
Digital data methods and historical criminology: with the increased digitisation of archival records, how do digital data methods interact with historical criminology?

Dr Zoe Alker, University of Liverpool and **Dr Alana Piper**, University of Technology, Sydney.

Digital technology has fundamentally changed the ways we approach academic, archival and public histories of crime and historical criminology. Digitisation alters the ways that we locate, analyse and present research, and teach upcoming scholars, a process of transformation that continues to unfold. As part of BSC's Historical Criminology network [#HCNet], scholars from the UK and Australasia met to discuss the capabilities and constraints of using digital data methodologies for historical criminology. The group firstly mapped out the existing landscape of digitised historic crime and justice data before drawing out the possibilities and challenges to future developments and international research collaborations.

Digital resources for crime history and historic criminology have multiplied since the publication of Old Bailey Online in 2003. From projects which aimed to recover hidden criminal lives [<https://www.digitalpanopticon.org/>; <https://www.londonlives.org/>; <https://ourcriminalancestors.org/>; <https://cockatooconvicts.wordpress.com/>; <https://criminalcharacters.com/>; <https://foundersandsurvivors.com/>] to projects which sought to expose the problematic machinery of criminal justice [<http://convictvoyages.org/>; <https://prosecutionproject.griffith.edu.au/>; <https://www.traffickingpast.uk/>; <http://esrcvictims.org/>], a wealth of social and criminal records from the eighteenth, nineteenth and twentieth centuries are now available to historians, criminologists and the wider public. Once we add to the mix the vast amount of online social and economic records, including newspapers and parliamentary papers, it is clear that historical criminologists are dealing with an abundance, rather than scarcity, of material. The collection of online records facilitates a rich variety of enquiries –

including analyses of social representations of crime, investigations of trial outcomes, and studies of life courses – to name but a few.

But, while this landscape of material initially appears to paint a picture of abundance in historic crime and criminal justice data [at least in the Western context], there are barriers to being able to use digital resources to respond to the shared critical values and research enquiries of historians and criminologists. While the diversity of digitised records has certainly illuminated a range of historical enquiries, what about those who stand in the shadows of digitisation? Digitising non-elite, non-Western resources is essential to capturing the histories of the marginalised and exposing the intricacies of power embedded within the criminal justice system. The under-representation of BAME and Indigenous groups in digital crime history – and criminology more broadly – invites a political commitment to working with communities, activists and archivists to recover previously hidden histories of Indigenous and BAME groups (for example, see projects including Real Face of White Australia). In turn, digitisation of material produced by the colonial criminal justice system risks the potential objectification of vulnerable groups – be they Indigenous peoples, juveniles or offenders generally. It is essential that researchers challenge systemic biases by providing audiences with contextual information that makes explicit how colonial violence is embedded within the records now so easily accessible, and by increasing accessibility further through multi-lingual OCR software translation. Roberto Catello (Liverpool Hope University) brought up the lack of ethical frameworks in digitisation in relation to Indigenous persons. In many instances researchers working with historical materials are not required to fill out ethics applications, but making ethics applications a standard component of historical studies would foster more ethical and inclusive approaches to digitising and presenting historic crime and justice data. It is crucial that researchers, communities and archivists centralise their shared political aim of exposing the elite bias and white privilege inherent in the criminal justice system – past and present.

The vast public interest in family history has led to a widespread commercialisation of social data including birth, marriage and death records, the census (in UK), and other records useful to historians of crime and justice, particularly newspapers. The commercialisation of data has brought as many constraints as possibilities. Susanne

Karstedt (Griffith University) pointed out that often crime historians will not just be working with crime data, but want access to other forms of socio-economic data (e.g. BMD records). While large-scale funded projects such as Digital Panopticon have synthesised social, legal and personal records, other forms of data useful to crime historians and criminologists are often restricted by paywalls implemented by sites such as Ancestry, British Library Newspapers Online and British Newspaper Archive. This brings together inter-related issues; putting publicly-owned archival data behind paywalls limits accessibility through access and affordability, but also impacts on functionality and useability. Commercial datasets often hide their digitisation processes and workflows behind user-friendly interfaces, and the poor quality of OCR and transcriptions on commercial sites means that online archives are often useful for family researchers who rely upon individual queries, but are not so practical for academic researchers who want to conduct large-scale queries or analyse the material in detail and with accuracy.

The group highlighted problems with interoperability, in particular, the ability to share and synthesise criminal justice and social records across international borders. Integrating criminal records requires expert knowledge in developing coding frameworks that are replicable and reusable across project, such as the streamlining of offence categories. HISCO's development of an international standardisation of occupational classification has been hugely influential in fostering international, comparative enquiries, and discussions commenced at the 2018 European Social Science History conference about the feasibility of a similar system of offence classifications to study the criminal justice system across international and jurisdictional boundaries. There is a need to make datasets open, freely accessible and transparent to further facilitate international collaborations. Some projects, including the Prosecution Project and Digital Panopticon, are already making use of research tools and repositories, including DataVerse and Figshare, to enable the publication and sharing of criminal justice datasets, but programming codes should also be shared, like Shoemaker and Alker's Convict Tattoos project codes, through sites such as GitHub, for the benefit of other researchers and to make research workflows transparent. Online archive repositories such as UK Data Service and Australia's Dataverse offer potential for storage and sustainability and the sharing of materials, but existing funding models don't support long-term storage or ongoing

administration of successful digital projects. More low-level, sustained funding that supports the ongoing maintenance of datasets and tools would be useful in ensuring projects' sustainability.

The meeting was extremely productive in bringing together crime historians and historical criminologists to share their experiences and highlight critical values and concerns. The Chairs thank the delegates for their contributions and hope that this will be the start of many future conversations and collaborations between the UK and Australasia.

Dr Zoe Alker (University of Liverpool) and Dr Alana Piper (University of Technology, Sydney)
