: From Historical GIS to Spatial Humanities: Deepening Scholarship and Broadening Technology," in *Toward Spatial Humanities: Historical GIS and Spatial History*, ed. Ian N. Gregory and Alistair Geddes (Bloomington & Indianapolis: Indiana University Press, 2014), ix-xix; Keith D. Lilley, "Mapping Truth? Spatial Technologies and the Medieval City: A Critical Cartography," *Post-Classical Archaeologies* 2 (2012): 201-224.

³⁰ Stephen Robertson, Shane White, and Stephen Garton, "Harlem in Black and White: Mapping Race and Place in the 1920s," *Journal of Urban History* 39, no. 5 (2013): 864-880.

³¹ Don Lafreniere and Jason Gilliland, "All the World's a Stage': A GIS Framework for Recreating Personal Time-Space from Qualitative and Quantitative Sources," *Transactions in GIS* 19, no. 2 (2015): 225-246; Tiffany Earley-Spadoni, "Spatial History, Deep Mapping and Digital Storytelling: Archaeology's Future Imagined through an Engagement with the Digital Humanities," *Journal of Archaeological Science* 84 (2017): 95-102.

been disposed toward a disciplinary fixation upon the built and modified landscape and the material culture record.³² There are, however, considerable opportunities for cross fertilization. As others have observed, attaching relationally organized complex data to spatial references can open whole new areas of historical enquiry.³³

Many colonial records have spatial metadata buried within them. Court records, shipping records, accounts of production, published parliamentary papers, and warrants for the movement of prisoners across the colony are common examples. Often, these data require extensive excavation (both figurative and literal) to extract their spatial value. The description lists drawn up when each transport vessel arrived in the colony, for example, contain details of each convict's "native place," and once transcribed, these can be geolocated, enabling professed literacy levels, variations in adult height, and former offences to be mapped.

Spatial references can be particularly important in the analysis of archival systems designed to aid the control and coercion of convict labor. Such documentation ensured that a prisoner's life was both circumscribed by the walls and regulations of their penal confines, and transixed by the clerk's pen to particular points in time and space. The size of the mark made by that pen varied with the resolution of the record, but from their pretransportation lives to their eventual freedom (or death under sentence), the prisoner's progression through the system was mapped according to a series of fixed points in time and space. Digital tools provide the opportunity to capture, analyze, and visualize these records to better understand processes such as the interplay between convict agency and management techniques across different landscapes of labor extraction. In effect, the systematic coding and linking of spatial references can be used to provide the digital archive with a floor plan.

For the duration of a convict's sentence, the conduct record was an active document, the predigital equivalent of a modern prisoner's ankle monitor. The record activated as soon as the prisoner entered the system and only fully deactivated when the words "Dead" or "Absconded" were scrawled across the bottom in red ink. Even when a sentence to transportation had been served, there was a risk that subsequent encounters with the court system might trigger further

³² Tuffin et al., "'Landscapes of Production and Punishment,'" 7–8.

³³ Stephen Robertson, "Putting Harlem on the Map," in *Writing History in the Digital Age*, ed. Jack Dougherty and Kristen Nawrotzki (Ann Arbor: University of Michigan Press, 2013), 186–197.

entries. The total number of charges itemized in these series is impressive: summaries of over 450,000 separate prosecutions are listed for 73,000 convicts. Each individual entry records:

- The date of the hearing.
- The name of the convict's employer or the location at which the convict was employed (or their status if free or holding a ticket of leave).
- A summary of the charge.
- The initials or names of the magistrates who heard the case.
- The sentence.
- Notes on whether the sentence was commuted or otherwise changed because of an executive order.

Digital technologies can be used to organize these data spatially, powerfully adding to the context of the record. There is evidence that distance played a part in the decision to bring a prosecution for example. The expense of bringing a charge against a convict working for a master on a farm located far from the nearest magistrate might act as a powerful disincentive. The convict, in effect, had more opportunity to bargain their way out of trouble on a remote property than they did in a more centrally located workplace.

Factors other than geography also influenced the decision to press charges. The loss of the convict's labor for the duration of the court case and any resultant punishment might weigh on the master's mind—especially if the prisoner possessed skills that were hard to replace. The services of the convict might be lost for months or even years if the magistrate decided that the offence merited a sentence to hard labor on the roads, a house of correction, or a penal station. As was alleged at the time, some masters used the prosecution process to fire convicts with less valued skills, knowing that any loss occasioned by a sentence would result in the provision of a replacement.³⁴ The availability of alternative labor was also likely to factor into the decision-making process. When wages for free workers were depressed and the costs of clothing and feeding a convict were high, the rate at which convicts were prosecuted increased.³⁵

The number of charges that accumulated on a convict's record mattered. The colonial administration used this information to determine whether a prisoner qualified for an indulgence such as permission to marry. An understanding of the complexity of the

³⁴ House of Commons, *British Parliamentary Papers*, XLII (1837–1838), Copy of Despatch, Note (G) Testimonials by Messrs. Backhouse and Walker. p. 25.

³⁵ Hamish Maxwell-Stewart, "Convict Labour Extraction and Transportation from Britain and Ireland 1615–1870," in *Convict Labour: A Global Regime*, ed. C. de Vito and A. Lichtenstein (Leiden, Brill, Studies in Global Social History, 2015): 182–193.

prosecution process, including the spatial and temporal factors that influenced the decision to prosecute, provides important pointers to the way these records can be read. Such understandings are considerably aided when records possess digital transactional histories that have been linked and enriched through the addition of spatial referencing systems.

Many of the summaries of charges listed in the conduct records identify the places in which the offences were allegedly committed, such as at the Government Garden, Port Arthur Penal Station—Port Arthur was a site of punishment that operated from 1830 to 1877.³⁶ The reference to the garden is not in itself locational. However, when the garden location is geolocated, the reference becomes geographically meaningful and can be plotted in Cartesian space. Digitized archival maps and plans can be georeferenced to locations in the landscape, using remnant fabric, archaeological data, or LiDAR-derived Digital Elevation Maps (DEMs) as spatial anchors.³⁷ Structures (including internal layouts where known), cultivation and activity areas, jetties, roads, tramlines (used to move timber, quarried stone and stores), and pathways can be digitized and added as additional layers. In this way, each digitized object can become a container that can be populated with data and linked to form networks.

Each of these containers can be represented in a GIS (Geographic Information System) as polygons, lines, and points. While spatially and temporally inert in themselves, these objects form electronic reference points to which other information can be attached. They can also be organized into layers in order to explore chronological change or arranged thematically. Information in the conduct records can thus be used to locate the places and spaces in which individual prosecutable actions occurred, enabling the exploration of offence patterning across the life of a settlement like Port Arthur. This might include analysis of charges for smoking, fighting, or possession of concealed food or other items such as fishing hooks. Such evidence enables an analysis of the way in which specific spaces were policed, and, of course, convicts' failed attempts to subvert regulation and control. This kind of analysis allows the results of archaeological excavation and a textual analysis of

³⁶ See the online resource: Richard Tuffin, Martin Gibbs, David Roberts, Hamish Maxwell-Stewart, David Roe, Jody Steele, Susan Hood, and Barry Godfrey, 2019, *Convict labour landscapes, Port Arthur 1830–1877*, www.convictlandscapes.com.au.

³⁷ Richard Tuffin and Martin Gibbs, "Repopulating Landscapes: Using Offence Data to Recreate Landscapes of Incarceration and Labour at the Port Arthur Penal Station, 1830–1877," *International Journal of Humanities and Arts Computing* 13, no. 1–2 (2019), 155–181.

the archive to be analyzed side by side. One might expect areas where many fragments of clay pipes were recovered, for example, to be associated with few or no prosecutions for smoking, while the reverse might be true in other areas.

Such data has a three-dimensional quality. The ability to locate multiple events within both space and time is extremely powerful. A particularly important aspect of multidimensional analysis is that it can be used to explore the moments when individual life courses intersect. Such techniques can be particularly informative in convict society where ties with biological kin were effectively destroyed by forced migration. In the absence of fathers, mothers, brothers, and sisters, convicts created fictive kinship networks. Examination of the points of intersection in the life courses of transported brides and grooms and the witnesses who signed their marriage certificates provides a means of exploring the process of colonial relationship formation.

CONCLUDING REMARKS

Digitization gives historians access to vast amounts of information, but more than that, it provides the potential to reconstruct long runs of data and read across multiple record series in parallel.³⁸ We have outlined how the systematic organization of digitized record collections can be used to improve the ecology of the archive and create resources such as branching locational narratives that should negate concerns about the decontextualization of the digital resource. Of course, the approaches we have employed come with their own sets of challenges. In particular, they suggest that historians need to rethink some of the ways in which they have traditionally worked.³⁹ The creation of tools to foster collaboration will necessitate conversations about appropriate standards for coding and describing datasets. There will also be a number of ethical complexities to negotiate.

Developing ways to help users understand the processes that shaped the colonial archive will also be an ongoing issue. Historians acknowledge the mutability of archival records as both a legacy

³⁸ James Barker, "A History of History through the Lens of Our Digital Present, the Traditions That Shape and Constrain Data-Driven Historical Research and What Librarians Can Do about IT," in *Laying the Foundation: Digital Humanities in Academic Libraries*, ed. John White and Heather Glibert (West Lafayette, IN: Purdue University Press, 2016), 15.

³⁹ Hitchcock, "Confronting the Digital," 19.

of colonial sensibility and a continuing part of postcolonial governmentality.⁴⁰ While it is readily accepted that archived documents are subjective and purposeful instrumental tools that were created to record and inform power relations, the selective processes that deposited particular data in repositories have been less widely discussed. The processes of selecting particular data for digitalization is the subject of even less discussion. No matter how documents may be read against or along the grain, if the collection has been curated in a way which omits certain data (and also omits to tell us the basis on which information has been selected), then our analysis—like the collection from which it is derived—will be incomplete. This does not obviate, of course, the imperative to analyze what has been digitized, to read it critically, and to interrogate the data; on the contrary, it makes the exploration and documentation of these processes even more necessary.

In this sense, far from creating new threats to the discipline of history, digitization exposes preexisting problems that lie at the heart of the historical process. Access to record groups has long been mediated by file descriptions, indexes, and other finding aids. Files whose contents have been comprehensively described by librarians and archivists are cited more frequently than those that are poorly described. While it is true that record groups that are digitally indexed are likely to be accessed far more than those that are not, this is hardly a new problem.⁴¹

Nevertheless, there is a danger that uncontextualized use of digitized data will exacerbate this issue. Crime history students in the United Kingdom, for example, can find data on Londoners much more easily than offenders and victims from elsewhere due to the successes of the Old Bailey Online and Digital Panopticon.⁴² There is a risk that selective digitization of data will lead us to neglect communities, nations, and regions. Digitized data made available through commercial sites which favor name-rich datasets for the genealogical market are likely to heighten such trends, creating resources that enable the study of settler populations at the expense of First Nations for example. Such commercial interests can also introduce a form of survivor bias, where records for a subsection of a population with a large

⁴⁰ Stoler, A., *Along the Archival Grain: Epistemic Anxieties and Colonial Common Sense* (Princeton University Press, 2009).

⁴¹ Richard Dunley and Jo Pugh, "How do Catalogues Make History?," paper delivered to the Digital Humanities Australasia Conference, Adelaide, 26 September 2018.

⁴² See <https://www.oldbaileyonline.org> [accessed 2 February 2020] and <https://www.digitalpanopticon.org> [accessed 2 February 2020].

numbers of descendants become more accessible and as a result are viewed as representative of past experience as a whole.⁴³

In order to mitigate these risks, it is necessary to account for how information is statistically and spatially distributed. Otherwise, it is difficult to place any selected observation within a meaningful context. The reality, however, is that the digital revolution has coincided with a marked decline in the use of statistical methods within the discipline of history. Statistical training has also disappeared from most undergraduate history teaching programs.⁴⁴ When historians literally have big data at their finger-tips, the discipline finds itself poorly equipped to meet the challenge of the digital.

Rethinking the way that history is taught at undergraduate level should thus form an important part of any wider historical digital agenda. There is no particular reason to fear a return to more intensive engagement with quantitative techniques. Courses that provide students with access to research datasets and the necessary skills to visualize and otherwise interrogate that data should have appeal. In an era where history is often seen as a nonvocational “luxury”—to quote a former British prime minister—engagement with digital data and its associated advantages and problematics has a wider logic.⁴⁵ There is room to envisage ways in which digitally centered undergraduate teaching might lead to improvements in the way that data is curated and documented. This might in turn lead to the creation of internships and other partnerships with archives and data providers.

There is certainly a wider need to provide users of digitized historical content with greater levels of contextual information. Making data more accessible could lead, for example, to the growth of what one of us has described as “digital dark tourism.”⁴⁶ The availability and speed at which we can now access digital crime data can lead to a dilitantish approach. The attraction of serious crimes, or notorious murders, celebrated criminals, and so, can be alluring. Those of us who spent long hours in the archives, ordering, waiting, retrieving, considering, transcribing, and then ordering another document—

⁴³ Hamish Maxwell-Stewart and Lydia Nicholson, “Penal Transportation, Family History and Convict Tourism,” in *The Palgrave Handbook of Prison Tourism*, ed. Jacqueline Z. Wilson, Sarah Hodgkinson, Justin Piché, and Kevin Walby (London: Palgrave Macmillan, 2017), 274–277.

⁴⁴ James Barker, “A History of History,” 18–24.

⁴⁵ Douglas Greenberg, “History is a Luxury: Mrs. Thatcher, Mr. Disney, and (Public) History,” *Reviews in American History* 26, no. 1 (1998): 294–311.

⁴⁶ Barry Godfrey, *Crime in England 1880–1945: The Rough and the Criminal, the Police and the Incarcerated* (London: Routledge, 2014), 96–104.

had a slower experience of finding data. It also gave us time to consider, reflect, and form opinions over a longer period (sometimes in discussion with whoever else happened to be in the archive café). The speed at which we can now select, access, consider, and discard digital data should not compromise sound and considered analysis. Nor should we encourage students to select and focus on the more dazzling crimes and cases, when they are, very often, the least important for developing theory. Just because the data can be digested quickly, it does not make it more nutritious. These are all appropriate issues to be discussed in undergraduate classrooms.

The classroom is also a good place to address the fallacy that only large well-funded research groups are equipped to work with “Big Data.” While it is true that large interdisciplinary research teams are required to generate datasets that contain millions of lines of data and thousands of life histories, engagement with such datasets should not cease at the point where the publically funded research projects that lead to their creation cease. Initiatives such as The Old Bailey Online, Digital Panopticon, and Founders and Survivors were designed to generate digital data for use by others, as well as to produce research findings in their own right.⁴⁷ The groups who produced the digital “product” may have been large, but the community of potential users is many, many times larger. That community should include undergraduates studying history and related disciplines.

Moreover, researchers do not necessarily need to analyze thousands of cases. Digital data is as useful to the microhistorian as it is to the “global picture” researcher. Indeed, it might be argued that it is essential in that it enables the case study to be located within the wider whole. “Thick” analysis can also be an effective way of exploring the deeper dimensions and complex dynamics that underlie systematic patterns visible in broad but, otherwise, thin sources.

There is another side to this particular coin. The comprehensiveness of life course data assembled through the linking of census, criminal, and other civil records, can lead to the assumption that historical lives were entirely driven by their interactions with the state. Official documents are compiled for specific purposes. They are not replete with details of the underlying motivations and experiences of the people that were subject to such state surveillance. Humans have complex relationships, impressive imaginations, and reasoning

⁴⁷ Founders and Survivors datasets have been made available through the Digital Panopticon and The Tasmanian Name Index <https://www.libraries.tas.gov.au/how-to/Pages/Names-Index-content.aspx> [accessed 2 February 2020].

abilities. While we will never be in a position to fully understand the behavior of individuals, micro studies that explore lives that are documented in greater detail than most provide an opportunity to see the limitations of wider processes of civil recording.⁴⁸ Indeed, the linking of information contained in digitized diaries and other qualitative records to relevant entries in large administrative datasets should form a crucial part of the development of future historical data ecologies. Spatial and networking tools could facilitate deeper engagement across these different types of source materials.

Historians have practiced the uncontextualized search for information for generations. As Lara Putman puts it, it is unusual for historians to “build in systematic checks against omitted variable bias.”⁴⁹ While we agree with those who argue that deep reading is an important part of the historical process, in a digital age, we would add that counting and charting form essential constituent components of reading.⁵⁰ By combining modes of working, historians can take advantage of digitization without running the risk of “disintermediated discovery.” While we will have to watch for selection bias, and other distorting processes including the acquisition of funding, transcription, and tagging exercises aimed at particular record groups at the expense of others, the necessary tools and resources already exist to train a new generation of historians and archaeologists to be alert to these issues.

In the end, we can mull over the advantages and disadvantages associated with the digitization of the archive; however, we should not be fooled into thinking that nonengagement is a choice—the horse has already bolted. We need to ensure that the rigor with which we approached our work in predigital times continues. Indeed, we would argue that it can and should intensify. The digital age is upon us and historians will increasingly rely on digital data. Yet, working with related disciplines, we have the opportunity to build complex datasets and create the necessary contextual tools to guide future users of those resources. If historians do not rise to the task of building such ethical data ecologies, they risk losing access to the keys to the digital archive. If this were to occur, there can be little doubt that the discipline would be collectively poorer.

⁴⁸ Barry Godfrey, *Crime in England 1880–1945: The Rough and the Criminal, the Police and the Incarcerated* (London: Routledge, 2014), 96–104. pp. 216.

⁴⁹ Laura Putman, “The Transnational and the Text-Searchable,” 384.

⁵⁰ Franco Moretti, *Graphs, Maps, Trees: Abstract Models for a Literary History* (London: Verso, 2005).

Barry Godfrey is a professor of social justice at the University of Liverpool, UK. He has published widely on the social history of crime and the criminalized in collaboration with researchers in the UK, Australia, and North America.

Caroline Homer is the manager of the State Library and Archive Service, Tasmanian State Government, Australia. The State Library and Archive Service provides access to the Tasmanian Archives, including the Tasmanian Convict records, which are listed on the UNESCO Memory of the World International Register.

Kris Inwood is a professor at the University of Guelph, Canada, where he holds appointments in both economics and history. He has a particular interest in the ways in which military and criminal justice records can be used to explore the impacts of settler colonialism.

Hamish Maxwell-Stewart is a professor of history at the University of New England, Australia. He has published many books and articles on overseas deployment of convict labour and has a growing interest in the use of digital data to explore life course and intergenerational outcomes.

Rebecca Read recently completed her doctorate in history at the University of Tasmania, Australia. She primarily works on differences in prosecution risk across colonial contexts.

Richard Tuffin is a postdoctoral research fellow in historical archaeology at the University of New England, Australia. He is particularly interested in the use of historical GIS systems to explore colonial labor landscapes.