

**ADDRESSING INFORMATION ASYMMETRY IN THE SOCIAL
CONTRACT: AN ARCHIVAL-DIPLOMATIC APPROACH TO OPEN
GOVERNMENT DATA CURATION**

James Lowry

**A thesis submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy**

Department of Information Studies

UCL

London, UK

2019

Declaration of Originality

I, James Lowry, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed: James Lowry
PhD candidate

Dedicated to William and Elaine Lowry

ABSTRACT

This thesis shows that the concepts and practices developed in the field of record-keeping can be applied to the curation of open government data to strengthen the trustworthiness of that data.

It begins by situating open government data in the context of the social contract, which operates through the exchange of information. The thesis develops the notions of the 'record-as-command' and 'data-as-command' to explain the dialogical but asymmetrical information relationship between the individual and the state, which is modelled as a principal-agent problem. Using concepts from information economics, the study argues that open government data is the latest monitoring mechanism in a long history of government secrecy and openness. This establishes the significance of the curation of open government data beyond technical questions.

The thesis then considers how trustworthiness has figured in thinking about record-keeping, identifying three core record-keeping controls; 1) metadata used to document 2) custodianship in 3) auditable systems. To show how these three broad controls have been put into practice, the study examines two examples of record-keeping guidance, one for paper records and one for hybrid records, which demonstrates the application of the three core controls across time and media. The study then looks for the presence or absence of these controls in government datasets published in Kenya and Australia. These analyses trace the datasets back to their source(s), at each step looking for evidence of custodial and curatorial interventions documented in metadata managed in auditable systems.

The study's contribution to open government data work is its demonstration of the value of record-keeping controls in the curation of data. Additionally, the study contributes new insights to information in the principal-agent problem of the social contract, contributes to archival theory and finds a need to foster critical data literacy in the body politic if open government data is to be read and used to correct information asymmetry.

IMPACT STATEMENT

The insights set out in this thesis could be put to beneficial use in the curation of open government data. As the thesis explains, the open data community is becoming increasingly interested in data quality, including data trustworthiness, and is seeking methods for improving open data. Some of that work is reinventing concepts and techniques that have been developed over many centuries in the field of record-keeping, and those record-keeping concepts and techniques could usefully be introduced into the open data community and its projects in order to make open data more trustworthy. The findings of the thesis could inform data curation practices in terms of thinking about how data is collected, aggregated, used, published and repurposed, and throughout these processes, documented in systems that allow custodianship to be audited. Such interventions would improve data quality by improving trustworthiness, which in turn could have substantial impacts on information asymmetry under the social contract. Improvements in open government data trustworthiness might have other benefits further down the data value chain, in terms of social enterprise, commercial activity, public policy development, public service delivery and the quality of the environment. These effects might be seen on the individual, communal and societal levels. These possible benefits are not restricted to local or national contexts, but may be effected wherever data processing could be improved. Beyond these technical and social impacts, the benefits of this research might include benefits within academia, including greater interdisciplinarity in the teaching of records and archives management and data curation and analysis. These impacts could be brought about by disseminating the findings through scholarly publications, but the primary focus of dissemination should be in easily accessible (free and online) guidance with an applied focus aimed at data curators and users, which is the audience best placed to implement the findings of this research.

ACKNOWLEDGEMENTS

I am grateful to my supervisors, Elizabeth Shepherd, Andrew Flinn and Elizabeth Lomas, for their advice and patience. For their support or their help with my research, thanks to Prestone Adie (Data Analyst, Open Data Initiative, ICT Authority, Kenya), Nisa Bayindir, Mark Brogan (Senior Lecturer, School of Computer and Security Science, Edith Cowan University), Alexandrina Buchanan (Senior Lecturer, Liverpool University Centre for Archive Studies), Marilyn Deegan (Professor, Department of Digital Humanities, Kings College London), Frederick Galtung (President, Integrity Action), Anne Gilliland (Professor, Department of Information Studies, UCLA), Erin Harris, Edward Kosgei (Head of Lands Administration, Lands Department, Ministry of Lands and Physical Planning, Kenya), Catherine Lowry, Gabriel MacCormack, Patricia MacCormack (Professor, School of Humanities and Social Sciences, Anglia Ruskin University), Sifa Mawiyoo (Open Data Specialist and GIS Technologist, Open Data Initiative, ICT Authority, Kenya), Peter Molnar (Senior Research Fellow, Centre for Media and Communications Studies, Central European University), Sandra Musoga (Senior Programme Officer, Transparency, Article 19, Kenya), Leonida Mutuku (Founder and CEO, InteliPro, Kenya), Francis Mwangi (Director, Kenya National Archives and Documentation Service), Emily Ndungi (former Principal Records Management Officer, Lands Department, Ministry of Lands and Physical Planning, Kenya), Liana Saif (then British Academy postdoctoral fellow at the Oriental Institute at the University of Oxford), Anthea Seles (then at the National Archives, UK), Ember Taylor (then at the University of Liverpool), Anne Thurston (Director, International Records Management Trust), Megan Venter (then at the University of Liverpool), Justus Wamukoya (Professor and Dean, School of Information Sciences, Moi University, Kenya), Richard Wato (Assistant Director, Kenya National Archives and Documentation Service) and Stacy Wood (Assistant Professor, School of Computing and Information, University of Pittsburgh).

TABLE OF CONTENTS

LIST OF FIGURES	11
LIST OF ACRONYMS	12
Chapter 1: Introduction and Research Design	15
1.1. Introduction	16
1.2. Justification of the Topic - An Archival Imperative	17
1.3. Research Questions	19
1.4. Methodology	20
1.5. Limitations	26
1.6. Definitions	28
1.7. Thesis Outline	31
Chapter 2: Information and the Social Contract	35
2.1. Introduction	36
2.2. Commands	36
2.3. The Record as Command	40
2.4. Bureaucracy as Apparatus	48
2.5. Information in the Societies of Discipline and Control	55
2.6. Conclusion	63
Chapter 3: Information Asymmetry in the Social Contract	67
3.1. Introduction	68
3.2. The Social Contract and Agency	68
3.3. Information Asymmetry	77
3.4. Information Economics Approaches to Information Asymmetry	82
3.4.1. <i>Selection Problems (Adverse Selection)</i>	82
3.4.2. <i>Incentive Problems (Moral Hazard)</i>	84
3.5. Information Asymmetry in the Social Contract	86
3.6. Conclusion	89
Chapter 4: Opening Government: A History of Monitoring	91
4.1. Introduction	92

4.2. The Development of a Culture of Secrecy in Government	92
4.3. Archival Access	98
4.4. Whistleblowing	102
4.5. Freedom of Information	106
4.6. Open Government Data and the Open Government Partnership	114
4.7. Conclusion	121
Chapter 5: Trustworthy Information	125
5.1. Introduction	126
5.2. Informational Trustworthiness in Historical Perspective	126
5.3. Archival Approaches to Trustworthiness	131
5.4. Defining Trustworthiness	132
5.4.1. <i>Reliability and Accuracy</i>	133
5.4.2. <i>Authenticity</i>	137
5.5. Accuracy, Reliability or Authenticity?	139
5.6. Establishing and Maintaining Authenticity	141
5.6.1. <i>Metadata</i>	141
5.6.2. <i>Systems</i>	142
5.6.3. <i>Custodianship</i>	144
5.7. A Note on Originality and Fixed Forms	150
5.8. Conclusion	153
Chapter 6: Practices for Establishing and Maintaining Authenticity in Information	157
6.1. Introduction	158
6.2. The Bengal Records Manual 1943	159
6.3. ICA-REQ	162
6.4. Controls for Authenticity: The Bengal Manual & ICA-Req in Comparison	167
6.4.1. <i>Registration</i>	167
6.4.2. <i>File Registration</i>	175
6.4.3. <i>Forward Diaries</i>	177
6.4.4. <i>Naming and Numbering Conventions</i>	178
6.4.5. <i>Foliation and Fly-Leafs</i>	180
6.4.6. <i>Tracking: Removal Slips, Receipts for Transfer and</i>	

<i>Assurances of Safe Custody</i>	182
6.4.7. <i>Cross-Referencing and ‘Stitching’</i>	184
6.4.8. <i>Copying</i>	185
6.4.9. <i>Security Classifications</i>	187
6.4.10. <i>Records Destruction</i>	189
6.4.11. <i>Into the Archives</i>	190
6.5. Conclusion	192
Chapter 7: Case Study: The Kenya Open Data Initiative	196
7.1. Introduction	197
7.2. Historical Development of Open Data in Kenya	197
7.2.1. <i>Secrecy in Kenyan Government</i>	197
7.2.2. <i>Access to Information: Open Data and Freedom of Information</i>	200
7.3. Current State of Kenya’s Openness Programme	210
7.4. The Role of Civil Society	216
7.5. Examination of Land Data	220
7.5.1. <i>Land information management</i>	220
7.5.2. <i>Examination of land dataset</i>	225
7.6. Conclusion	230
Chapter 8: Case Study: data.gov.au	234
8.1. Introduction	235
8.2. Historical Development of Open Data in Australia	235
8.2.1. <i>Secrecy in Australian Government</i>	235
8.2.2. <i>Access to Information: Open Data and Freedom of Information</i>	237
8.3. Current State of Australia’s Openness Programme	247
8.4. The Role of Civil Society	249
8.5. Examination of Land Data	251
8.5.1. <i>Land information management</i>	251
8.5.2. <i>Examination of land dataset</i>	255
8.6. Conclusion	265

Chapter 9: Applying Record-Keeping Controls to Government Data Curation	267
9.1. Introduction	268
9.2. Comparison between Kenya and Australia	268
<i>9.2.1. Metadata</i>	269
<i>9.2.2. Documented Custodianship</i>	272
<i>9.2.3. Auditable Systems</i>	276
9.3. The Application of Record-Keeping Controls to Datasets	281
9.4. Conclusion	286
Chapter 10: Conclusion	288
10.1. Introduction	289
10.2. Summary of Thesis	289
10.3. Conclusions	300
10.4. Critical Data Literacy: The Need for Further Research	315
10.5. Conclusion	317
Bibliography	319

LIST OF FIGURES

Figure 1. Information economics approaches to addressing pre- and post-contract principal-agent problems.

Figure 2. Left-hand docket stamp, the Bengal Manual, 1948.

Figure 3. Right-hand docket stamp, the Bengal Manual, 1948.

Figure 4. KODI interface on 11 January 2017.

Figure 5. KODI sign in error message.

Figure 6. KODI sign in error message (close up).

Figure 7. Additional Information metadata element set for *Land Management Practices* dataset.

Figure 8. Relative Standard Errors of Selected Estimates

LIST OF ACRONYMS

AAE	Australian Agricultural Environments
ABF Act	Australian Border Force Act
ABS	Australian Bureau of Statistics
AGIFT	<i>Australian Governments' Interactive Functions Thesaurus</i>
ALRC	Australian Law Reform Commission
ANZSIC	Australia and New Zealand Standard Industrial Classification
API	application programming interface
ARNECC	Australian Registrars' National Electronic Conveyancing Council
AU	African Union
CDF	Constituency Development Fund
CFOI	(UK) Campaign for Freedom of Information
COP	Chain of Preservation
CSIRO	Commonwealth Scientific and Industrial Research Organisation
CSV	comma separated value (file format)
DARPA	(US) Defense Advanced Research Projects Agency
DDF	De-identification Decision-Making Framework
DDI	Data Documentation Initiative
DEFRR	Digital Evidence First Responder
DLM Forum	Document Lifecycle Management Forum
D-Notice	Defence Notice
DoD	(US) Department of Defense
DROID	Digital Record Object Identification
EASSY	Eastern Africa Submarine Cable System
ECA	Economic Commission for Africa
EDRMS	Electronic Document and Records Management System
ERM	electronic records management
FOI	Freedom of Information
GCHQ	(UK) Government Communications Headquarters
GDR	German Democratic Republic
GIF	Graphics Interchange Format
GIS	geographic information system
GPSDD	Global Partnership for Sustainable Development Data

KIPPRA	Kenya Institute for Public Policy Research and Analysis
KNADS	Kenya National Archives and Documentation Service
KNBS	Kenya National Bureau of Statistics
KODI	Kenya Open Data Initiative
IBA	International Bar Association
ICA	International Council on Archives
ICA-Req	International Council on Archives Principles and functional requirements for records in electronic office environments
InterPARES	International Research on Permanent Authentic Records in Electronic Systems
IPS	Information Publication Scheme
IRM	(OGP) Independent Reporting Mechanism
IRMS	Integrated Records Management System
IRMT	International Records Management Trust
ISO	International Standards Organisation
JPEG	Joint Photographic Experts Group (file format)
LaMPS	Land Management Practices Survey
MI5	(UK government) Security Service
MI6	(UK government) Secret Intelligence Service
MoReq	(EU) Model Requirements for the Management of Electronic Records
MP	Member of Parliament
NAA	National Archives of Australia
NAP	(OGP) National Action Plan
NBC	(US) National Broadcasting Company
NGO	non-government organisation
NOARK	Norsk arkivstandard (Norwegian Archives Standard)
NSK	Neue Slowenische Kunst (New Slovenian Art)
NSOs	National Statistics Offices
OAIC	Office of the Australian Information Commissioner
OAIS	Open Archival Information Systems
ODB	Open Data Barometer
ODI	Open Data Institute
OGD	Open Government Declaration

OGP	Open Government Partnership
OHCS	Office of the Head of the Civil Service
PCAW	Public Concern at Work
PDF	Portable Document Format
PDF/A	Portable Document Format / Archival
PEXA	Property Exchange Australia (system)
PIDA	Public Interest Disclosure Act
PRO	Public Record Office
PROV-DM	PROV (for 'provenance') Data Model
PROV-O	PROV (for 'provenance') Ontology
RCRG	Records Continuum Research Group
RDIM	Records/Document/Information Management: Integrated Document Management System for the Government of Canada
RSE	relative standard error
PSI	public sector information
SAILIS	South Australian Integrated Land Information System
SDMX	Statistical Data and Metadata Standard
TAI	Transparency and Accountability Initiative
TEAMS	The East African Marine System
TNA	The (UK) National Archives
UBC	University of British Columbia
UBC-MAS	University of British Columbia Master of Archival Science
UCLA	University of California, Los Angeles
UNECA	United Nations Economic Commission for Africa
URL	unique resource locator
W3C	World Wide Web Consortium
XML	eXtensible Mark-up Language

Chapter 1

Introduction and Research Design

1.1. Introduction

A file in the Graphics Interchange Format (GIF) was characterised as a deadly weapon in the criminal charge made against John Rayne Rivello after he tweeted a GIF with a strobe effect to journalist Kurt Eichenwald, triggering a seizure.¹ In the *London Review of Books*, William Davies wrote of ‘weaponised paperwork’ creating hostile environments for migrants.² Files, paperwork, have always been weaponisable. In legal and bureaucratic systems, information can be deployed in assertive, aggressive and violent ways: it can be a type of force. In the civil sphere, the way that this force moves back and forth between actors both enacts and exposes how power operates at a given time or place; analyses of records and data and record-keeping and data curation processes can therefore reveal power asymmetries. Access to this powerful, weaponisable commodity that is information is often asymmetrical in the citizen/state relationship because of the organisational and computational power of the state’s bureaucracy to amass, sift, aggregate and correlate data. Efforts to correct this information asymmetry, broadly conceived of here as the open government movement, often focus on access to quantities of government data, rather than critiquing the qualities of government data. At the heart of this thesis is the premise that assurances of data trustworthiness have a powerful effect on information asymmetry, which correlates with power asymmetries in the environment of the social contract. This thesis will show that the latest instantiation of government openness – open government data – better enables the power potential of information to be enjoyed by individuals and communities if assurances of trustworthiness are afforded by the systems and practices of data curation that bring datasets from their source(s) into the public domain.

Once in the public domain, data often appear unproblematic, but many machines, processes, people, ideas and biases influence data during its collection, curation

¹ Kang, C., *A Tweet to Kurt Eichenwald, a Strobe and a Seizure. Now, an Arrest.*, New York Times, <https://www.nytimes.com/2017/03/17/technology/social-media-attack-that-set-off-a-seizure-leads-to-an-arrest.html> [last accessed 23 November 2018].

² Davies, W., ‘Weaponising Paperwork,’ *London Review of Books*, Vol 40 Number 9 (10 May 2018).

and publication. Observing the same of astronomical data, Matthew Stanley wrote that

A table of ancient eclipses appears to a casual reader as a simple chart of numbers indicating time and place... They have all the crispness and precision one expects from astronomy. But in the history of those numbers lies hidden significant intellectual and interdisciplinary struggle whose outcome was far from unanimous within the astronomical community, and continues to bedevil the calculation of lunar motions today.³

‘These controversies’, Stanley says, ‘...are rarely made public in the published data’.⁴ This thesis will show that the making public of influences on and decisions about data is significant for information symmetry.

To begin that argument, this chapter introduces the topic of the thesis, providing a justification for the study of information in the social contract and the need to bring open government data and record-keeping into conversation for improved informational trustworthiness. The chapter then sets out the research questions, explains the methodology, and identifies the limitations of the research. The chapter provides definitions to the key concepts used in this thesis, though they are elaborated in subsequent chapters, and concludes with an outline of the structure of the thesis.

1.2. Justification of the Topic - An Archival Imperative

In 2011, Arofan Gregory, Executive Manager of the Open Data Foundation, a US-based non-profit organisation, published a paper that called for closer collaboration between open government data initiatives, the linked data community, statisticians and the research data community.⁵ Through this inter-

³ Stanley, M., ‘Where is that Moon, Anyway? The Problem of Interpreting Historical Solar Eclipse Observations’, in Gitelman, L. (ed.) *Raw Data is an Oxymoron*, (Cambridge: MIT Press, 2013) p.86.

⁴ Stanley, ‘Where is that Moon, Anyway?’, p.86.

⁵ Gregory, A., ‘Open Data and Metadata Standards: Should We Be Satisfied with “Good Enough”?’ , Open Data Foundation (16 June 2011).

disciplinary working, Gregory argued, our ability to interpret datasets would be improved and the value of the data would be increased. He claimed that users of 'open data could deploy generic tools for accessing, understanding, and comparing data coming from many sources. While many different technologies promise this functionality, in truth it is only the combination of technology and expertise coming from these different communities which can make it a reality'.⁶ Gregory's call to interdisciplinarity is as necessary in 2019 as it was in 2011, as the open data community has only started to recognise the value of principles and techniques long established in the field of record-keeping.

For the purposes of this thesis, the field of record-keeping means that profession practised by information and records managers, archivists, and the many related administrative roles historically demarcated as registrars, clerks and secretaries, taking in a technical body of knowledge developed in the scholarly discipline of archival studies and the applied discipline of archival science, and the related fields of digital preservation and curation, as well as the clerical techniques that have evolved in office work. As I will show, the record-keeping profession has developed, over hundreds of years, technical controls to safeguard and demonstrate the trustworthiness of information. Importing this expertise into the mix of open government data, linked data, statistics and research data management will further increase the value of open government data because it will facilitate the interrogation of open data for trustworthiness. The exclusion of this body of knowledge from relevant interdisciplinary work, more correctly posited as a failure of the record-keeping profession to see the movements occurring around it and proffer its own expertise, is not a new phenomenon. Colin Burke's history of American intelligence information systems shows how the collaboration of librarians, computer scientists and intelligence experts on the information organisation and retrieval problems of a heterogeneous bureaucracy under Cold War imperatives took years to identify provenance as a valuable organising principle: a matter of fact for archivists since at least 1841.⁷ This thesis

<http://odaf.org/papers/Open%20Data%20and%20Metadata%20Standards.pdf> [last accessed 26 February 2019].

⁶ Ibid, p.4.

⁷ Burke, C. *America's Information Wars: The Untold Story of Information Systems in America's Conflicts and Politics from World War II to the Internet Age* (Lanham: Rowman and Littlefield,

seeks to connect two areas of work – record-keeping and data curation - in order to strengthen open government data.

This thesis is particularly focused on open government data because it is the newest example of a mechanism for opening government information to the public. While earlier approaches to openness, such as Freedom of Information (FOI) legislation, often recognised the contribution of records and record-keeping, the value of record-keeping has not been made evident in the open data context.⁸ This is because FOI laws envisaged searches for and retrieval of information from records, whereas open data initiatives see data derived from numerous and diverse sources. For open government data, in contrast, I will argue that the contribution of record-keeping is in making that data more like records. The potential contribution of the record-keeping field has perhaps been less obvious in relation to open data because records are not necessarily the primary source of data for open data portals and civic technologies, and access, rather than quality, has been the main concern of the open data movement to date.

1.3. Research Questions

These issues will be explored by a study of the way information works under the social contract, how information asymmetry has been theorised, and an historical overview of openness initiatives, of which open government data is only the latest. Following this work, the thesis will then turn to the primary research questions for this study:

- How has the record-keeping profession defined trustworthiness in information?

2018). Sweeney, S. 'The Ambiguous Origins of the Archival Principle of "Provenance"', *Libraries & the Cultural Record*, Vol. 43, No. 2 (2008) pp.193-213.

⁸ Ministry of Justice, *Lord Chancellor's Code of Practice on the management of records issued under section 46 of the Freedom of Information Act 2000*.

<http://www.nationalarchives.gov.uk/documents/information-management/foi-section-46-code-of-practice.pdf> [last accessed 28 February 2019].

- What controls has the record-keeping profession developed to offer assurances of trustworthiness and the means of assessing trustworthiness in information?
- Can these controls be used to improve the trustworthiness of open government data?

The answers to these questions will demonstrate that open government data curation practices can be improved to offer assurances of data trustworthiness, with positive consequences for information symmetry under the social contract.

1.4. Methodology

This research used mixed methods to answer these questions; methods that included literature-based research, historical research using primary and secondary sources, case studies, archival-diplomatic analyses and comparative analyses. These choices were informed by the needs of the particular component parts of the thesis; for instance, the theoretical framework required analysis of the literature while the thesis that archival-diplomatics can improve data quality required a diplomatic analysis of datasets. Ontologically, however, the methods are linked by emphases on historicisation and comparison.

Firstly, the research employed a wide-ranging literature review of political philosophy and information studies to explore the functions of information under the social contract. The findings of this review are used in Chapter 2 to construct a picture of the socio-political context for the work on trustworthiness and open government data. Secondly, the research involved a review of the theorising of information asymmetry and agency, taking in work in the fields of information economics and agency theory. This afforded a model - the principal-agent problem model - for thinking about information problems in the social contract, serving as a framework and set of terms for describing and analysing the dynamics at work around open government data. The form of this model and its alteration for use in this research will be elaborated in Chapter 3.

An historical enquiry into the development of the open government movement then used primary and secondary sources to contextualise open government data. The search strategy for this element of the work comprised a search of search engines, online databases and library catalogues for the terms ‘secrecy’, ‘government secrecy’, ‘culture of secrecy’, ‘freedom of information’, ‘access to information’, ‘right to information’, ‘open government’, ‘open government movement’, ‘open government data’, ‘open data’, ‘open data movement’ and ‘open government partnership’. Checking references within the works brought up in the literature review ensured that the key sources on the topic were identified and analysed.

A literature search was also undertaken in relation to the trustworthiness of information. This search focused on the record-keeping literature, since it is that professional discourse that, I posit, is helpful in improving open government data trustworthiness. The search included terms ‘trustworthiness’, ‘authenticity’, ‘reliability’ and was conducted in search engines, online databases and library catalogues as well as technical standards issued by the International Standards Organisation’s Technical Committee 46 (Information and Documentation) Subcommittee 11 (Archives / Records Management) and the International Research on Permanent Authentic Records in Electronic Systems (InterPARES) project, as a well-known and expansive investigation of the question of trustworthiness in digital records.⁹ Again, reviews of the works returned in search results revealed further sources that were examined until the key texts had been identified and reviewed. This work resulted in the definition of key concepts related to the trustworthiness of information, and their relationships to each other. These concepts are articulated in archival-diplomatics, defined here as the combination of archival studies of digital records and diplomatics; this thesis, therefore, employs archival-diplomatics as a lens for examining the problem of data trustworthiness.

Archival-diplomatics is a part of the research methodology for this thesis. Using its terms and concepts, the research investigated how traditional approaches to

⁹ InterPARES Project <http://interpares.org> [last accessed 28 February 2019].

government record-keeping have incorporated controls that provide some assurance of trustworthiness in records. This part of the study consisted of a comparative exegesis of a registry manual, as an instantiation of traditional record-keeping guidance, and an international standard for digital record-keeping systems: the 1943 Bengal Records Manual, which documents an elaborate registry system for the management of paper records, and ICA-Req, an international standard for functional requirements for hybrid digital / paper records management systems.¹⁰ The Bengal Manual was chosen because, as discussed in Chapter 6, it is highly detailed example of a registry manual and its analysis effectively shows the connections between the concepts that are central to this thesis: the documentation of custodial activity in auditable systems. Similar manuals and systems will have existed in other parts of the Commonwealth, such as Australia and Kenya, but it has not been possible for me to establish if copies survive. As described in chapter 6, contemporary registry manuals, such as can be found in Kenya today, for instance, describe less detailed processes and controls and therefore offer less insight into how the controls work in concert to assure trustworthiness. ICA-Req was chosen as a point of comparison because its details allows a close consideration of the linkages between these concepts. The comparison of the manual and the standard identified individual controls and how they work in combination to effect a system that protects informational authenticity. This was undertaken to show that the core principles identified in the preceding literature review have been operationalised in different ways over time, but have remained consistent. Randolph Head has suggested that comparative methods 'are applicable in a wide range of disciplines in the human sciences, and cover a wide range of approaches'.¹¹ Comparative analysis is an appropriate method here because I aimed to find out if the principles could be applied in diverse contexts, since open data curation occurs across diverse contexts and comprises diverse practices: my comparison of the manual and standard show that the principles are consistent across space, time and records formats.

¹⁰ Board of Revenue, Government of Bengal, *The Bengal Records Manual 1943* (Alipore: Bengal Government Press, 1943). International Standards Organisation, *ISO16175:2011 - Information and documentation - Principles and functional requirements for records in electronic office environments* (ISO 2011).

¹¹ Head, R., 'A Comparative Case-Study Approach to Historical Archives in Europe,' in Gilliland, A., McKemmish, S. & Lau, A. (eds.) *Research in the Archival Multiverse* (Melbourne: Monash University Publishing, 2016) p.439.

Having established the validity of the controls for trustworthiness identified through the literature review in this way, the research then included two case studies of open government data initiatives, to demonstrate the application of the principles to open government data. Case studies are an appropriate method for this research because they enable analyses of specifics and are 'intended to allow for transferability of findings based on contextual applicability'.¹² The literature on case studies as a research method includes a number of typologies of case studies, which are useful in thinking through the design and application of the method in this research on open government data. In 1995, Robert Stake identified three types of case study: intrinsic case studies, which study a particular site, situation or entity to understand that individual case; instrumental case studies, in which a single case is used to illuminate a wider phenomenon; and collective case studies, which study multiple cases to understand a phenomenon.¹³ The overall approach to the cases in this research is collective, but each study is treated separately and then compared, so that each case is really an instrumental case study as defined by Stake: it is a single case used to understand the wider phenomenon of record-keeping controls in data curation. This research design diverges from Stake's thinking about case studies, though, in that Stake saw case studies as analysing situations without looking for evidence for or against preconceived theories. As described above, this research identified record-keeping controls and looks for their presence or absence in the data curation practices identified through the case studies. In this respect, case studies are used in this research in line with Robert Yin's thinking about case studies. Yin argues that case study design should involve the selection of cases that will corroborate or challenge a theoretical proposition.¹⁴ This is an important feature of the case study component of the research design here, because the thesis seeks evidence that record-keeping controls for informational trustworthiness have some relationship with the trustworthiness of open government data. The choice to include two case studies was informed by what Charles Ragin has called 'focused comparison', which suggests that relevant

¹² Pickard, A., *Research Methods in Information* (London: Facet Publishing, 2007) p.93.

¹³ Stake, R., *The art of case study research* (Thousand Oaks: Sage Publications, 1995).

¹⁴ Yin, R., *Case study research: Design and methods* (5th ed.). (Thousand Oaks: Sage Publications, 2014).

cases should be selected according to the presence of the object of study to an extent that allows a close reading of the details of cases.¹⁵ The two case studies follow the rubrics of Alexander George and Andrew Bennett's method of 'structured, focused comparison', in which the objects of study are identified within the cases in advance – in contrast to Stake's observational approach – and in which each case is structured in such a way that direct comparisons are possible across cases, even as they may accumulate in future.¹⁶ In this research, the case studies are focused on the three record-keeping controls of metadata, custodianship and systems, and they are structured as described below.

Head, writing about the historiography of archives, states that 'Explicit comparison in historical research most often takes the form of case studies, which preserve the historical discipline's concern for context and particularity while allowing comparisons among properly contextualised phenomena across cases'.¹⁷ Although the primary focus of my research is not historical, the case studies present open government data in historical context and are, in their way, a kind of historiography of records, since they describe the technical details of accumulations of information. Each case study begins with an historical overview of the development from government secrecy towards openness before describing the relevant open government data portal as a recent focus of openness efforts. These were written following a review of sources including laws, government policies and strategies, government and non-governmental organisations' reports, websites and blogs. The case studies then look at particular datasets available through the portals: in both cases these concern land use. Land use data was selected because land management is one of the primary objectives and constituent parts of the modern state and an important nexus of citizen/state interactions. Contextual information on land information management is provided in the case studies, again, written following an analysis of published literature and official documentation, before the studies turn to an analysis of the curation of the sample land use dataset. The case studies map

¹⁵ Ragin, C., *The Comparative Method: Moving beyond Qualitative and Quantitative Strategies*. (Berkeley: University of California Press, 1987).

¹⁶ George, A. & Bennett, A., *Case Studies and Theory Development in the Social Sciences*, (Cambridge: MIT Press, 2005).

¹⁷ Head, 'A Comparative Case-Study Approach to Historical Archives in Europe', p.439.

the journeys of the data and identify strengths and weaknesses in the controls for the protection of the data's trustworthiness. This analysis began with an analysis of the datasets' metadata and used this metadata analysis to track the life of the dataset, identifying blank spots left by partial metadata.

The case studies have been restricted to Commonwealth countries to facilitate a closer comparison, given that the countries share legal and administrative traditions. This is particularly important in the study of the trustworthiness of government information because the legal tradition shared by Commonwealth countries is the same from which archival-diplomatics draws heavily in its conceptualisation of authenticity and reliability, as constituent aspects of trustworthiness, and the administrative tradition manifests in similar practices around record-keeping. Kenya and Australia were chosen as sites for case studies because, as Commonwealth countries, they share the legal and administrative traditions that provide the context for information management practices. In particular, their colonial legacy includes cultures of government secrecy, record-keeping through formal registries, and common law understandings of records as evidence. These aspects of their governments are important to this study as they relate to openness, information management and concepts of trustworthiness, allowing comparisons that focus instead on data curation and the extent of record-keeping controls. The two case studies concern national open data initiatives in Kenya and Australia, which have been leaders in information governance in some capacity – Kenya, because its open data initiative was the first open government data initiative in sub-Saharan Africa and was seen, at one time, as the leading open data platform in the region, and Australia because of the advanced state of digital record-keeping in that federal government.¹⁸ A comparison of data released through Kenyan and Australian portals therefore allowed some insights into pre-existing information governance

¹⁸ The perception of Kenya as a regional leader in open data is discussed in World Wide Web Foundation, *Open Data Barometer, Third Edition, Regional Report, Africa*, (May 2016) p.6. <http://opendatabarometer.org/doc/3rdEdition/ODB-3rdEdition-AfricaReport.pdf> [accessed 25 May 2018]. Australia's high profile contributions to the field of record-keeping in the later part of the twentieth century and into the twenty-first century is evident in the literature but particular examples include the work of the Records Continuum Research Group (RCRG), based in the Monash University Faculty of Information Technology and standards-setting work, which has been behind many of the international standards for records management.

infrastructures and controls that affected the trustworthiness of the available data. The two sites were also chosen because they have open government portals in relatively advanced states of development, having started in the early days of the movement towards open government data. However, these case studies are examples, and archival-diplomatic analyses could be usefully applied to opened datasets in many different cultural, juridical and administrative contexts.

It is important to note that at several places in this thesis, I draw on my own knowledge of events that I witnessed or participated in, but which are not documented in other sources. These are: 1) the work between the (UK) National Archives and the International Records Management Trust on the UK government's Open Government Partnership National Action Plan (2013-2015) commitment on records management. I was the International Records Management Trust representative for that collaboration during my employment with that organisation. 2) Some information about the records management practices and open data landscape in Kenya was collected while I was consulting on a project related to archival policy and unrelated to the research for this thesis. Those staff sources have been acknowledged in this thesis, but they were not interviewed for this research. 3) In the Kenyan case study, I reference work that I did with Justus Wamukoya for the International Records Management Trust in 2010-2011, which means that the detail in the discussion of records management practices is slightly richer in the Kenyan case study than in the Australian case study, for which there is no comparable data from that time.

A comparative approach is again taken to the analysis of the findings of the case studies, which helps to illustrate the significance of the presence or absence of record-keeping controls in data curation. Finally, the findings of this comparative analysis are brought into conversation with the arguments around information in the social contract to underscore the significance of the findings for information asymmetry in socio-political settings.

1.5. Limitations

This thesis is concerned with the trustworthiness of information, and though it attempts to show that trustworthiness in open government data can be improved, it recognises that trustworthiness is always contingent on perspectives and circumstances that preclude claims of universality in any definition of trustworthiness and its component parts. As Heather MacNeil observed, ‘...we should acknowledge that the methods for assessing reliability and authenticity, and the generalizations on which they are built, are not essential or transcendent verities but human constructs that have been shaped within a particular historical and cultural context’.¹⁹

In addition to this limitation, there are also areas of work related to the topic of this thesis that will be excluded from the study because of the limitations of space. In particular, three endeavours are clearly related and warrant investigation in relation to the findings of this thesis. These are metadata ontologies for provenance, the possibilities of fuzzy hashing for assessing reliability as conceived of in archival-diplomatics, and experiments in what I have termed ‘bit-level diplomatics’ to authenticate digital images.

An ontology that standardises and operationalises the exchange of provenance metadata exists.²⁰ It was published in 2013 by the World Wide Web Consortium (W3C) as the PROV Ontology (PROV-O), and referenced in a broader set of recommendations in 2017, called *Data on the Web Best Practices*, that also included guidance on versioning and preservation, amongst other things.²¹ PROV-O is designed to effect the PROV Data Model (PROV-DM).²² It would be useful to see where, if and how PROV-DM and PROV-O are being used to provide provenance metadata for open government data, particularly if this is being done in human-readable forms that would allow users to understand more of the context of the data they are using. An awareness of the significance of

¹⁹ MacNeil, H., ‘Trusting Records in a Postmodern World’, *Archivaria*, 51 (2001) p.46.

²⁰ Lebo, T., Sahoo, S. & McGuinness, D., *PROV-O: The PROV Ontology*. World Wide Web Consortium (30 April 2013) <https://www.w3.org/TR/prov-o/> [last accessed 22 October 2018].

²¹ World Wide Web Consortium, *Data on the Web Best Practices*. <http://www.w3.org/TR/dwbp/> [last accessed 22 October 2018].

²² Moreau, L., Missier, P. (eds.), *PROV-DM: The PROV Data Model*. 30 April 2013, World Wide Web Consortium (30 April 2013). <http://www.w3.org/TR/2013/REC-prov-dm-20130430/> [last accessed 22 October 2018].

provenance is also evident in the Data Documentation Initiative (DDI), and there is a need to look closely at how far that initiative has integrated the principles from record-keeping and archival-diplomatics, and whether DDI and similar social science data standards can be effectively implemented in open government contexts.²³

Technical tools have been developed to assess the authenticity of digital records. For instance, the file profiling performed by DROID (Digital Record Object Identification) tests digital objects for identity, by generating a checksum that can be used to identify identical copies of a record, and integrity, if users regularly generate a checksum to check that the file is unchanged or uncorrupted.²⁴ According to the InterPARES ontology for trustworthiness, the reliability of records can be assured by the completeness of a record and the conformance of its making to predefined processes of record-making. In other words, it must be similar to other records of its type. DROID cannot test for reliability, since it cannot check for similarity, only for identicalness. Fuzzy hashing is a technique for identifying similar files - could fuzzy hashing therefore be used to assess those signs of reliability defined by archival-diplomatics? If the process of record-making dictates a certain form that can be used as an exemplar against which other records can be assessed for their conformance, and if that form allows users to assess the record's completeness, can completeness and similarities of form be detected through fuzzy hashing with enough accuracy to offer an assurance of reliability? This will depend very much on the process and form of the record and defining the similarity range that would indicate creation through the same process and/or its completeness.

On 18 April 2018, NBC News posted an article about the US Defense Advanced Research Projects Agency's (DARPA) experiments in detecting fake news.²⁵ The article reports a media forensics project called MediFor that has developed a tool

²³ Data Documentation Initiative <http://www.ddialliance.org> [last accessed 25 October 2018].

²⁴ The National Archives, *DROID*. <http://www.nationalarchives.gov.uk/information-management/manage-information/policy-process/digital-continuity/file-profiling-tool-droid/> [last accessed 2 March 2019].

²⁵ Barrett, M. & Ling Kent, J., 'Inside the government agency designing tech to fight fake news', NBC, (19 April 2018). <https://www.nbcnews.com/tech/tech-news/inside-government-agency-designing-tech-fight-fake-news-n865586> [last accessed 22 October 2018].

to 'identify where an image's statistics, called a JPEG dimple, differ from the rest of the photo'.²⁶ Another DARPA tool examines videos and 'detects light levels and the direction from which the light is coming, using arrows to point out the differences and prove that the original videos were shot at different times. The indicator turns red, marking a literal red flag when the content is suspicious'.²⁷ These tools are expected to be in their final stages in 2020. These advances represent a kind of 'bit-level diplomatics', where forgery, annotation and manipulation are detected through an inspection of the digital object and its construction. In the same way that contradictory charters are examined for their authenticity using diplomatic principles, these tools should allow true copies of digital records to be identified. DARPA's work seems to have focused on audiovisual materials, but can the media forensic techniques now being developed be applied to other forms of records and information, and are there potential applications to open government data?

These three areas of work intersect with various aspects of the findings of this thesis and warrant further exploration, but their examination is out of scope here, where the focus is on open government data curation practice.

1.6. Definitions

The following terms recur throughout this thesis and therefore warrant definition from the outset.

'Archival-diplomatics' refers to a body of concepts that developed around the importation of traditional, formal diplomatics into archival studies in the 1990s. The Oxford English Dictionary gives Webster's Dictionary's (1828) definition of diplomatics as the '...science of diplomas, or of ancient writings, literary and public documents, letters, decrees, charters, codicils, etc., which has for its object to decipher old writings, to ascertain their authenticity, their date, signatures,

²⁶ Ibid.

²⁷ Ibid.

etc'.²⁸ Concepts from diplomatics were introduced into studies of digital records in the field of archival studies in the 1990s through scholars such as Luciana Duranti and Heather MacNeil and projects such as UBC-MAS (named after the University of British Columbia's Masters of Archival Science programme) and InterPARES, as described in Chapter 5. The term 'archival-diplomatics', sometimes referred to as 'digital diplomatics', distinguishes this more modern application of diplomatics to digital records from traditional diplomatics.

Information asymmetry refers to an imbalance in the quantity or quality of information accessible by two or more parties. In this thesis, the interest is in information asymmetry between states and individuals. A full explanation of the term and its development and use is provided in Chapter 3.

Open data is data that conforms to the Open Definition, developed and promoted by the Open Knowledge Forum, a public online forum maintained by the international non-government organisation the Open Knowledge Foundation.²⁹ The Open Definition states that 'open' as used in 'open data' refers to data that '...anyone can freely access, use, modify, and share for any purpose'.³⁰

The Open Knowledge Forum presents a definition of 'open government data' as an ancillary definition, and one that is still in development. Currently, the definition states that

Open government data and content is material that is:

- "Open" as defined by this site's Open Definition— in essence material (data) is open if it can be freely used, reused and redistributed by anyone.
- Produced or commissioned by government or government controlled entities.³¹

²⁸ Oxford English Dictionary, 'Diplomatics' (n.d.). <http://www.oed.com/view/Entry/53206?redirectedFrom=diplomatics#eid6620476> [last accessed 4 December 2018].

²⁹ *The Open Definition*. <https://opendefinition.org> [last accessed 2 March 2019].

³⁰ *Ibid.*

³¹ The Open Definition, 'open government data' (n.d.). <http://opendefinition.org/government/> [last accessed 2 March 2019].

The term 'social contract' is used in this thesis to denote a communal life characterised by shared submission to authority and conformance to social norms. Bypassing debates in political philosophy over the nature and value of forms of social contracts, I use the term as a generic descriptor for individual / state relations that are, as I argue in Chapter 2, binding.

The connected concepts of trustworthiness, accuracy, reliability, authenticity, identity and integrity are defined and explained fully in Chapter 5 using definitions created by the InterPARES research project, but, for reference, the trustworthiness of a record is the 'accuracy, reliability and authenticity of a record'.³² Accuracy is 'the degree to which data, information, documents or records are precise, correct, truthful, free of error or distortion, or pertinent to the matter'.³³ Reliability is the 'trustworthiness of a record as a statement of fact' and authenticity is the 'quality of a record that is what it purports to be and that is free from tampering or corruption'.³⁴ Authenticity depends on integrity, which is the 'wholeness and soundness' of a record and identity, which is 'the distinguishing character of a record, that is, the attributes of a record that uniquely characterize it and distinguish it from other records'.³⁵

1.7. Thesis Outline

Following this chapter on research questions and methods, I will argue, in Chapter 2, *Information and the Social Contract*, that records are the means of enacting the dialogical power relation of the social contract. I will do this by establishing that the social contract is characterised by a cycle of command and compliance, and that these are performative utterances, according to John Searle's classifications of illocutionary speech acts; in other words, they are

³² InterPARES 2, *Terminology Database*, http://www.interpares.org/ip2/ip2_terminology_db.cfm [last accessed 15 March 2019].

³³ Ibid.

³⁴ Ibid.

³⁵ InterPARES 1 Authenticity Task Force, 'Appendix 2 - Requirements for Assessing and Maintaining the Authenticity of Electronic Records', in InterPARES 1 Project, *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project* (InterPARES Project, 2002) pp.1-2.

speech acts that do things or give effect to things in the world. I will then argue that since the modern period the speech acts of government have increasingly been written in the form of records, and I use Cornelia Vismann's work to explicate 'records-as-commands'; records as the written instantiation of the performative utterance. Records enable power to operate at a distance, and I show how formal record-making and record-keeping bind individual subjects to states and their bureaucracies, and limit or control the parameters of individual/state dialogues. I then draw on record-keeping literature to explore how these power dynamics are represented or embedded in the archive that is constituted out of these records. Archives – storehouses of information – become the databanks of the state, as seen in the information acquisition project of colonialism. I suggest that an essential characteristic of this information is authenticity, since bureaucracies depend on the information in records for much of their activity. I then explore how bureaucracies enact the creation and transmission of records, arguing that bureaucracies act as apparatuses, or in Foucault's terms, *dispositifs*. These assemblages are obscure and often secretive, supersubjects comprised of subjects, which are conduits for transmitting and receiving information. Finally, I consider the power enacted by bureaucracies and their records during what Deleuze called a transition from a society of discipline to a society of control. Under this transition, record-making and record-keeping, and how individuals are recorded and datafied, is changing. This leads me to build on the 'record-as-command' concept and argue for the notion of 'data-as-command'. I argue that Jacques Ranciere's 'dissensus' and 'impertinent dialectic' are viable responses to the changing sociopolitical environment, but responses that will necessarily have to engage with records-and data-as-commands.

This provides a theoretical context for the third chapter, *Information Asymmetry in the Social Contract: A Principal Agent Problem*. Agency relationships have been theorised in the social sciences and I use this theory to frame the social contract as a principal-agent problem. The principal-agent problem model concerns the delegation of agency by a principal, and the mismatch between the goals of the principal and the agent. I use Mitnick's typology of agent behaviours to show how an agent's 'self-goals' can diverge from the goals of the principal, before

analysing the canonical form of the model described by Miller. A brief literature review then demonstrates the variety of the model's applications to social and political subjects, and the insights gained by applications that have relaxed one or more of the six assumptions in the canonical model. I work through the six assumptions in relation to the social contract to show that three of the assumptions are applicable, but that the information asymmetry assumption is the most significant for understanding the social contract because the contract is enacted through the exchange of information, as established in Chapter 2. I describe the development of a theory of information asymmetry and the emergence of the field of information economics, which radically changed economics by showing the powerful results of the operation of information in the economy. I describe the problems of information asymmetry as conceptualised in the economics literature. These are categorised as selection problems, or the problem of adverse selection, and incentive problems, or moral hazard. The literature posits signaling and screening as solutions to selection problems and metering and monitoring as solutions to incentive problems. After describing these approaches to solving information asymmetry problems, I argue that an incentive problem exists in the social contract, and that monitoring, rather than metering, is the appropriate approach to correcting that problem.

In the fourth chapter, I set the study in its historical context by providing an overview of the history of monitoring in the social contract. This includes a discussion of the culture of secrecy that has been observed in the British civil service and manifested in Official Secrets legislation, the opening of government information through archival access, the introduction of whistleblowers protection legislation, Freedom of Information legislation, and the growth of the open data movement, particularly as driven forward by the Open Government Partnership. This helps to establish the significance of this study of open government data for understanding information asymmetry in the social contract. The analysis finds the tensions of the social contract mirrored in the tensions around access to government information, strengthening the claim that the social contract is given effect by records- and data-as-commands, and the argument that dissensus and impertinent dialectics implicate information control and use.

The research looks to archival-diplomatics for a language and ontology for understanding trustworthiness in information (authenticity and reliability), which is elucidated in the fifth chapter, *Trustworthy Information*. Here, I detail the introduction of diplomatic principles into the study of digital records to explain the development of archival-diplomatic conceptualisations of informational trustworthiness, and I condense this work down into three broad controls that support assurances of trustworthiness. These are 1) custodianship documented in 2) metadata within 3) auditable systems.

In Chapter 6, *Practices for Establishing and Maintaining Authenticity in Information*, I show how these three broad controls have been given effect in record-keeping practices. The chapter details traditional and digital registry practices to understand how record-keeping systems have supported presumptions of authenticity. In Chapters 7 and 8, I present two case studies of national open data regimes – Kenya and Australia - in order to establish the presence or absence of these three controls for authenticity, particularly in relation to sample datasets. In Chapter 9, *Applying Record-Keeping Controls to Government Data Curation*, I present an analysis of the findings using a comparative approach, and in Chapter 10 I relate the analysis to the broader context elaborated earlier in the thesis to show the significance of the findings for information asymmetry in the social contract. The thesis concludes that the introduction of record-keeping controls into data curation practices would improve the trustworthiness of open government data, that this would positively affect information asymmetry in the social contract, but that publicity in data curation and critical data literacy on the part of data users are essential elements of a corrected or correcting – perhaps dissensual and perhaps impertinent - principal-agent relationship.

Chapter 2

Information and the Social Contract

First, the cord of the *harpedonaptēs* marks out a field and with its flexibility surrounds it: can anything be defined without it? To this object, second, it attaches the subject, as if to its knowledge or to its property. And third, it informs others, contractually, of the situation produced by the enclosure: can there be collective forms of behaviour without this? These practices concern, respectively, form, energy, and information; they are, if you will, conceptual, material, and judicial; geometric, physical, and legal. Bonds of knowledge, of power, and of complexity. All in all, its triple tress links me to forms, to things, and to others, and thus initiates me into abstraction, the world, and society. Through its channel pass information, forces, and laws.³⁶

Michel Serres, *The Natural Contract*

³⁶ Serres, M., *The Natural Contract* (Ann Arbor: University of Michigan Press, 1995) p.107.

2.1. Introduction

In this chapter, I will elaborate the concept of the 'record as command'. Drawing on studies of bureaucracy and governmentality, I will show that for political power to function, commands and responses to commands are necessary and that, in the state, this feedback loop is enacted in writing: this allows power to operate at a distance, documenting and controlling subjects, territories and events. Records, as written instantiations of performative utterances, manufacture, represent and inform authority. They are essential to regimes of disciplinary writing that enact power relations, and to be outside of this regime of documentation is to be outside of the regime of law. The state's bureaucracy, which is a writing apparatus, privileges what it has itself written down, which makes authenticity a key requirement of information employed in bureaucracies.

In order to understand how information moves back and forth between authority and the subject under the social contract, I will look at bureaucracies as a mechanism for giving effect to governance processes, identifying them as *dispositifs* with diverse and changing natures and constitutions, which can be open or closed, effective or moribund, etc. etc. I will use notions of control and valorizing information to show how interactions with bureaucracies help to sustain and empower them, before considering the role of information in the shift from a society of discipline to a society of control as conceptualised by Deleuze, and in the impertinent dialectic that Ranciere has proposed. These considerations contextualise Chapter 3, in which I discuss the information asymmetry that exists between states and their subjects.

2.2. Commands

'Mangia!'

('Eat!')

This command, given by the Duke (Paolo Bonacelli) to Renata (the unnamed victim played by Renata Moar) in Pier Paolo Pasolini's last film, *Salò; or the 120*

Days of Sodom, takes place in a discrete moment that crystallises the political analogy of Pasolini's masterpiece: it is a command to engage in coprophagia.³⁷ In the instant of the Duke's command, there is nothing outside of the command – a pronouncement of law; an expression of *imperium*. Renata seems to have no recourse beyond the supremacy of the *signori* who, as in Sade's *Les 120 journées de Sodome ou l'école du libertinage*, represent the 'establishment' (the Duke, Bishop, President (Magistrate) and Banker in Sade; the Duke, Bishop, President and Magistrate in Pasolini).³⁸ Resistance to the command would criminalise her in a regime of unbridled sovereign power.

The moment of the Duke's command is made discrete by the silences that flank the command. The cinematography uses long shots that put the spectator at a distance and works with the *mise-en-scène* - with the actors arranged at the edge of the room – to position the viewer as a spectator and therefore an accomplice in Renata's abuse.

To break out of the schema of projection-identification, Pasolini purposely limited the sympathy that the *victimmi* garnered, thus avoiding the easy catharsis that cleanses the palate after aesthetic or ethical trauma. If anything, even as the film tempts many critics to adopt the perspective of the censor, it also imposes upon the viewers the perspective of the *signori*.³⁹

This complicity on the part of the spectator takes on new meaning in view of the film's sociopolitical setting. In transposing Sade's scenario to the Repubblica Sociale Italiana, the final iteration of Mussolini's fascist state, Pasolini articulated fascism's inherent sadomasochism. Christopher Roberts has proposed that Pasolini meant to extend this analogy to the entire history of sociopolitical organisation, and film studies scholar Naomi Greene attests to this intention:

³⁷ Pasolini, P. P., *Salò o Le 120 giornate di Sodoma* (London: British Film Institute, 1975).

³⁸ Sade, D.A.F., *The 120 Days of Sodom* (London: Arrow Books, 1990).

³⁹ Roberts, C., 'The Theatrical Satanism of Self-Awareness Itself: Religion, Art and Anarchy in Pasolini's *Salò*', *Angelaki*, 15.1 (2010) pp.29–43.

The obsession with precise formulas and bureaucratic regulations that characterizes his libertines represented, [Pasolini] declared, the strategies embraced by all power in its pure arbitrariness, that is, its own anarchy.⁴⁰

Through this idea of power's own anarchy Pasolini invites the viewer to consider how power is constituted and maintained. In Renata's abuse, the audience witnesses how power is buttressed by the complicity of her fellow *victimmi*, who outnumber the *signori* but acquiesce in the face of the initially unspoken threat of violence. In extrapolation, consider Jean-Jacques Rousseau's observation, in 1762, that this relationship of authority, the social contract, is 'perhaps never... formally promulgated... everywhere tacitly acknowledged and received'.⁴¹ The body politic is governed and governable not through its explicit consent to a contract, but through its tacit acceptance of a form of government, its institutions and processes. The body politic surrenders power. It can be commanded.

If the command, a directive according to John Searle's classifications of illocutionary speech acts, is a performative utterance, if it changes reality, then answering back is also a performative utterance.⁴² Jussi Parikka has argued that '...all acts of language are acts of power: attempts to draw territories, define borders, impose operations. Order-words are assemblages of enunciation that mobilize and transform bodies and hence create new affects (actions and passions) for them'.⁴³ If we can accept Parikka's position that all acts of language are acts of power, then responses to commands are performative utterances. But how does the response to a command function as an act of power? Tracing the development of record-keeping in German government, Cornelia Vismann observed that '...feedback enabled the government to assume control and

⁴⁰ Greene, N., 'Salo: The Refusal to Consume' in Rumble, P. & Testa, B. (eds.) *Pier Paolo Pasolini: Contemporary Perspectives* (Toronto: University of Toronto Press, 1994) pp.232–42.

⁴¹ Rousseau, J.-J., *A Treatise on the Social Contract* (London: T. Becket and P.A. De Hondt, 1764).

⁴² Searle, J., 'A Taxonomy of Illocutionary Acts', *Language, Mind and Knowledge*, 7 (1975) pp.344–69.

⁴³ Parikka, J., *Digital Contagions: A Media Archaeology of Computer Viruses* (Oxford: Peter Lang Publishing, 2007) p.139.

guidance functions',⁴⁴ and JoAnne Yates observed the same dynamic in her study of the emergence of 'systematic management' in the American private sector; 'systems were established, operated, evaluated, and adjusted – that is to say, managed or controlled – all on the basis of flows of information and orders' and '...internal communication came to serve as a mechanism for managerial coordination and control of organisations'.⁴⁵ In the private sector, as in government, control required a return of information; it could not be exercised by a one-way flow of commands. This return of information is a form of the body politic's submission to power, a kind of call and response. Of the power dynamic in *Salò*, Gabrielle Lesperance wrote:

At the same time that there exists an uneven bureaucracy in *Salò*, the film's physical symmetry could be considered overwhelming. This degree of glacial perfection has both political and semiotic implications. While *Salò*'s symmetry gives the impression of complete dominance by man over his surroundings, it, in fact, reveals the limits of human control.⁴⁶

That is, control is constructed by the parties to the control relationship. The boundaries of power are defined by the assertion of the *signori* and the acquiescence of the *victimmi*. When individuals, together constituting the body politic, comply *en masse*, the tacit acceptance of power can be conceived as an expression of the 'general will'. Silence - refusing to answer - is not possible if the concept of the general will is to be sustained, and in this way the general will binds subjects together. The philosopher Michel Serres considers the exile of Anaxagoras of Clazomenae from Athens as an illustration of this function of the general will:

⁴⁴ Vismann, C., *Files: Law and Media Technology*, trans. by Winthrop-Young, G. (Stanford: Stanford University Press, 2008) p.95.

⁴⁵ Yates, J., *Control through Communication: The Rise of System in American Management* (Baltimore: Johns Hopkins University Press, 1993) pp.xvii-xv.

⁴⁶ Lesperance, G., 'Beginning to Think about *Salò*', in Iannucci, A. (ed.) *Dante, Cinema, and Television* (Toronto: University of Toronto Press, 2004), p.98.

By what right, then, does some citizen criticize Anaxagoras? By this fundamental legal right that founds the existence of the city and that is sometimes called the social contract. If, because you're observing the planets, you lose interest in your country, then you break the contract that unites us, and thus society must logically exclude you, condemn you at least to exile and at most to death.⁴⁷

Serres argues that in order to be general, the general will must be all-encompassing, and that when one opts out, one is in effect pronouncing one's own sentence: 'The contract,' he writes, 'in its logic, knows no mercy'.⁴⁸ For political philosopher Jacques Ranciere, the 'social bond is maintained by this endless manufacture of acquiescence'.⁴⁹ To do anything but acquiesce is to move out of juridical personhood and into a state of bare life. Command, compliance, refusal in silence, rejection of the general will; these are all operations of information exchange.

2.3. The Record as Command

Inasmuch as a command is performative according to speech act theory, recording a command is performative. Michel Foucault's notion of disciplinary writing - that individuals can be bound up in a 'network of writing' – constitutes a tool for the government of subjects.⁵⁰ A record is a written command. Vismann has elaborated a theory of the record where 'Every file note indirectly contains a command. Reporting the execution of an order triggers the next one'.⁵¹ In his study of bureaucracy in Pakistan, Matthew Hull wrote that the file '...is a technology for materially enacting an authoritative decision, for making a decision out of various utterances and actions'⁵² and Akhil Gupta's study of bureaucracy's structural violence in India foregrounds the performativity of writing, particularly in

⁴⁷ Serres, *The Natural Contract*, p.67.

⁴⁸ Ibid.

⁴⁹ Ranciere, J., *On the Shores of Politics*, (London; New York: Verso Books, 2007) p.83.

⁵⁰ Foucault, M., *Discipline and Punish: The Birth of the Prison*, trans by A. Sheridan, 2nd ed. (New York: Vintage Books, 1995) p.190.

⁵¹ Vismann, *Files*, p.8.

⁵² Hull, M., *Government of Paper: The Materiality of Bureaucracy in Urban Pakistan* (Berkeley: University of California Press, 2012) p.127.

the context of government: ‘...writing is itself a form of action and is perhaps the most important of bureaucratic activities’.⁵³ The writing down of the command, and other speech acts, also confers authority on the speech, as Hull suggested when he wrote that ‘the requests, complaints, decisions, understandings, permissions, evasions, and refusals of these conversations eventually must find their way onto the paper of files to have a life beyond talk’.⁵⁴ This idea that writing, and particularly bureaucratic writing, makes speech actionable in a substantive way was also expressed by Vismann when she wrote that ‘Within the imagined chain of replacements for the spoken language, *supplements* [after Derrida], files are the closest to the presence of speech’ and that files ‘compensate for the transitoriness of speech but without abandoning presence’.⁵⁵ These two references to ‘presence’ suggest the representational potency of the file. The file bears the marks of authority; logos, stamps, signatures and seals. It is through these symbols of power that power can be represented beyond its immediate presence, and it is no longer necessary for the institution or personification of power to be present at the site of the command’s execution. Jeffrey Sconce has observed that:

As many commentators on modernity have noted, in the era’s new administration of the social, “power” increasingly emanated from elsewhere – the capital, the courts, the university – institutions beyond the immediate experience and comprehension of the individual subject.⁵⁶

Documents were used to communicate directives to these subjects at a distance, but they were also increasingly used to identify individual subjects. A relationship developed between the physical body and its documentation, further fleshing out the juridical person found in traces across records of their activities and interactions with the state. In addition to the subject’s inscription in registers,

⁵³ Gupta, A., *Red Tape: Bureaucracy, Structural Violence, and Poverty in India* (Durham: Duke University Press, 2012) p.188.

⁵⁴ Hull, *Government of Paper*, p.113.

⁵⁵ Vismann, *Files*, pp.8-9.

⁵⁶ Sconce, J., ‘On the Origins of the Influencing Machine’ in Huhtamo, E. & Parikka, J. (eds.) *Media Archaeology: Approaches, Applications, and Implications* (Berkeley: University of California Press, 2011) p.82.

licenses, certificates and dossiers, documentation of the person in the form of identification documents allowed the individual to be a data point within a documentary regime that mirrored and enacted a government of physical entities (bodies, goods, records, territories). This relationship became symbiotic, and the physical body became readable in a formal sense: 'The body "documented" identity such that information was always co-present with the appearance of a person'.⁵⁷ Identification documents such as passports moved with physical persons, or as Craig Robertson puts it in his history of the American passport, passports and visas 'not only controlled the flow of information by stopping the movement of "knowledgeable" bodies, they also facilitated governing over distance by making knowledge mobile; the passport contained information for officials to read'.⁵⁸ These moves to codify the physical body in documents are the roots of biometrics, and today biometric technologies are perfecting the documentation and interpretation of irises, ears, gaits, typing styles, odours and voices. This is the datafication of the human body, and relates to the notion, developed in surveillance studies, of the 'data double', which Kevin Haggerty and Richard Ericson explain as a 'distinctively hybrid composition' that is 'broken down by being abstracted from its territorial setting. It is then reassembled in different settings through a series of data flows. The result is a decorporealized body, a 'data double' of pure virtuality'.⁵⁹ This data double appears in records as the object of control. In a sense, information is essential to constituting the juridical body – the person who exists before the law must first exist in documentation. As Judith Butler wrote: 'Constituted as a social phenomenon in the public sphere, my body is and is not mine. Given over from the start to the world of others, it bears their imprint, is formed within the crucible of social life; only later, if at all, do I lay claim to my body as my own, if, in fact, I ever do'.⁶⁰ In the modern state, the physical person is constituted as a juridical person through documentation created with and by others. Once the data double exists within a

⁵⁷ Robertson, C., *The Passport in America: The History of a Document* (Oxford: Oxford University Press, 2010) p.183.

⁵⁸ Ibid. p.188.

⁵⁹ Haggerty, K. & Ericson, R., 'The Surveillant Assemblage', *British Journal of Sociology*, 51:4, 2000, p.611.

⁶⁰ Butler, J., *Prekarious Life: The Powers of Mourning and Violence*, (London; New York: Verso Books, 2004) p.82.

documentary regime, a regime of disciplinary writing, a level of systemic control is established.

As power could be projected over time and space through official documentation, some forms of documentation carried authority with them: thinking about the verification of personal identities, Robertson notes that 'documents verified documents, and the authority of the state remained to some extent continuous...'.⁶¹ The interrelated processes and documents of government form a web or chain that theoretically returns eventually to authority. The file or record is a sign of an absent referent (state authority), removed in two senses: by physical space and by a sequence of processes. A similar notion of records standing in for authority is drawn out of Sigmund Freud's work by Jacques Derrida when he describes the 'topology of archives' as 'what ought to exclude or forbid the return to the origin',⁶² and drawing on the archaeological parables in Freud's 1896 study of hysteria:

...the very success of the dig must sign the effacement of the archivist: *the origin then speaks by itself*. The *arkhe* appears in the nude, without archive. It presents itself and comments on itself by itself. "Stones talk!" In the present. *Anamnesis* without *hypomnesis*! The archaeologist has succeeded in making the archive no longer serve any function. It *comes to efface itself*, it becomes transparent or unessential so as to let the origin present itself in person. Live, without mediation and without delay.⁶³

So, the archive stands in for the original thing; it is a representation of power. Records, archives, files, while supplements, are also constitutive of power. Files, Ilana Feldman says, 'can be authoritative and can impart authority to the government that does the filing without conferring legitimacy on that

⁶¹ Ibid. p.212.

⁶² Derrida, J., *Archive Fever: A Freudian Impression*. (Chicago: University of Chicago Press, 1998) p.92.

⁶³ Ibid.

government'.⁶⁴ Feldman argues that the functioning of a system of bureaucratic writing, form filling and filing create a sense of authority and bind subjects in that cycle of command and compliance required by power: 'Under conditions where government was tenuous and lacked a stable ground, it was the repetitions of filing procedures, the accumulation of documents, and the habits of civil servants that produced the conditions of possibility for authority'.⁶⁵ Feldman takes the notion of the archive standing in for power and suggests that, in the case of Gaza's bureaucracy:

...reiteration did not stand in for an original belief that had lost its power. To the extent that belief was invoked, people were asked not so much to believe that bureaucracy stood for something else as to believe in it for itself. Perhaps even more vital than such belief, though, they were asked to participate in its workings. It was, again, authority (of bureaucracy itself) rather than legitimacy (of the regime) that bureaucratic repetition promoted.⁶⁶

As well as representing power, then, these disciplinary writings can help to manufacture authority for regimes or at least, as Feldman has shown, for the machineries of government.

These records and files may eventually become archives of the state that also have roles in supporting power relations. Kirsten Weld has articulated the paradox of the state archive that appears neutral but is shaped through sequences of systems, practices and choices:

Any archive contains far less than it excludes, as archivists know, and every archive has its own history – one that conditions the ability to interact with it, write from it, and understand the larger systems of power, control, and legibility that record keeping necessarily enables. The Enlightenment notions undergirding the concept of state archives, as both *a part of* and

⁶⁴ Feldman, I., *Governing Gaza: Bureaucracy, Authority, and the Work of Rule, 1917–1967*, (Durham: Duke University Press, 2008) p.32.

⁶⁵ *Ibid*, p.3.

⁶⁶ *Ibid*, p.16.

apart from modern societies, represent these institutions as neutral storehouses of foundational documents. In practice, however, the politics of how archives are compiled, created, and opened are intimately tied to the politics and practices of governance, and are themselves historical in a way that transcends the content written on their documents' pages.⁶⁷

Ann Stoler, Verne Harris, Randall Jimerson and others have articulated the relationship between archives and power. Stoler, writing about reports from Dutch colonial administrators in Sumatra concerning a family murdered in 1876, notes that these 'storied reports were fashioned cultural accounts *with political effects* that precluded some conclusions and encouraged others'.⁶⁸ Harris argued that 'the archival record is at once expression and instrument of power'.⁶⁹ On the role of the archive in the power dynamics of apartheid South Africa, he wrote that by 'their silences and their narratives of power, their constructions of experience, apartheid's memory institutions legitimised apartheid rule'.⁷⁰ Jimerson's study of George Orwell's writings draws out some of the issues arising from the connection between archives and power. Jimerson writes:

... Orwell explored the danger that totalitarian leaders would entrench their power by distorting historical knowledge through control of written records and human memory... Authentic records – the very stuff of archives – provide one of the strongest bulwarks against totalitarianism.⁷¹

It is also the case that authentic records provide one of the strongest bulwarks for totalitarianism, as may be seen from the record-keeping programmes of numerous totalitarian regimes. In 1990, the imminent fall of the German Democratic Republic (GDR) prompted *Aktion Riesswolf* (Operation Shredder),

⁶⁷ Weld, K., *Paper Cadavers: The Archives of Dictatorship in Guatemala* (Durham: Duke University Press, 2014) p.13.

⁶⁸ Stoler, A. L., *Along the Archival Grain: Epistemic Anxieties and Colonial Common Sense* (Princeton: Princeton University Press, 2010) p.232.

⁶⁹ Harris, V., 'The Archival Sliver: Power, Memory, and Archives in South Africa,' *Archival Science*, 2.1 (2002) p.84.

⁷⁰ *Ibid.*, p.69.

⁷¹ Jimerson, R. C., *Archives Power: Memory, Accountability, and Social Justice*. (Chicago: American Library Association, 2009) p.142.

which would ultimately fail in its aim to destroy the regime's records, which were proof of the use of information to monitor and control the population.⁷² Paul Lihoma has described the official interest in the National Archives of Malawi under Banda's one-party regime between 1964 and 1994, over which period obtaining permission to use the archives became increasingly demanding, eventually requiring Banda's personal permission, a situation arising from his desire to censor writing that drew on sources in the state archives.⁷³ In both cases, the regimes demonstrated a desire for total knowledge, which aided monitoring and control. Knowledge requires the ongoing acquisition of information, or as Lisa Gitelman has written: 'Filling in and filing away are the ways that bureaucracies collect and connect; like the micrologics of enclosure and attachment, they are part of a repertoire of techniques through which bureaucracies come to know'.⁷⁴ Thomas Richards has described the connection between government and knowledge acquisition in the context of British imperialism:

Much Victorian thought participated in seeing the state as central to human life, and more, in imagining a kind of complete documentary knowledge of human life that would exist solely for the state. The threat uniting the thought of Bentham and Mill with the thought of Russell and Keynes and C.P. Snow was the idea that knowledge is inconceivable without the state: that the question of the state is a question of knowledge, especially scientific knowledge; that the classing of knowledge must be underwritten and directed by the state in its various capacities; that all epistemology became and must remain state epistemology in an economy of controlled information.⁷⁵

⁷² Miller, J., 'Settling Accounts with a Secret Police: The German Law on the Stasi Records' *Europe-Asia Studies*, 50.2 (1998) pp.305–330.

⁷³ Lihoma, P., 'The Influence of the One Party Regime on Archives in Malawi: 1964 to 1994' in Lowry, J. & Wamukoya, J. (eds.), *Integrity in Government through Records Management* (Farnham: Ashgate, 2014) pp.85-98.

⁷⁴ Gitelman, L., *Paper Knowledge: Toward a Media History of Documents*, (Durham: Duke University Press, 2014) p.32.

⁷⁵ Richards, T., *The Imperial Archive: Knowledge and the Fantasy of Empire* (London; New York: Verso, 1993) p.74.

Imperial governance is only possible through the acquisition of information – the ongoing transmission of data across physical territories. The materiality of records is intertwined with the materiality of the state. Jack Goody made this clear when he wrote that in ‘the vast majority of regimes, ancient or modern, whether capitalist or socialist, the postal service is organized by, and in an important sense on behalf of, the state, carrying its paperwork by priority and without payment. Right at the beginning such a system involved the upkeep of roads and canals so that deliveries could be effected in reasonable time.’⁷⁶

Today, this connection between the physical infrastructure of the state and its ability to transmit commands in writing has a digital counterpart that arguably extends the reach of authority, but records, whether made on physical or digital carriers, operate as commands only insofar as they bear the marks of authority.

In order for records to perform their role in this power relationship, records must be what they purport to be – they must be authentic. Establishing the authenticity of records would be the purpose of a new science that emerged in Europe following the publication of Mabillon’s *De re diplomatica*.⁷⁷ This science, ‘diplomats’, would develop principles for the analysis of the relationship of the record to the authority that was said to have created it, through the analysis of the record’s marks of authority. These were to be found in the record’s extrinsic elements, which ‘constitute the first and most obvious proof of their authenticity’.⁷⁸ Such elements include: medium - Matfei and McAndrew cite as an example a royal diploma of Childebert I (King of Franks, VIth century), which, if written on parchment instead of papyrus, would indicate inauthenticity⁷⁹; script and seals - ‘The characteristic of the public seal is to be *‘bene cognitum et famosum’* (well known and famous) in order to be recognised by all subjects’⁸⁰; and additional notes. These marks of authority also included the record’s intrinsic

⁷⁶ Goody, J., *The Logic of Writing and the Organization of Society* (Cambridge: Cambridge University Press, 1986) p.96.

⁷⁷ Mabillon, J., *De re diplomatica* (1681).

⁷⁸ Matfei, M. & McAndrew, I., *Lineage of Elements Included in the Template for Analysis (pre-InterPARES): From Traditional Diplomats to Contemporary Archival Diplomats* (University of British Columbia, 2000) p.3.

http://www.interpares.org/display_file.cfm?doc=ip1_lineage_of_elements.pdf [last accessed 25 April 2018].

⁷⁹ Ibid.

⁸⁰ Ibid., p.4

elements, the content and composition of the document, including protocol (invocation, superscription, inscription, salutation), text (preamble, notification, exposition, disposition, final clauses), eschatocol (date, appreciation, subscription) and language. In diplomatics, each element is compared with its norm to determine if it is a genuine mark of authority. In a sense, the official record is a visual, material manifestation of the authority that produces it; the *supplement* closest to speech. Through diplomatics, which seeks to establish if a record has originated from a particular chancery (bureaucracy), the legitimacy of the record can be established. The legitimacy of the record confers legitimacy on the command that it contains.

As Jimerson alludes to when he says ‘the very stuff of archives’, establishing and maintaining authenticity in records is the province of archival science. Vismann observes that the law ‘operates not *in mundo* but in the medium of literality; it believes only what is written – more precisely, what it has itself written down’.⁸¹ Ensuring that records are indeed what the law ‘has itself written down’ is the purview of record-keeping. This is the function that the archive performs for the law; it not only houses the records, but it safeguards their authenticity so that they can serve as evidence that can be relied upon by the bureaucracy. In this way, records become, as archival theorists Joan Schwartz and Terry Cook said, ‘value-laden instruments of power’.⁸² Perceived authenticity is essential in conferring authority on records. Authenticity is the crucial characteristic of records that make them tools of power. Who is doing this writing and authenticating? If Vismann is right that a letter ‘is nothing but a transmittable order,’ how is the order, the command, transmitted from the site of power, and how is proof of its execution returned? It is through the apparatus of bureaucracy.

2.4. Bureaucracy as Apparatus

Bureaucracy is the apparatus through which the state exercises power, and its materiel is information. Gregg Hetherington’s study of ‘guerilla auditing’ – the use

⁸¹ Vismann, *Files*, p.56.

⁸² Cook, T. & Schwartz, J. M., ‘Archives, records, and power: From (postmodern) theory to (archival) performance,’ *Archival Science*, 2.3 (2002) p.178.

of records and archives by Paraguayan tenant farmers to defend their rights – showed that ‘at its core, the state functions through the creation, circulation, and endless interpretation of documents’ and that for the farmers, ‘documents do not store information so much as make it possible, and state power is therefore not about seeing or inscribing so much as it is about controlling who reads what and under what circumstances’.⁸³ Though Hetherington downplays the importance of inscription somewhat in this statement, his suggestion that documents make information possible underscores the generative quality of bureaucratic writing. Hull posits that writing is central to the bureaucratic project: ‘From the official bureaucratic point of view, a person who is an agent but not an author, who causes things to happen without writing or being written about, is improper at best, corrupt at worst.’⁸⁴ It is the role of the bureaucrat to write. The state, according to Vismann:

...acts through its civil servants and thus ideally through every citizen. As if it were a supersubject composed of subjects, the state is credited with all its administrative operations and utterances. Once civil servants embody the state, there is no administrative act, no matter how trivial, that does not carry the sovereignty of the entire state with it.⁸⁵

What Vismann describes here is an amorphous entity; beyond the ‘supersubject composed of subjects’, she alludes to administrative practices that evoke a whole set of contingent systems, actors and actions, a *dispositif*, after Foucault, ‘a thoroughly heterogeneous ensemble consisting of discourses, institutions, architectural forms, regulatory decisions, laws, administrative measures, scientific statements, philosophical, moral and philanthropic propositions... The apparatus itself is the system of relations that can be established between these elements’.⁸⁶ This complex apparatus is a writing machine. Akhil Gupta noted that

⁸³ Hetherington, K., *Guerrilla Auditors: The Politics of Transparency in Neoliberal Paraguay* (Durham: Duke University Press, 2011) p.151.

⁸⁴ Hull, *Government of Paper*, p.130.

⁸⁵ Vismann, *Files*, p.110.

⁸⁶ Foucault, M., *Power/Knowledge: Selected Interviews and Other Writings, 1972-1977*, trans. by Gordon, C., Marshall, L., Mepham, J. & Soper, K. (New York: Pantheon Books, 1980) p.194.

one of the 'most striking aspects about references typically made of the state, apart from its purposive nature, is the unitary character of the object being described', and called for closer consideration of the forms and processes of state machinery.⁸⁷ Gupta is right that the state is not unitary, that it is, as Vismann says, a supersubject, and one that incorporates, quite literally, subjects beyond the bureaucrat. The *dispositif* that is bureaucracy operates 'ideally through every citizen',⁸⁸ Vismann says, and according to Matteo Pasquinelli, 'bureaucracy descends into the bodies of the workers' so that it is difficult to establish the boundaries of bureaucracy. Foucault wrote that it is a 'system of relations' rather than a discreet body. Jensen and Meckling's theory of the firm similarly saw organisations as 'legal fictions which serve as a nexus for a set of contracting relationships among individuals'.⁸⁹ Presaging Vismann's idea of bureaucracies as supersubjects that might even envelop citizens, Jensen and Meckling argued that:

... it makes little or no sense to try to distinguish those things that are "inside" the firm (or any other organization) from those things that are "outside" of it. There is in a very real sense only a multitude of complex relationships (i.e., contracts) between the legal fiction (the firm) and the owners of labor, material and capital inputs and the consumers of output.⁹⁰

The subject's engagement with a particular office or service is the means through which they are entangled in a relationship with bureaucracy, becoming part of the system of relations. Gupta has observed this characteristic: 'Leaving aside extraordinary events, the boundary between state and society may actually be constructed through the everyday practices of state offices and representations created by officials'.⁹¹ His examination of the presence of structural violence in a public sector responsible for a large portfolio of care programmes led him to

⁸⁷ Gupta, *Red Tape*, p.44.

⁸⁸ Vismann, *Files*, p.110.

⁸⁹ Jensen, M. C. & Meckling, W. H., 'Theory of the firm: Managerial behavior, agency costs and ownership structure,' *Journal of Financial Economics*, 3.4 (1976) p.310.

⁹⁰ *Ibid.*, p.311.

⁹¹ Gupta, *Red Tape*, p.55.

argue 'that structural violence is enacted through the everyday practices of bureaucracies, and one therefore needs to look closely at those everyday practices in order to understand why violence coexists with care...'.⁹² These everyday practices, engagements and transactions with subjects, establish or reinforce a tacit contract, so that each person, though never signing a 'social contract', submits to government by 'signing' (tacitly or actually) any number of minor contracts with the bureaucracy. This assigns power to bureaucrats in a way that Judith Butler shows to be problematic for those excluded from the social contract. Writing about the Guantanamo Bay detention camps, Butler says

Those who decide on whether someone will be detained, and continue to be detained, are government officials, not elected ones, and not members of the judiciary. They are, rather, part of the apparatus of governmentality; their decision, the power they wield to "deem" someone dangerous and constitute them effectively as such, is a sovereign power, a ghostly and forceful resurgence of sovereignty in the midst of governmentality.⁹³

Sovereign power, then, is diffused through bureaucracies. When the boundaries of bureaucracy are porous, and individuals move into positions of power and back out as subjects, often constituted as both, where are the limits of power, and how should the subject know how to navigate them? Marx saw bureaucracy as a 'particular closed society within the state'⁹⁴ which extends its power through secrecy and mystery, and this opaque character of bureaucracy is brought out in Vismann's analysis of Kafka's *The Trial*, in which the central character, Joseph K., is arrested for an unidentified crime and prosecuted by an unnamed organisation. In reference to K.'s difficult interaction with the bureaucracy of this obscure judicial system, Vismann says that there is 'no place from which the entire architecture of barriers can be grasped, none that renders comprehensible the master plan that appears to control everything.'⁹⁵ Gupta has observed the same trait in Indian bureaucracy: 'There is no Archimedean point from which to

⁹² Ibid., p.33.

⁹³ Butler, *Precarious Life*, p.59.

⁹⁴ Marx, K., *Critique of Hegel's "Philosophy Of Right."*, trans by O'Malley, J. (Cambridge: Cambridge University Press, 1977) p.46.

⁹⁵ Vismann, *Files*, p.18.

visualize the state, only numerous situated knowledges⁹⁶ and that ‘...the materiality and solidity of the state dissolve under scrutiny. The state consists of congeries of institutions with diversified levels, agencies or bureaus, agenda, functions, and locations’.⁹⁷

The idea of an incomprehensible bureaucracy is illustrated in the theoretical state of the Neue Slowenische Kunst (NSK) art collective, where organisation is based on Yugoslavian self-management structures.⁹⁸ The development of the complexity of the NSK organogram, from a group of three founding units (Laibach (music), IRWIN (visual arts), and the Theatre of the Sisters of Scipion Nasice (theatre), to more complex forms, parodies the growth of the Yugoslavian self-management form of government into an unwieldy and impenetrable network of organs, by what Alexei Monroe called an ‘extreme process of institutional proliferation’.⁹⁹ The NSK organogram grew more complex between its 1982 and 1984 iterations. Monroe wrote that it ‘took the process of alternative institutionalization to its (il)logical formal extreme, recapitulating and attempting to transcend the institutional anarchy of the period and the fantastically complex, deliberately opaque web of state and parastate organisations within the late Yugoslav system’.¹⁰⁰ On this ‘deliberately opaque web’, Monroe cites Mark Thompson’s argument that the ‘self-management system was expressly structured in order to conceal and diffuse the bases of power’, giving rise of a Kafkaesque, labyrinthine bureaucratic apparatus.¹⁰¹

Deleuze and Guattari found in Kafka’s work revelations of the secret architectures of power and bureaucracy as the complex apparatus of power is present across Kafka’s work.¹⁰² Examining Kafka’s *The Castle*, Reza Banakar wrote:

⁹⁶ Gupta, *Red Tape*, p.104.

⁹⁷ *Ibid.*, p.45.

⁹⁸ Zukin, S., *Beyond Marx and Tito: Theory and Practice in Yugoslav Socialism* (Cambridge: Cambridge University Press, 1975).

⁹⁹ Monroe, A., *Interrogation Machine: Laibach and NSK*. (Cambridge; London: MIT Press, 2005) p.25.

¹⁰⁰ *Ibid.*, p.107.

¹⁰¹ *Ibid.*, p.78.

¹⁰² Deleuze, G. & Guattari, F., *Kafka Toward a Minor Literature*, 9th ed. (Minneapolis: University of Minnesota Press, 1986).

... K.'s desire for justice should be seen above all as another double thought, which captures the tension between the individual and his/her community: to satisfy his desire for justice, K. must gain an audience with the Castle authorities and convince them to recognize his particular situation and circumstances. The only medium of action available to K., however, is the Castle's impersonal bureaucracy, which K. must access through Momus's deposition. K. must thus submit to procedures that establish their objectivity precisely by disregarding the singularity of his position.¹⁰³

This engagement with bureaucracy is significant. Vismann observed that in 'sixteenth-century German cities, outgoing and incoming writs were collected in such a way as to make up a file. Accordingly, the monological rule changed and became – for all its asymmetry – dialogical.'¹⁰⁴ Hull expressed this idea when he wrote that 'Governing paper is central to governing the city. And paper is also the means by which residents acquiesce to, contest, or use this governance.'¹⁰⁵ This dialogical rule is the written rendering of the command and compliance embodied by the coprophagia at Salo.

Srikant Sarangi and Stefaan Slembrouck assert that bureaucracy depends on information exchange because the management of people is done by proxy through information about them, and that interactions between citizens and public bureaucracy is dependent upon processes of information seeking and exchange.¹⁰⁶ Pasquinelli sees an interdependency between bureaucracy and information in Romano Alquati's study of the 'organic composition of capital at the Olivetti factory', which produced typewriters and mainframe computers.

Pasquinelli writes:

¹⁰³ Banakar, R., 'In Search of Heimat: A Note on Franz Kafka's Concept of Law,' *Law & Literature*, 22 (2010) p.479.

¹⁰⁴ Vismann, *Files*, p.95.

¹⁰⁵ Hull, *Government of Paper*, p.1.

¹⁰⁶ Sarangi, S. & Slembrouck, S., *Language, Bureaucracy, and Social Control* (New York: Routledge, 2014).

Alquati first encountered the cybernetic apparatus (what today could be easily identified with a digital network) as an extension of the internal bureaucracy of the factory, which monitors the production process by means of *control information*. Bureaucracy descends into the bodies of the workers via the mediation of the circuits of cybernetics and machinery. In this context, Alquati introduces the concept of *valorizing information*, the flow running upstream and feeding the circuits of the whole factory. Such valorizing information is continuously absorbed by machinery and finally condenses into products.¹⁰⁷

Pasquinelli identifies in Alquati's work the two-way flow of information ('control information' and 'valorizing information') that allows the bureaucracy of the factory to function:

The internal bureaucracy of the factory is a specific division of labor that is mirrored, implemented, and extended by cybernetics. In fact, the important insight advanced by Alquati is that of an organic continuum merging bureaucracy, cybernetics, machinery, and the division of labor: cybernetics unveils the machinic nature of bureaucracy and conversely the bureaucratic role of machines working as feedback apparatuses to control workers and capture their know-how on the productive process.¹⁰⁸

This embedded or embodied nature of information *in* subjects binds or connects subjects into supersubjects. The ability of information to connect and bind has been observed in numerous studies of bureaucracy. For instance, Hull has shown how the circulation of files acts to draw people into matters across hierarchies, organisations, or across the porous state / civil society divide¹⁰⁹: 'A property document and a government file may inhabit the same world of

¹⁰⁷ Pasquinelli, M., 'To Anticipate and Accelerate: Italian Operaismo and Reading Marx's Notion of the Organic Composition of Capital,' *Rethinking Marxism*, 26.2 (2014) p.183.

¹⁰⁸ Ibid., p.184.

¹⁰⁹ Hull, *Government of Paper*, p.18.

bureaucratic inscription, but they circulate differently and gather around themselves different people and things'.¹¹⁰

Records bind people to each other, to institutions and to obligations. Once records (as evidence of subject–state interactions) accrue, they stand between the seat of power and the body politic. The archive not only contains evidence, but, as a corpus of records, is itself evidence – evidence of the tacit acceptance of the social contract, because the records are, as Vismann argues, dialogical. It no longer matters if the seat of power is empty, if the original (social) contract is missing, because a series of *supplements* (the archives) have been created. As archival theorist Heather MacNeil put it, 'in a bureaucratic environment recordkeeping represents a kind of controlled speech', and this speech is dialogical.¹¹¹ Thompson wrote of the Yugoslavian self-management system 'From the point of view of power, anyone who opposed the system was still supporting it, as long as he or she kept participating'.¹¹² To govern the body politic, bureaucracy needs the participation of the body politic; the body must respond to commands. Engaging in the dialogue, writing back, is the act of acquiescence by which the body politic surrenders to the command, just as K. allows himself to be swallowed up by the obscure machinery of the law in *The Trial*. If the state is constituted through writing,¹¹³ the 'wheels of government grind to a halt without a file'¹¹⁴ and documents 'embodied and congealed state power'¹¹⁵ but these are dialogical products, what are the political implications of not writing back?

2.5. Information in the Societies of Discipline and Control

Benito Mussolini declared '*Tutto nello Stato, niente al di fuori dello Stato, nulla contro lo Stato*' (everything in the state, nothing outside the state, nothing against the state).¹¹⁶ We may see this as a climax of what Deleuze termed the 'society of

¹¹⁰ Ibid., p.20.

¹¹¹ MacNeil, H., *Trusting Records: Legal, Historical and Diplomatic Perspectives* (New York: Springer, 2000) p.62.

¹¹² Thompson, M., *A Paper House* (New York: Pantheon, 1992) p.39.

¹¹³ Gupta, *Red Tape*, p.143.

¹¹⁴ Ibid., p.146.

¹¹⁵ Ibid., p.208.

¹¹⁶ Mussolini, B., *Discorsi Del 1927*, 2nd ed. (Milan: Alpes, 1928) p.157.

discipline' – climax because the state was at its most monumental, the social contract most rigidly enforced, and bureaucracy at its most incontrovertible. And documented bodies were brought into and excluded from that state and contract. As suggested earlier in this chapter, in such circumstances, the refusal to write back is an act that propels the subject into bare life, an exclusion zone outside of Athens, beyond the law and outside of juridical personhood. According to Deleuze:

The disciplinary societies have two poles: the signature that designates the *individual*, and the number or administrative numeration that indicates his or her position within a *mass*. This is because the disciplines never saw any incompatibility between these two, and because at the same time power individualizes and masses together, that is, constitutes those over whom it exercises power into a body and molds the individuality of each member of that body.¹¹⁷

The concurrence of individualising and massing tendencies in societies of discipline is apparent in bureaucratic writing, the former tendency in the letter and case file and the latter in the compilation of statistics, for instance from census forms. This tension between individualising and massing echoes Banakar's identification of the paradox of K's position in *The Trial*: his singularity must be sublimated to the court's objectivity in order to be recognised.

Deleuze proposed the Second World War as the start of a transition between the society of discipline and the society of control:

In the societies of control... what is important is no longer either a signature or a number, but a code: the code is a *password*.... The numerical language of control is made of codes that mark access to information, or reject it. We no longer find ourselves dealing

¹¹⁷ Deleuze, G., 'Postscript on the Societies of Control,' *October*, 59 (1992) p.5.

with the mass/individual pair. Individuals have become “*dividuals*”, and masses, samples, data, markets, or “*banks*”.¹¹⁸

This supposed shift is not yet entirely reflected in the nature of written individual-state interactions, but the move to increasingly interconnected e-government portals such as that established in Estonia, where citizens use passwords to access a central service hub fuelled by matched data, is arguably a move in this direction.¹¹⁹ These systems unify information about subjects and in so doing, bind subjects more closely to the state as they become more and more richly documented. Similarly, it could be argued that the proliferation of civic tech apps reposition political subjects as prosumers more closely tied to the state by the data they feed it. Jefferson Adams observed of the information management practices of the GDR:

While most modern organizations have a propensity to generate their own peculiar acronyms and abbreviations, this tendency reached almost manic proportions in the various divisions of the security apparatus. Much of the reason was simply to keep information as compartmentalised as possible, and undecipherable to those not directly involved.¹²⁰

Adams is describing an informational panopticon, where information is rigorously organised and searchable and the archive is a site of anonymised power. Adam Smith’s invisible hand theorem suggests that no central authority could acquire or process the information necessary for regulating markets to the satisfaction of individuals’ needs more effectively than they would be regulated to that end as an effect of individuals’ actions in a free market. However, technology is opening the possibility of the realisation of the dreams of the GDR, the informational panopticon, in a way Smith could not have foreseen; centralised e-government portals and data-driven civic tech are only two examples. Finn Brunton and Helen

¹¹⁸ Ibid.

¹¹⁹ Government of Estonia, *E-Estonia*, (n.d.). <https://e-estonia.com> [last accessed 18 January 2019].

¹²⁰ Adams, J., 'Probing the East German State Security Archives,' *International Journal of Intelligence and Counter Intelligence*, 13.1 (2000) p.26.

Nissenbaum see a general trend that is driven by information technologies and directly relates to the individual's place and participation in the social contract, enjoyment of rights, etc:

Things we once thought were private – if we thought of that at all – become open, visible, and meaningful to new technologies. This is one aspect of the information asymmetry that shapes our practices of privacy and autonomy: we don't know what near-future algorithms, techniques, hardware, and databases will be able to do with our data. There is a constantly advancing front of transition from meaningless to meaningful – from minor life events to things that can change our taxes, our insurance rates, our access to capital, our freedom to move, or whether we are placed on a list.¹²¹

New technologies and the emergence of this big data environment allow the regime of documentary writing to take on more granularity, with more data points, more richly documented persons and actions, and discrete information exchanges opened up to comparison and reconciliation, and it becomes possible to speak of 'data as command'. Paul Virilio speaks of 'global information' when he suggests an impending information war, 'which will make general what espionage and police surveillance inaugurated long ago, though they were unable to draw, as we are today, on the limit-acceleration of "global information"'.¹²² In the emerging society of control, global information goes beyond the dialogical collection of information: technology, psychology, decision sciences, behavioural economics and marketing are converging around the collection, aggregation and repurposing of data about individuals and their preferences that is volunteered by them to different and seemingly unconnected systems.¹²³ This is often unknown

¹²¹ Brunton, F. & Nissenbaum, H., *Obfuscation: A User's Guide for Privacy and Protest* (Cambridge: MIT Press, 2015) p.50.

¹²² Virilio, P., *The Information Bomb*, 2nd ed. (London; New York: Verso, 2005) p.140.

¹²³ For instance, in 2013 Kosinski, Stillwell and Graepel showed the accuracy with which sexual orientation, race and political allegiances can be predicted from an analysis of a data subject's Facebook 'likes'. See: Kosinski, M., Stillwell, D., & Graepel, T., 'Private traits and attributes are predictable from digital records of human behavior,' *Proceedings of the National Academy of Sciences*, 110.15 (2013) pp.5802–5805. In 2000, the economist Joseph Stiglitz, proposing some lines of enquiry for the field of information economics, predicted: 'Some [of the advances] will entail an integration of economics with other social sciences – with psychology, on, for instance, how individuals process information, form expectations, and select among possible signals; and with sociology, on, for instance, the creation of social knowledge and signaling conventions.'

or poorly understood by the 'data subjects' who, as demonstrated by the existence of the United Kingdom's programme of data collection through the Government Communications Headquarters' (GCHQ) Tempora operation, exist within an informational panopticon. On Friday, 21 June 2013 *The Guardian* reported that:

GCHQ has secretly gained access to the network of cables which carry the world's phone calls and internet traffic and has started to process vast streams of sensitive personal information which it is sharing with its American partner, the National Security Agency (NSA)... This is all being carried out without any form of public acknowledgement or debate.¹²⁴

This informational panopticon was predicted by Virilio in 2000, when he wrote that

...with the network of networks, the Internet, comes the imminent establishment of real *networks of transmission of the vision of the world...* which will contribute... to developing the **panoptical** (and permanent) tele-surveillance of planetary sites and activities...¹²⁵

This echoes Deleuze's statement that the 'disciplinary man was a discontinuous producer of energy, but the man of control is undulatory, in orbit, in a continuous network'¹²⁶: The subject in a continuous network, open to tele-surveillance, in a panopticon that allows information about them to be gathered, but limits their access to that information. The scale of modern government dataveillance practices belies this vision as merely a trope of dystopian science fiction. Of GCHQ's Tempora operation, *The Guardian* reported:

Stiglitz, J. E., 'The Contributions of the Economics of Information to Twentieth Century Economics,' *The Quarterly Journal of Economics*, 115.4 (2000) p.1471.

¹²⁴ MacAskill, E., Borger, J., Hopkins, N., Davies, N. & Ball, J., 'GCHQ taps fibre-optic cables for secret access to world's communications,' *The Guardian*, (21 June 2013) <http://www.theguardian.com/uk/2013/jun/21/gchq-cables-secret-world-communications-nsa> [last accessed 25 April 2018].

¹²⁵ Virilio, *The Information Bomb*, p.121.

¹²⁶ Deleuze, G., 'Postscript on the Societies of Control,' *October*, 59 (1992) p.5.

The documents [leaked by Snowden] reveal that by last year GCHQ was handling 600m "telephone events" each day, had tapped more than 200 fibre-optic cables and was able to process data from at least 46 of them at a time. Each of the cables carries data at a rate of 10 gigabits per second, so the tapped cables had the capacity, in theory, to deliver more than 21 petabytes a day – equivalent to sending all the information in all the books in the British Library 192 times every 24 hours.¹²⁷

This hunger for data has geopolitical drivers. Jean-Francois Lyotard recognised the gravity of the implications of this connection between information and state power, anticipating the current prominence of data as a resource for both business and government:

Knowledge in the form of an informational commodity indispensable to productive power is already, and will continue to be, a major – perhaps *the* major – stake in the worldwide competition for power. It is conceivable that the nation-states will one day fight for control of information, just as they battled in the past for control over territory, and afterwards for control of access to and exploitation of raw materials and cheap labour. A new field is opened for industrial and commercial strategies on the one hand, and political and military strategies on the other.¹²⁸

The transition between societies of discipline and control is marked, according to Deleuze, by a 'crisis of the institutions', which is 'the progressive and dispersed installation of a new system of domination'.¹²⁹ The state's increasing appetite for information, acknowledged by Lyotard and demonstrated by Tempora, or the United States' National Security Agency's programme of telephone and internet interceptions, might be read as a paroxysm of this crisis; the recent high profile

¹²⁷ MacAskill et al, 'GCHQ taps fibre-optic cables'.

¹²⁸ Lyotard, J-F., *The Postmodern Condition: A Report on Knowledge*, trans. by Bennington, G. & Massumi, B. (Manchester: Manchester University Press, 1984) p.5.

¹²⁹ Deleuze, G., 'Postscript on the Societies of Control', p.7.

whistleblowing cases as signs of resistance to an emerging system.¹³⁰ According to Virilio, this is 'the information bomb, capable of using the interactivity of information to wreck the peace between nations'.¹³¹ If Deleuze has accurately described the current trajectory of history, how should the subject respond? In *On the Shores of Politics*, Ranciere wrote:

Exhibition in place of appearance, exhaustive counting in place of imparity, consensus in place of grievance – such are the commanding features of the current correction of democracy, a correction which thinks of itself as the end of politics but which might better be called post-democracy.¹³²

Ranciere has outlined a revitalisation of the adversarial aspect of democracy. He observes moments of 'impertinent dialectic'¹³³ when a 'community of equals' forms through the assertion of rights, by answering back, though not necessarily according to the forms and structures of the bureaucracy, that is, not in its own terms, but through different means, more direct and public. The result of this dynamism is a constant state of 'self-correction':

... the good regime is characterized precisely by the fact that it is always a mixture of constitutions, a constitutional marketplace. A regime without mixture, a regime which wants all its laws and institutions to resemble its basic principle, condemns itself to civil war and ruin because of the very unilateralism of this principle. In order to approach perfection, each regime must therefore correct itself, striving to welcome opposing principles, to make itself unlike itself. There is never such a thing as a good regime, in fact, only regimes off course engaged in the perpetual work of self-correction – one might almost say of self-dissimulation.¹³⁴

¹³⁰ This refers to the United States National Security Agency's (NSA) so-called 'warrantless wiretapping' and to Edward Snowden's famous 2013 disclosure of United States government documents relating to various national and international surveillance programmes.

¹³¹ Virilio, *The Information Bomb*, p.63.

¹³² Ranciere, *On the Shores of Politics*, p.98.

¹³³ Ibid., p.91.

¹³⁴ Ibid., p.42.

The impertinent dialectic is, *inter alia*, a technique of *dissensus*, which Ranciere says is not a 'conflict of interests, opinions or values; it is a division inserted in 'common sense': a dispute over what is given and about the frame within which we see something as given.'¹³⁵ The impertinent dialectic it is not possible without a change in the power dynamic that currently operates through bureaucracy and the record, which are both sites of dissensus even as they figure in dissensual renegotiations of power.

Every way in which we interact with government through information becomes, then, deeply political. If, as I have argued, information is the materiel of bureaucracy, the information we supply contributes to the acquisition of knowledge/power and the feedback required of commands, the qualities of that information bear directly on the operations of the machinery of government. If the individual subject is disadvantaged by the obscurity of bureaucracy's rules and operations, how we bring those rules and operations to light through information sharing practices has effects on power imbalances. Hetherington observed that when 'guerilla auditors encounter documents, they unleash undisciplined interpretations and hence novel possibilities into the situations that the documents purport to describe'.¹³⁶ Documents, according to Hetherington, are 'artifacts with many qualities which can be momentarily actualized as information in the multiple readings, negotiations, and tussles into which they are introduced. The insight of guerilla auditors is to never see documents as the end point, but as the site of possibility, not a store of information as a static thing but as a tool for making it as a political effect'.¹³⁷ What are the radical uses of government records? Who has rights in and over records? What will constitute an archive and how will it be used? As datafication continues, who and what will be data points to be individualised and massed together? Robertson observed that 'Digital verification has moved the claimed authority of identification closer to that of a "mechanical objectivity", if not a scientific objectivity, as proponents claim that the assessment of individuals is less necessary. This has been achieved as new

¹³⁵ Ranciere, J., *Dissensus: On Politics and Aesthetics* (London: Bloomsbury, 2010) p.77.

¹³⁶ Hetherington, *Guerrilla Auditors*, p.9.

¹³⁷ *Ibid.*, p.166.

technologies have been employed to offer a new set of relationships between the body, the archive, and the document'.¹³⁸ What is the nature of these technologies? When information processing is automated, how do algorithms operate, how have machines been trained and what does this mean for government, power and the subject?

These and many other questions emanate out of an understanding of the record-as-command. Rita Raley has claimed that for 'every system of disciplinary power... there is a "countervailing" response from those in precarious, subordinate, or marginal positions, which is to say that dataveillance and counterveillance must be seen as inextricably connected'.¹³⁹ If the exchange of information is the essence of the state-subject power relation, the way the subject treats information, both as its writer and its reader, is political and essential to any impertinent dialectic. Monroe wrote that '... no matter how discrete, fixed, or closed a regime / system / machine appears to be, it always contains within its coding possibilities of escape, supersession, obsolescence, disintegration, or mutation'.¹⁴⁰ When the regime, system or machine is bureaucracy, records are foremost among those possibilities.

2.6. Conclusion

Acquiescence and complicity are features of the social contract, and although those terms often bear negative connotations, as with the term 'bureaucracy', they can only be positive or negative in particular instances and from particular perspectives. Whatever their character, acquiescence and complicity in the context of the social contract require performative utterances in the form of commands and responses to commands, which, in the modern state, are usually issued in writing. It is documentation that makes the physical person into a juridical person, inside the regime of law and rights. This documentation is not fixed, but accrues and moves around its subject as s/he interacts with the state through its bureaucracy. This documentation is the product of disciplinary writing.

¹³⁸ Robertson, *The Passport in America*, p.249.

¹³⁹ Raley, R., 'Dataveillance and Countervailance', in Gitelman, L. (ed.) *Raw Data is an Oxymoron*, (Cambridge: MIT Press, 2013) p.131.

¹⁴⁰ Monroe, *Interrogation Machine*, p.8.

I have shown that the record-as-command is a form of disciplinary writing that allows power to operate at a distance, by standing in for authority, and to designate as unique and controllable people, actions, places, relationships, etc.: to make of everything a data point. These records are thereby made performative in an immediate and localised sense and in a broader political sense. Describing tenant farmers' encounters with official records, Hetherington says that they 'began to look past the politics of land titles and to see the whole bureaucracy as a series of material entanglements built around papers'.¹⁴¹ Records bring bureaucratic control to relationships, persons and territories in a material sense because they are performative utterances in writing. Hull's observation that bureaucracies are 'among the most consciously materialized of social collectives – painstakingly fabricated in the layouts of offices, the writings of functionaries, the stampings of clerks, the movement of files – because they are designed to unify and control individuals conceived as either naturally independent and refractory or entangled in other collectives'¹⁴² illustrates the controlling and embodied nature of bureaucracies, which we can extend to records. Hull has suggested that government records join people with bureaucracies while excluding them from their authorised production.¹⁴³ The operation of authoritative archives is another example of the state monopoly on the command function of records.

Beyond this, records as deployed by bureaucracies can also manufacture power in two distinct ways. Firstly, as Feldman has shown, they, and the processes they animate, engage subjects in relationships that channel discourses and actions in controlled ways, so that governance is done regardless of the validity or strength of the legal or actual basis of authority. Secondly, as is widely recognised in the literature but perhaps best explicated by Richards, records contain the information required as a knowledge base for governments. Record-making and record-keeping allow data to be collected, aggregated, analysed and used in various ways by governments. To be outside of this regime of documentation is to be outside of the social contract, that is, to leave behind juridical personhood

¹⁴¹ Hetherington, *Guerrilla Auditors*, p.143.

¹⁴² Hull, *Government of Paper*, p.129.

¹⁴³ *Ibid*, p.22.

and approach bare life. But to the law, and anyone who wishes to operate within it, these records must be authentic. They must be what they purport to be if they are to be used to good effect by bureaucracies and the subjects documented in them and/or engaged in the processes that make use of them.

Bureaucracies are *dispositifs* that are multivalent and have porous borders. As Vismann says, they are supersubjects composed of subjects. They are staffed by individuals that are also subject to their procedures and rules. Their power is derived, *inter alia*, from the acquiescence and complicity of their staffs and their subjects, which is made material through control information and valorising information, as defined by Alquati. Even individual instances of bureaucracies cannot be characterised definitively, since their priorities and modes of operation are always in flux as a result of their supersubject nature. Butler has illustrated the negative aspect of this phenomenon, where appointed, rather than elected officers take positions of power that can have direct and serious consequences for the lives of others. In other instances, bureaucracies are benevolent, functional, even radically progressive.

If Deleuze's notion of a shift from a society of discipline to a society of control is in progress, the nature of information use by bureaucracies may follow his description, shifting from the primacy of the signature and the number, which individualises and masses the subject respectively, to the primacy of the password, where individuals become 'dividuals', which I interpret to mean discernible as a set of data points. This datafication of the individual, as cited above, is already seen at the nexus of marketing, psychology and algorithm design. Deleuze suggests this shift is to be seen in a 'crisis of the institutions'; we may interpret growing government hunger for data as evidenced in dataveillance programmes as a sign of crisis, but power's need for data is consistent: this panoptical impulse has a parallel with the archival impulse of imperialism, for instance.

One response to this is dissensus, Ranciere's idea of a radical remaking of democracy via an impertinent dialectic. According to Ranciere, political subjects are 'not social groups but forms of inscription that (ac)count for the

unaccounted'¹⁴⁴ and '...rights are inscriptions, a writing of the community as free and equal... the Rights of Man are the rights of those who make something of that inscription, deciding not only to 'use' their rights but also to build cases to verify the power of the inscription'.¹⁴⁵ These inscriptions – social groups, rights – in fact, all data points in the modern state – might be re-inscribed in different ways and places. How disciplinary writing in general and records in particular are used, abandoned or rethought under this impertinent dialectic is only starting to be seen in the technological infrastructures established for whistleblowing, the theorising of rights in records, co-creation and participatory archives, and open data practices such as data journalism and the use of civic technologies. As Ranciere observes, the 'strength of those rights lies in the back-and-forth movement between the initial inscription of the right and the dissensual stage on which it is put to the test'.¹⁴⁶ Following my development of the idea of the record-as-command, building on the work of Vismann, Feldman, Gupta and others, information can be seen as such a dissensual stage.

¹⁴⁴ Ranciere, *On the Shores of Politics*, p.43.

¹⁴⁵ Ibid., p.76.

¹⁴⁶ Ibid., p.81.

Chapter 3

Information Asymmetry in the Social Contract

If the 'present' is indeed the axis of symmetry of passing time, this **omnipresent** centre now controls the totality of the life of the 'advanced' societies and we have at all costs to avoid the 'breaking' of that axis, which would lead back to the 'past', to dead memory and – who knows? – to remorse. Have we not in recent years seen the emergence of a vogue for repentance, a wave of guiltless officials apologising for offences committed by their predecessors, but less worried, it would seem, about the crimes they may currently be committing? Or, alternatively, we must avoid that sudden temporal 'symmetry-breaking' which could project us into the future – something we have been partially cured of by the failures of economic planning.¹⁴⁷

Paul Virilio, *The Information Bomb*

¹⁴⁷ Virilio, *The Information Bomb*, p.126.

3.1. Introduction

Having established the dialogical power of information, and argued that information should be recognised as a site for dissensus, I will now explore how information currently operates under the social contract. In this chapter I will show that information asymmetry is a feature of the social contract. I will look at the treatment of information asymmetry in political science and information economics to understand how those fields have approached the correction of information asymmetry. I will describe the ontology of information problems and solutions delineated in information economics and argue that, of the solutions to information asymmetry advanced in the economics literature, 'monitoring' is vital to correcting information asymmetry in the context of the social contract.

3.2. The Social Contract and Agency

The social contract is a relationship of agency. The character of the contract depends on the nature and extent to which agency is delegated or relinquished and to whom, such that totalitarian and democratic regimes display agency relations while maintaining very different characters, particularly in terms of force and participation. The relationship of agency is:

...one of the oldest and commonest codified modes of social interaction. We will say that an agency relationship has arisen between two (or more) parties when one, designated as the agent, acts for, on behalf of, or as a representative for the other, designated the principal, in a particular domain of decision problems. Essentially all contractual arrangements, as between employer and employee or the state and the governed, for example, contain important elements of agency.¹⁴⁸

'Agents' and 'principals' appeared in the lexicon of the social sciences at the middle of the twentieth century. Barry Mitnick provides an overview:

¹⁴⁸ Ross, S.A., 'The Economic Theory of Agency: The Principal's Problem,' *The American Economic Review* 63.2 (1993) pp.134–39.

In the accounting and control literature, Cooper (1949, 1951) discussed agents inside the firm; in economics Downs (1957) referred to agents in his economic theory of democracy and Arrow (1963) referred to agents and to delegation to agents in his discussion of characteristic problems in medical care and the response of institutions to those problems; in political science, Pitkin (1967) and Tussman (1960) used agent-principal language in works on political philosophy; and, in sociology, Swanson (1971) described collective society using such terms.¹⁴⁹

In the 1970s, a theory of agency emerged as a way of examining problems arising from the delegation of work to an agent by a principal. The origins of agency theory are contested. Michael Jensen and William Meckling's work¹⁵⁰ is often cited as the origin of the theory, but Mitnick contests this,¹⁵¹ describing a separate but nearly concurrent development of agency theory from the perspectives of institutional theory (his own work) and economics (Stephen Ross' work) from 1972/3. In both of its strains, agency theory focuses on principal-agent problems and how they can be solved. Principal-agent problems are those problems that arise from the delegation of agency from a principal to an agent. As examples, a business owner (the principal) hires a manager (the agent) to run the business, or an electorate (the principal(s)) elects a candidate (the agent) to represent its / their interests in parliament. These relationships become problematic when the goals of the two parties diverge. Agency theory 'endeavors to surface contractual problems arising from the assumption that the agents will behave opportunistically if their interests conflict with those of the principals.'¹⁵²

Mitnick has defined four types of behaviour that the agent might choose to engage in; the types are defined by the extent to which the agent pursues its 'self-goals' or 'other-goals' (goals of the principal).

¹⁴⁹ Mitnick, B. M., *Origin of the Theory of Agency: An Account By One of the Theory's Originators*, (2013) <http://dx.doi.org/10.2139/ssrn.1020378> [last accessed 17 August 2018].

¹⁵⁰ Jensen, M. C. & Meckling, W. H., 'Theory of the firm: Managerial behavior, agency costs and ownership structure,' *Journal of Financial Economics*, 3.4 (1976) pp.305–360.

¹⁵¹ Mitnick, *Origin of the Theory of Agency*, pp.1-16.

¹⁵² Sharma, A., 'Professional as Agent: Knowledge Asymmetry in Agency Exchange,' *The Academy of Management Review*, 22.3 (1997) pp.758–98.

The *pure fiduciary* will choose that collection of specifications which contains the highest returns to his other-goals, regardless of the level of return to his self-goals. The lexicographic, or "lexical" fiduciary will after choosing the collections with the highest return to the principal, choose that collection which also has the highest return to his self-goals. These types may be contrasted with types of agents who choose for their self-goal first: The pure self-interest agent chooses that collection of specifications which contains the highest returns to his self-goals, regardless of the level of return to his other-goals. The lexical self-interest agent chooses, first, those collections of specifications which contain the highest returns to his self-goals and, second, from that group of collections, that collection which has the highest return to his other-goals.¹⁵³

Mitnick posits that, through policing, the pure self-interest agent can be encouraged to become a pure fiduciary agent.¹⁵⁴ The global financial crisis of 2007-2008 provides an example of large-scale lexical self-interest behaviour, which was demonstrated by banks that were operating in an environment in which high-risk behaviour (such as subprime lending, in the United States) was incentivised, resulting in 'predatory lending'; the agents (the banks) acted in their self-interests in extending credit to the principals (risky clients). Agency theory holds that, although economic inefficiency is inevitable in principal-agent relationships, inefficiency can be reduced through improving the contracts between parties, and that the best contracts align the interests of the agent with those of the principal.

¹⁵³ Mitnick, B. M., 'The Theory of Agency: The Policing "Paradox" and Regulatory Behavior', *Public Choice*, 24 (Winter 1975) p.28.

¹⁵⁴ Ibid.

What does a principal-agent problem model look like? Gary Miller defines the 'canonical' form of the principal-agent problem as those with the following 'core assumptions'¹⁵⁵:

1. Agent impact. Agent impact describes the ability of the agent to take actions that determine, in combination with a risk variable, an outcome for the principal. This is fundamental to the principal-agent relationship in that no such relationship could exist without an agent empowered sufficiently as to have an impact on the principal.

2. Information asymmetry. Information asymmetry describes the imbalance of information, in terms of quantity and / or quality, between the parties to a contract. In the principal-agent problem model, the imbalance favours the agent, whose actions are difficult to monitor. For Joseph Stiglitz, information asymmetry is a primary cause of the principal-agent problem:

Lack of information and the economies of specialization... force owners (shareholders) to delegate responsibility for gathering information and making decisions to managers. But that same lack of information gives managers discretion to act in their own interests – as opposed to the interests of the shareholders.¹⁵⁶

3. Asymmetry in preferences. This assumption describes differences between the preferences of the agent and the principal, creating risk for the principal. This gives rise to what Mitnick calls 'self-goals' (goals of the agent) and 'other-goals' (goals of the principal) – the basis of his classification of behaviours outlined above.¹⁵⁷

4. Initiative that lies with a unified principal. Miller writes that 'The principal acts rationally based on a coherent set of preferences, and is able to move first by

¹⁵⁵ Miller, G. J., 'The Political Evolution of Principal-Agent Models,' *Annual Review of Political Science* 8 (2005) pp.205-6.

¹⁵⁶ Stiglitz, J. E., 'The Contributions of the Economics of Information to Twentieth Century Economics,' *The Quarterly Journal of Economics*, 115.4 (2000) p.1467.

¹⁵⁷ Mitnick, 'The Theory of Agency', p.3.

offering a contract'.¹⁵⁸ In this assumption, the principal has the power to offer or withhold a contract from the agent. The principal initiates the principal-agent relationship by offering a contract that delegates powers to the agent.

5. Backward induction based on common knowledge. Miller writes:

Principal and agent share knowledge about the structure of the game, effort costs, probability distribution of outcomes, and other parameters. Just as important, they share common knowledge of the agent's rationality; both know that the agent will prefer any incentive package with an expected utility slightly more than the agent's opportunity cost. This leads to backward induction by the principal. The principal can infer the agent's best response function from known parameters and use backward induction to identify the best possible outcome, subject to that function.¹⁵⁹

6. Ultimatum bargaining. The bargaining power in the principal-agent relationship is assumed to be in the hands of the principal, or the 'principal is presumed to be able to impose the best possible solution from the agent's correctly inferred best response function'.¹⁶⁰

Steve Sorrell et al note that principal-agent problems 'pervade both markets and organisations',¹⁶¹ and the spread of the application of the model into the study of organisations and other forms of contractual relationships was rapid. By the early 1980s, the principal-agent problem model was being applied to the study of the public sector. An early example is Barbara Spencer's work on asymmetric information in government budget-setting.¹⁶²

¹⁵⁸ Miller, 'The Political Evolution of Principal-Agent Models', p.206.

¹⁵⁹ Ibid.

¹⁶⁰ Miller, 'The Political Evolution of Principal-Agent Models', p.206.

¹⁶¹ Sorrell, S., et al, *The Economics of Energy Efficiency: Barriers to Cost-Effective Investment* (Cheltenham: Edward Elgar, 2004) p.41.

¹⁶² Spencer, B. J., 'Asymmetric Information and Excessive Budgets in Government Bureaucracies: A Principal and Agent Approach,' *Journal of Economic Behavior and Organization* 3.2 (1982) pp.197–224.

Agency theory has been controversial. In her defence of the theory, Kathleen Eisenhardt observed that it had been called ‘trivial, dehumanizing and even “dangerous”’.¹⁶³ Despite Michael Lubatkin’s more recent criticism¹⁶⁴ of the use of agency theory in corporate governance research on the grounds of its behavioural assumptions (rebutted by Thomas Lange in the same year),¹⁶⁵ agency theory is now enjoying new currency in response to developments in the configuration of relationships in public service delivery, for instance new forms of public-private or public-third sector partnerships to deliver social services, outsourcing and privatisation:

Explicitly or implicitly, principal-agent theory has become the dominant framework to understand the nature of delegation and interaction among actors who are involved in exchanging resources in the hollow state.¹⁶⁶

Using a principal-agent approach to look at contracting out local government services in Ghana, Nicholas Awortwi found that well-constructed contracts were not sufficient to ensure an alignment between the preferences of the principal (government) and agent (private firm), but that outsourcing provided an arena for the asymmetry in preferences between the public and the public servant to become apparent, with public servants tempted into agency that is harmful to the interests of their principals (citizens) by the opportunity for corruption.¹⁶⁷ Studying

¹⁶³ Eisenhardt, K. M., ‘Agency Theory: An Assessment and Review,’ *The Academy of Management Review* 14.1 (1989) pp.57–74.

¹⁶⁴ Lubatkin, M. H., ‘A Theory of the Firm Only a Microeconomist Could Love,’ *Journal of Management Inquiry*, 14.2 (2005) pp.213–16.

¹⁶⁵ Lange, T., ‘A Theory of the Firm Only a Microeconomist Could Love? A Microeconomist’s Reply to Lubatkin’s Critique of Agency Theory,’ *Journal of Management Inquiry* 14.4 (2005) pp.404–6.

¹⁶⁶ Lamothe, M., ‘Redesigning the Hollow State: A Study of Florida Child Welfare Service Reform through the Lens of Principal-Agent Theory,’ *International Journal of Public Administration* 34.8 (2011) pp. 497. Lamothe uses the term ‘hollow state’ in the sense that it is used by Milward and Provan: ‘The hollow state is a metaphor for the increasing use of third parties, often nonprofits, to deliver social services and generally act in the name of the state’ (Milward, H. B. & Provan, K. G., ‘Governing the Hollow State,’ *Journal of Public Administration Research and Theory* 10.2 (2000) pp.359–80). An example of comparable work outside of the capitalist economic system is Peng’s ‘Chinese Villages and Townships as Industrial Corporations: Ownership, Governance, and Market Discipline,’ *American Journal of Sociology* 106.5 (2001) pp.1338–70, which was also written at a time of changing public-private relationships.

¹⁶⁷ Awortwi, N., ‘Contracting Out Local Government Services to Private Agents: An Analysis of Contract Design and Service Delivery Performance in Ghana,’ *International Journal of Public Administration* 35.13 (2012) pp.886–900.

child welfare service reform in Florida as a principal-agent problem, Lamothe found:

Consolidating myriads of contracts into the hands of a few big vendors and standardizing contract terms across the lead agencies result in efficiency and effectiveness in managing and overseeing contracts by the principal government, but at the same time, unite the contract agents to mobilize their resources and confront their masters more aggressively as their common interests solidify their forces.¹⁶⁸

These applications demonstrate that diverse insights can be gained by applying the principal-agent problem model to socio-political subjects. Eisenhardt,¹⁶⁹ Peter Wright, Ananda Mukherji and Mark Kroll,¹⁷⁰ and others have advocated more diverse applications of the theory, in less rigidly canonical forms. Applications of the model outside of economics often diverge from the canonical form. Miller has noted that ‘Many more political science applications have relaxed one or more assumptions. The result has been a flowering of insight about informational asymmetries and incentives that has made enormous contributions to various subfields of political science.’ It is useful to consider each of the core assumptions of the canonical form and whether or how they have been relaxed during applications to socio-political subjects, and what this suggests for its application to the social contract.

The notion that initiative lies with a unified principal (Miller’s fourth assumption) is anathema to the social situation, where the principal is not unified, but multiple and diverse: there are often millions of principals. This observation was central to Cook’s criticism of Wood’s principal-agent model of political control of bureaucracy as failing to recognise the ‘multiple principal nature of the U.S.

¹⁶⁸ Lamothe, ‘Redesigning the Hollow State’, p.510.

¹⁶⁹ Eisenhardt, ‘Agency Theory’.

¹⁷⁰ Wright, P., Mukherji, A. & Mark, M. J., ‘A Reexamination of Agency Theory Assumptions: Extensions and Extrapolations,’ *The Journal of Socio-Economics* 30.5 (2001) pp.413–29.

system'.¹⁷¹ The principal-agent relationship(s) in social contracts is not 'one to one' but 'many to one', or indeed 'many to many', if we consider the profusion of the organs of the state (or the boundaries of bureaucracy, reflecting the problem of defining the firm outlined in Chapter 2). Under this fourth assumption, Miller posits that the principal is able to offer a contract. In the context of the social contract, the principal(s) seems, *prima facie*, not to be in a position to offer a contract, nor, recalling Serres' treatment of Anaxagoras of Clazomenae, to extricate itself from the the pre-existing contract: to 'offer' a new contract is to revolt, or approach bare life – it would be a dissensual act.

Another issue arising from the lack of a unified principal in the socio-political context is the assumption of backwards induction. The principal and the agent do not 'share knowledge about the structure of the game, effort costs, probability distribution of outcomes, and other parameters'.¹⁷² Rather, the problem of information asymmetry in the social contract often extends to the basic parameters of the principal-agent relationship: citizens' knowledge of their rights and the duties of the government can vary considerably from citizen to citizen, to say nothing of people outside the formal demarcation of citizenship. That information asymmetry can extend to the basic parameters of the social contract is demonstrated in Akhil Gupta's study of Indian bureaucracy. He recounts an experience during his fieldwork in which local officials have convened a 'camp' (a kind of clinic or open day) to accept applications for state pensions from applicants fulfilling the eligibility criteria - those who are over sixty, without an adult son, and who own less than one and a half acres of land. Regarding an applicant who gave his age as approximately 55 years old, Gupta wrote:

It was almost certain that the applicants and headmen knew little about the eligibility criteria for the scheme and that information about eligibility was being passed around outside the building through rumor and speculation. Nothing else could explain why the applicant in the incident described above had stated his age as being less than

¹⁷¹ Wood, B., 'Principal-Agent Models of Political Control of Bureaucracy,' *The American Political Science Review* 83.3 (1989) p.965.

¹⁷² Miller, 'The Political Evolution of Principal-Agent Models', p.206.

the minimum required for the scheme.¹⁷³

Where bureaucracies are opaque and as they increasingly involve parastatal agencies and private service providers, information asymmetry first becomes apparent in the difficulties experienced by individuals in ascertaining which organ or official s/he must interact with, and on what terms. Knowledge of the parameters of 'the game' are not necessarily shared.

The assumption about ultimatum bargaining is entirely reversed in the context of the social contract. The individual principals do not have the ability to impose an ultimatum, and may struggle to bargain at all. Instead, the state has the power of ultimatum bargaining, through the influence of its bureaucracy, and ultimately through recourse to violence. As Max Weber wrote, the state is the institution that has 'the monopoly of violence in a given territory', and that monopoly firmly repositions ultimatum bargaining as the privilege of the state.

Other core assumptions survive an application to the social contract. Agent impact clearly persists; the actions of government help to determine payoffs to its citizens, the principals. Asymmetry in preferences also transfers to the social application of the model, since the preferences of the agent cannot align with the preferences of numerous and diverse principals. Finally, information asymmetry persists. As I have argued, information is the *sine qua non* of the social contract. 'Information asymmetry' describes an imbalance between the quantity and quality of information available to the parties to a contract; the principal and the agent. In the social contract, where the citizen is the principal and the government is the agent - a relationship that depends on the exchange of information (Alquati's control information and valorising information) - information asymmetry can have significant political and personal implications.

In summary, the core assumptions (1) agent impact, (2) information asymmetry, and (3) asymmetry in preferences are applicable to a study of the social contract, while assumption (4) initiative that lies with a unified

¹⁷³ Gupta, *Red Tape*, p.11.

principal, (5) backward induction based on common knowledge, and (6) ultimatum bargaining, are not applicable or should be considered to operate in reverse under the social contract. Why is information asymmetry the most significant feature of the model for this thesis? Roy Radner has discussed how partial information on the part of the principal can negatively affect agent behaviour,¹⁷⁴ and information economics may suggest that bringing the goals of principals and agents closer together requires a consideration of all components of the model, but when the social contract operates almost entirely through information, information asymmetry becomes the most critical aspect of the model. Victoria Lemieux found that principal-agent accountability mechanisms were vital to counteracting information asymmetry¹⁷⁵, and, as Stolowy et al have stated:

Information is at the heart of any contractual relationship. It allows investors to make “informed” investment decisions and to verify whether other contracting parties fulfill their obligations. As such, information is crucial for the realization of any accountability relationship.¹⁷⁶

Essentially, information is constitutive of the relationship of the social contract and any attempt to align the goals of the agent with those of the principals will require efforts towards symmetrical information.

3.3. Information Asymmetry

Information asymmetry was defined and theorised in the field of economics in response to problems with established economic models. Adam Smith’s *The Theory of Moral Sentiments* (1759) sketched out the ‘invisible hand’

¹⁷⁴ Radner, R., ‘Hierarchy: The Economics of Managing’, *Journal of Economic Literature*, Vol. 30, No. 3 (1992) pp.1382-1415.

¹⁷⁵ Lemieux, V., *Competitive Viability, Accountability and Record Keeping: A Theoretical and Empirical Exploration Using a Case Study of Jamaican Commercial Bank Failures*, doctoral thesis, University College London, 2001.

¹⁷⁶ Stolowy, H. et al, ‘Information, trust and the limits of “intelligent accountability” in investment decision making: Insights from the Madoff case,’ *Les Cahiers de Recherche* 956, (2011) <https://EconPapers.repec.org/RePEc:ebg:heccah:0956> [last accessed 25 April 2018].

theorem, which he would more fully develop in 1776 in *The Wealth of Nations*:

The proud and unfeeling landlord views his extensive fields, and without a thought for the wants of his brethren, in imagination consumes himself the whole harvest ... [Yet] the capacity of his stomach bears no proportion to the immensity of his desires ... the rest he will be obliged to distribute among those, who prepare, in the nicest manner, that little which he himself makes use of, among those who fit up the palace in which this little is to be consumed, among those who provide and keep in order all the different baubles and trinkets which are employed in the economy of greatness; all of whom thus derive from his luxury and caprice, that share of the necessaries of life, which they would in vain have expected from his humanity or his justice...The rich...are led by an invisible hand to make nearly the same distribution of the necessaries of life, which would have been made, had the earth been divided into equal portions among all its inhabitants, and thus without intending it, without knowing it, advance the interest of the society....¹⁷⁷

The invisible hand theorem - that resources are *naturally* efficiently distributed in a free market – would be linked with the notion of competitive equilibrium by the Italian economist Vilfredo Pareto in 1909. ‘Pareto optimality’, which describes an allocation of resources that cannot be changed to benefit one individual without disadvantaging another, became a cornerstone of twentieth-century economics. In 1951, Arrow and Debreu formulated the Welfare Theorems, which state that there is an equivalence between Pareto efficient outcomes and competitive price equilibria. The welfare theorems support Smith’s invisible hand theorem. These works by Smith, Pareto, Arrow and Debreu form the theoretical underpinning of modern free market economics, and the general equilibrium theory in particular.

¹⁷⁷ Smith, A., *The Theory of Moral Sentiments*, (New York: Penguin Classics, 2010) p.161.

These theories work on a number of assumptions. Arrow and Debreu presented proof of an equilibrium for an integrated model of production, exchange and consumption.¹⁷⁸ The Arrow-Debreu model is key to general equilibrium theory, since it is used to prove that general equilibrium is possible. Amongst the assumptions behind Arrow-Debreu is the assumption of 'perfect competition', where, through rational action, consumers pursue utility and producers pursue profit. Perfect competition is also characterised by 'perfect information'. The Oxford Dictionary of Finance and Banking defines perfect information as:

...a hypothetical state in which each participant in a market has complete up-to-date information about products and prices, and can therefore make perfectly rational choices.¹⁷⁹

In 1962, George Stigler published *Information in the Labor Market*, in which he proposed that the exploration of the problems of information in the economy might be valuable:

The amounts and kinds of information needed for the efficient allocation of labor, whether judged from the viewpoint of the laborer, the employer, or the community, extend far beyond the determination of wage rates. The kinds and amounts of skill men should acquire pose parallel informational problems, and so too do the non-monetary conditions of employment. The traditional literature has not done these problems justice.¹⁸⁰

In proposing new avenues for research, Stigler was cautious about overstating the importance of information problems:

¹⁷⁸ Arrow, K. J. & Debreu, G. 'Existence of an Equilibrium for a Competitive Economy,' *Econometrica* 22.3 (1954) pp.265–90.

¹⁷⁹ Law, J. & Smullen, J. (eds.), *A Dictionary of Finance and Banking*, 5th ed., (Oxford: Oxford University Press, 2014).

¹⁸⁰ Stigler, G. J., 'Information in the Labor Market,' *Journal of Political Economy* 70.5 (1962) pp.94–105.

It is doubtful that justice would be more closely approached by making exaggerated claims of the importance of the problem of information. There is no exaggeration however, in the suggestion that the analysis of the precise problems of information and of the methods an economy uses to deal with them appears to be a highly rewarding area for future research.¹⁸¹

In the early 1970s, a new generation of American economists began to question the assumptions made in mainstream economic theory, including the assumption of perfect information. Akerlof, Spence, Stiglitz and others became interested in information problems, which led to the emergence of the field of information economics. By 1972/3, when Ross and Mitnick were beginning to develop agency theory, the notion of information asymmetry had already been introduced into economics by Akerlof, in his 1970 paper *The Market for 'Lemons': Quality Uncertainty and the Market Mechanism*.¹⁸² Akerlof showed the effect of asymmetric information on markets using the example of used cars. Before buying a used car, the prospective buyer, not knowing the history of the car or the level of care shown by its previous owner, has even chances of buying a good car or a 'lemon'.

After owning a specific car, however, for a length of time, the car owner can form a good idea of the quality of this machine; i.e., the owner assigns a new probability to the event that his car is a lemon. This estimate is more accurate than the original estimate. An asymmetry in available information has developed: for the sellers now have more knowledge about the quality of a car than the buyers.¹⁸³

Akerlof showed how markets could be driven out of existence by a downward trend set in motion by the sale of inferior products and driven by declining

¹⁸¹ Ibid. p.104.

¹⁸² Akerlof, G. A., 'The Market for "Lemons": Quality Uncertainty and the Market Mechanism,' *The Quarterly Journal of Economics* 84.3 (1970) pp.488–500.

¹⁸³ Ibid., p.489.

customer expectations. His work underscored the significance of guarantees, licensing and branding, or more broadly, reputation and trust.

Arrow observed, in 1999, that 'Economic analysis in the last thirty years has been devoted in good measure to the analysis of the strategic implications of differences in information among economic agents'.¹⁸⁴ The ramifications of this analysis would be more profound than Stigler imagined when he sounded his note of caution about 'making exaggerated claims of the importance of the problem of information'. Writing in 2000 about the assumptions about information in mainstream economic models, Stiglitz stated:

Modern information economics turned these presumptions on their head: even small information costs can have large consequences, and many of the standard results—including the welfare theorems—do not hold even when there are small imperfections of information.¹⁸⁵

The effects of these investigations into information asymmetry on economic theory were profound:

The problems encountered were so serious that a whole generation of general equilibrium theorists gave up momentarily the grandiose framework of GE to reconsider the problem of exchange under asymmetric information in its simplest form, i.e., between two traders, and in a sense go back to basics.¹⁸⁶

¹⁸⁴ Arrow, K., 'Information and the Organization of Industry', in Chichilnisky, G. (ed.) *Markets, Information and Uncertainty: Essays in Economic Theory in Honor of Kenneth J. Arrow* (Cambridge: Cambridge University Press, 1999) p.20.

¹⁸⁵ Stiglitz, 'The Contributions of the Economics of Information to Twentieth Century Economics' p.1443.

¹⁸⁶ Laffont, J. J. & Martimort, D., *The Theory of Incentives: The Principal-Agent Model* (Princeton; Oxford: Princeton University Press, 2002) p.4.

Akerlof, Spence and Stiglitz would be jointly awarded a Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel in 2001 for their work on information problems.¹⁸⁷

3.4. Information Economics Approaches to Information Asymmetry

Stiglitz has stated that information economics has focused on two central problems: identifying characteristics (the selection problem) and monitoring behaviour (the incentive problem). The selection problem relates to information about the characteristics of the item being transacted, which may be, as in Akerlof's *A Market for Lemons*, the characteristics of a used car, or it may be staff productivity or the likely return on stocks. Stiglitz noted:

Smith, in anticipating later discussions of adverse selection, wrote that as firms raise interest rates, the best borrowers drop out of the market. If lenders knew perfectly the risks associated with each borrower, this would matter little; each borrower would be charged an appropriate risk premium. It is because lenders do not know the risk properties perfectly that this process of adverse selection has important consequences.¹⁸⁸

The incentive problem, also known as the moral hazard problem, is about behaviour, such as the behaviour of borrowers or workers. In the socio-political context, access to government information is important to the individual in addressing both the selection and incentive problems.

3.4.1. Selection Problems (Adverse Selection)

Selection problems arise from information asymmetry before contracting. The classic illustration of adverse selection is found in Akerlof's model of the second-

¹⁸⁷ Nobel Media AB, *The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2001* (2000) http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/2001/ [accessed 25 April 2018].

¹⁸⁸ Stiglitz, 'The Contributions of the Economics of Information to Twentieth Century Economics', p.1442.

hand car market. The potential buyer is disadvantaged by their lack of information about the mechanical condition of the available cars, and may more easily be led into buying a car in poor condition (the eponymous 'lemon'). Political applications of the concept of adverse selection are numerous. Eric Maskin and Jean Tirole,¹⁸⁹ Scott Ashworth,¹⁹⁰ Timothy Besley,¹⁹¹ Brandice Canes-Wrone and Kenneth Shotts,¹⁹² and James Alt, Ethan Bueno de Mesquita and Shanna Rose¹⁹³ have worked with theoretical models for elections with adverse selection, for instance.

The economics literature offers 'signaling' and 'screening' as ways of overcoming information asymmetry before a contract is signed (or a purchase or exchange made), that is, as solutions to selection problems. In 1981, Weiss and Stiglitz drew a distinction between signaling and screening, terms which had previously been used interchangeably by some economists.¹⁹⁴ Their distinction concerned which party moved first to offer a contract; signaling was used to describe scenarios in which the informed party moved first ('the more desirable informed agents signal who they are'¹⁹⁵), and screening was used for scenarios in which the uninformed party moved first ('contracts are designed to screen the more desirable agents from the less desirable ones'¹⁹⁶). The terms will be used here according to Stiglitz's and Weiss' distinction.

Signaling is the transmission of information from the informed party to the uninformed party to indicate, for examples; quality, ability or trustworthiness.

¹⁸⁹ Maskin, E. & Tirole, J., 'The Politician and the Judge: Accountability in Government,' *American Economic Review* 94.4 (2004) pp.1034–54.

¹⁹⁰ Ashworth, S., 'Reputational Dynamics and Political Careers,' *Journal of Law, Economics, and Organization* 21.2 (2005) pp.441–66.

¹⁹¹ Besley, T., *Principled Agents?: The Political Economy of Good Government* (Oxford: Oxford University Press, 2007).

¹⁹² Canes-Wrone, B. & Shotts, K. W., 'When Do Elections Encourage Ideological Rigidity?' *American Political Science Review* 101.2 (2007) pp.273–88.

¹⁹³ Alt, J., de Mesquita, E.B. & Rose, S., 'Disentangling Accountability and Competence in Elections: Evidence from U.S. Term Limits,' *Journal of Politics* 73.1 (2011) pp.171–86.

¹⁹⁴ Weiss, A. & Stiglitz, J., 'Sorting Out the Differences Between Signaling and Screening Models' (Cambridge: National Bureau of Economic Research, Inc., 1990) <http://www.nber.org/papers/t0093.pdf> [accessed 25 April 2018].

¹⁹⁵ *Ibid.*, p.3.

¹⁹⁶ *Ibid.*

We see banks signaling their trustworthiness by the size of their edifices, guarantees signaling a firm's confidence in the quality of its products, owner-managers of firms signaling their confidence that the firm is not overvalued by restricting the number of their own shares they sell.¹⁹⁷

Spence makes a distinction between the signalers intrinsic ('observable, unalterable') attributes, which he calls 'indices', 'reserving the term *signals* for those observable characteristics attached to the individual that are subject by manipulation by him'.¹⁹⁸ Therefore, signals are those things done on purpose to transmit information.

Screening is the action of uninformed parties that is intended to filter suitable agents from unsuitable agents. The notion of screening has been applied outside of economics, to sociology (e.g. criminal gang recruitment practices),¹⁹⁹ law (e.g. the consequences of preliminary injunctions for final settlements),²⁰⁰ and advertising (e.g. the social significance of public relations).²⁰¹ Regarding the political, the questioning of political candidates at public meetings is an example of screening; the uninformed party (the voter) moves first to screen the suitable candidate from their less suitable competitors, before the contract is offered in the form of a vote.

3.4.2. Incentive Problems (Moral Hazard)

Whereas selection problems occur before a contract is signed, the incentive problem occurs when a party acts opportunistically after a contract is signed. It concerns hidden actions and the information asymmetry arising therefrom. In the

¹⁹⁷ Stiglitz, 'The Contributions of the Economics of Information to Twentieth Century Economics', p.1453.

¹⁹⁸ Spence, M., 'Job Market Signaling,' *The Quarterly Journal of Economics* 87.3 (1973) pp.355–74.

¹⁹⁹ Densley, J., 'Street Gang Recruitment. Signaling, Screening, and Selection,' *Social Problems* 59.3 (2012) pp.301–21.

²⁰⁰ Jeitschko, T.D. & Byung-Cheol, K., 'Signaling, Learning, and Screening Prior to Trial: Informational Implications of Preliminary Injunctions,' *The Journal of Law, Economics, & Organization* 29.5 (2013) pp.1085–1113.

²⁰¹ Halff, G. & Gregory, A., 'What Is Public Relations to Society? Toward an Economically Informed Understanding of Public Relations,' *Public Relations Review* 41.5 (2015) pp.719-725.

principal-agent problem model, incentive problems concern the risk that the agent will prioritise its goals above the principal's goals. This is also known as moral hazard.

Since the late nineteenth century, the phrase 'moral hazard' has been in use in insurance, where it was used to describe the risk of immoral behaviour, but over time its meaning has changed. Allard Dembe and Leslie Boden's history of moral hazard traces its development and credits Jacques Dreze (1961), Kenneth Arrow (1963) and Mark Pauly (1968) with pioneering its analysis in modern economics following engagement with the St Petersburg Paradox,²⁰² which was brought to the attention of American economists by George Stigler and Leonard Savage in the early 1950s. The St Petersburg Paradox is a probability problem first resolved in Daniel Bernoulli's *Specimen Theoriae Novae de Mensura Sortis* (1738), by factoring in utility, which, in economics, is a measure of usefulness in satisfying a want, and which is fundamental to the study of moral hazard. Jean-Jacques Laffont and David Martimort agree with Dembe and Boden that the first important example in the modern literature was Arrow's 1963 article:

Barnard's work emphasized the need to induce appropriate effort levels from members of the organization -the moral hazard problem- and to create authority relationships within the organization to deal with the necessary incompleteness of incentive contracts. We will then have to wait for Arrow (1963) to introduce in the literature on the control of management the idea of moral hazard borrowed from the world of insurance. This work will be further extended by Wilson (1968) and Ross (1973) who will redefine it explicitly as an *agency problem*.²⁰³

²⁰² Translated into English by Louise Sommer and published as *Exposition of a New Theory on the Measurement of Risk* in *Econometrica* in 1954. The journal's editor noted that 'In view of the frequency with which Bernoulli's famous paper has been referred to in recent economic discussion, it has been thought appropriate to make it more generally available by publishing this English version'.

²⁰³ Laffont & Martimort, *Theory of Incentives*, p.13.

Anurag Sharma points to two approaches, in the economics literature, to dealing with moral hazard:

In addressing contractual difficulties that arise from information asymmetry and anticipated agent opportunism, agency theorists recognize two basic attributes of exchange with which the principals must contend: (1) cost of monitoring and (2) cost of metering...²⁰⁴

Metering describes the concept of measuring the outcome of an action by an agent relative to the expectations of the principal. It involves looking at the end result, rather than observing the means by which the result was reached or the actions of the agent. Jennifer Coats has considered the role of asymmetric information in making decisions about outsourcing government services, proposing that the decisions can be approached according to the extent to which metering that service is possible, and proposing three broad classes of service:

Full information: goods/services for which output reflects measures of the level of provision

Asymmetric Information I: goods/services for which output/outcomes cannot reflect measures of the level of provision

Asymmetric Information II: research and development of goods/services whose outcomes reflect the level of provision, but have highly unpredictable costs.²⁰⁵

In the second and third classes, where metering is not sufficient to address information asymmetry, monitoring becomes necessary. Sharma states:

...principals can design mechanisms to monitor agent behavior and to ensure that the agent actually is behaving as stipulated in

²⁰⁴ Sharma, A., 'Professional as Agent: Knowledge Asymmetry in Agency Exchange,' *The Academy of Management Review* 22.3 (1997) p.761.

²⁰⁵ Coats, J. C., 'Applications of Principal-Agent Models to Government Contracting and Accountability Decision Making,' *International Journal of Public Administration* 25.4 (2002) p.444.

the contract. In this way the principal can reduce information asymmetry vis-a-vis the agent and can impose sanctions if the agent deviates from expected behavior.²⁰⁶

In the economics literature, monitoring is often assumed to be prohibitively costly.

3.5. Information Asymmetry in the Social Contract

The social contract is a principal-agent problem typified by information asymmetry. Interactions with and through government information have the potential to change the nature or terms of the social contract, but this potential is limited by the information asymmetry between the individual and the state. The state information bound up in paper files, databases, content management systems, surveillance recordings, intra-organisational social media platforms, and registers, indexes and manuals of all kinds, is the material manifestation of a dialogue between the individual and the state, or as Ilana Feldman puts it: ‘One of the ways filing contributes to governing authority is by defining a space and style of interaction among people, whether civil servants or members of the public. Filing delimits both the terrain of possibility and modes of objection’.²⁰⁷

The economics definitions of the problems arising from asymmetric information, and the associated measures to address them, provide a framework for thinking about how information asymmetry in the public sphere can be reduced. Before the contract, individuals may face selection problems as electors, and may offer a contract after signaling from or screening of candidates. After the contract, the moral hazard or incentive problems arising from the principal-agent relationship might be solved through metering – comparing expectations with outcomes – or monitoring – oversight of agent actions.

²⁰⁶ Sharma, ‘Professional as Agent’, p.761.

²⁰⁷ Feldman, *Governing Gaza*, p.32.

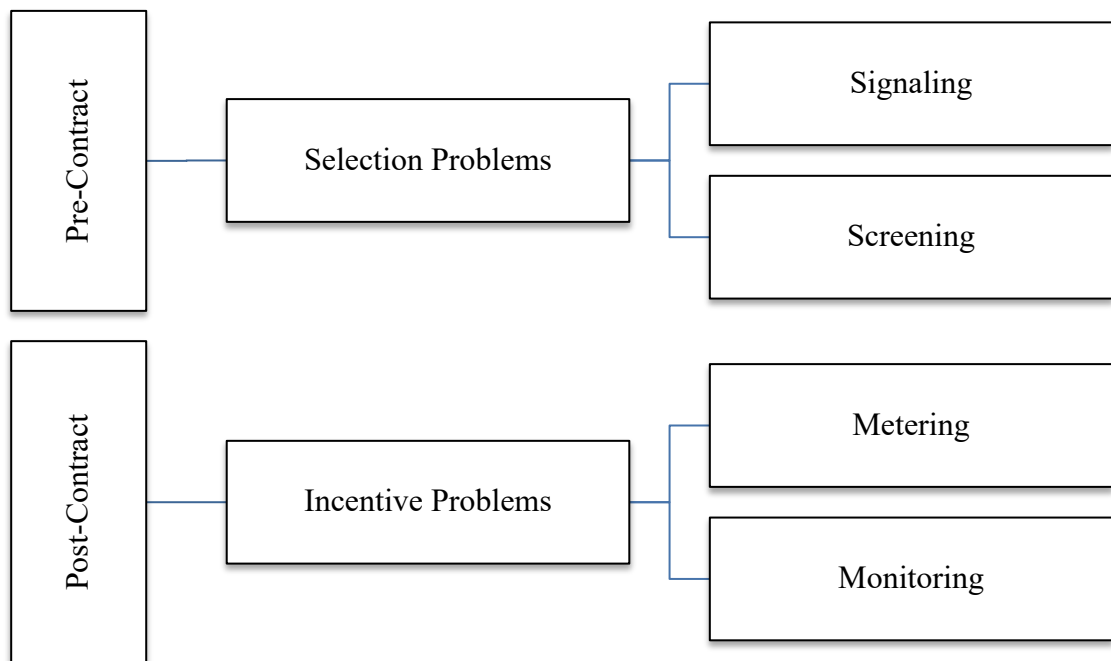


Figure 1: Information economics approaches to addressing pre- and post-contract principal-agent problems.

Individuals could be said to face selection problems when they have the opportunity to elect representatives, and could be said to apply their own screening criteria in choosing a representative, and political parties and candidates for election are certainly able to signal their abilities to act as representatives. This concerns the particular people moving into and out of particular positions, not the nature of the contract between the individual and the state, and the individuals' power or lack of power in interactions with the apparatus of bureaucracy. In this respect, individuals do not face selection problems, because these only properly arise before a contract is entered into. We are too late for selection problems because the principal-agent relationship has evolved in history. Signaling is less problematic. How do bureaucracies signal? If a signal is a purposeful transmission of information, then purposeful releases of government information, such as open data, press releases, or the opening of archives, are signals.

Concerning the incentive problem, in relation to the social contract the extent to which metering is possible varies. On a large and public scale, if a community

(principal) expects a road to be repaired, and the road is repaired, then the agent has satisfied expectations. If the community wants to know which firm was contracted to fix the road, and how that firm was chosen, it cannot rely on metering. It must be able to witness the decision-making, or evidence of it, through monitoring. The level of scrutiny facilitated by monitoring is evoked by Gupta in his study of the allocation of state pensions in India:

Paying attention to processes within the state – for example, procedures of planning and implementation, communication or lack thereof across hierarchies, the rule-following actions of particular officials – opens up possibilities of interpretation different from those we might posit were we to look only at finished outcomes....²⁰⁸

It is monitoring that facilitates accountability, in the sense of rendering an account. Post-contract, moral hazard is the primary problem arising from the information asymmetry that exists between individuals and the state, the principals and the agent, and monitoring becomes indispensable in redressing that information asymmetry.

3.6. Conclusion

Principal-agent relationships are susceptible to selection problems and incentive problems, but the historical nature of the social contract precludes individuals from confronting selection problems except in cases of sudden, often violent regime change. Instead, the individual is more likely to encounter those problems termed incentive problems in information economics. These problems may take many forms with varying degrees of significance in the lives of individuals – at the extreme end of the spectrum incentive problems result in corruption, abuse and neglect, and the erosion of rights and entitlements.

Sharma has argued that goal alignment between principals and agents is difficult because information asymmetry invariably favours agents. Information

²⁰⁸ Gupta, *Red Tape*, p.13.

economics offers ways out of the information asymmetry implicit in principal-agent problems. Signaling and screening are approaches to selection problems, and while useful in various ways in the socio-political context (for instance, in the election of representatives), they tend not to materially alter the interface at which the individual encounters the state, which is bureaucracy. In contrast, metering and monitoring, which information economics offer as solutions to incentive problems, allow individuals some insight into the activities of their agents. They do this by increasing the quantity and quality of the information available to individuals, which has implications for the power relations between individuals and the state.

Metering compares expectations or promises with outputs. In cases where metering has failed, is difficult, or is insufficient for addressing asymmetric information, monitoring becomes necessary. Monitoring, which enables principals to see what agents are doing or have done, requires mechanisms that may be difficult and expensive to establish and operate, and that may be resisted by agents, as I will show in the next chapter, which reviews the history of the emergence of monitoring mechanisms in the United Kingdom's public sector.

Finally, we can agree with Kregg Hetherington's statement, in which there are traces of Jacques Ranciere's constitutional marketplace:

Transparency as the key tool of political reform in the post-Cold War era built on a longer discussion in economics about the role information plays in markets. Transparency made it possible for citizens to make better decisions about their governments, only because politics now worked like a market, in which it had long been argued that better distributions of information made economic actors more rational.²⁰⁹

²⁰⁹ Hetherington, *Guerrilla Auditors*, p.156.

Chapter 4

Opening Government: A History of Monitoring

Archival culture is state politics; the keeping of records, granting of access to them, and denial of access to them lie at the heart of all systems of government.²¹⁰

Kirsten Weld, *Paper Cadavers: The Archives of Dictatorship in Guatemala*

²¹⁰ Weld, *Paper Cadavers*, p.52.

4.1. Introduction

When viewed as a monitoring mechanism, rather than as a resource for the development of new services and products, open government data should be seen in a broader context of government secrecy and openness, as its endorsement and promotion by government follows a long history of contention around access to public sector information. This chapter will provide an overview of the historical development of secrecy and openness in the government of the United Kingdom, and its practical manifestation as approaches and methods for public monitoring of government as an agent. The chapter will look at the development of a culture of secrecy in the British public sector, observing two themes in the existing literature that position the formalisation of government secrecy as a consequence of new information and communication technologies - first record-making on paper and later digitally - and as a consequence of the socio-economic factors (primarily colonial expansion and the growth of the welfare state) that necessitated the opening of public sector employment to lower socio-economic classes than had previously been engaged in government work and, as a number of scholars argue, conditioned to discretion by their class privilege. The chapter argues that the culture of secrecy in British government was transplanted to colonial administrations and has therefore been typical of government in Commonwealth countries, though information asymmetry is characteristic of all principal-agent relations. The chapter observes the institutionalisation of secrecy through bureaucratic instruments, Official Secrets legislation, and the Defence Notice (D-Notice) system and, as the face of sovereign power has changed, the move towards openness first through archival access to records, then whistleblowing protection, Freedom of Information, and most recently open data and the Open Government movement. The key contributions of the chapter are to show how information access relates to features of the social contract, particularly its participatory aspect, and how monitoring mechanisms work in concert and towards dissensus.

4.2. The Development of a Culture of Secrecy in Government

The ancient concept of *arcana imperii*, found in Tacitus and since interpreted variously as ‘state secrets’ or the hidden knowledge required to rule, was defended by Jean Bodin in France and Robert Filmer in England as important to kingly authority and the integrity of the state.²¹¹ Lawrence Quill has shown how secrecy has been essential to state power since the emergence of the modern state in the sixteenth century,²¹² citing discussions about the importance of the competitive acquisition of knowledge by states about other states and the *demos* in work from Hobbes to Foucault, and describing the modern ‘surveillance society’ as intrinsically secretive and permanently contested:

In fact, the power of the state, understood as the monopoly over secrets and secret collection, is proportional to the relative powerlessness of the citizens and their inability to keep secrets or discover those they are not supposed to know. This power, I would argue, is in a constant state of flux. Despite the relative permanence and durability of state institutions and practices, the (unofficial) leak, the data breach, and the actions of the whistle-blower all threaten the integrity of states and are... usually dealt with in brutal fashion.²¹³

The formalisation of state secrecy – both in terms of a demarcation of knowledge that is secret or may be open and the demarcation of behaviours and processes around information as legal or illegal - has been spurred on by developments in technologies of information and communication. Discussing pre-revolutionary France, Roberts wrote that as ‘a practical matter, secrecy was easily preserved at a time when the only method of distributing information was by manuscript (that is, handwritten) texts’.²¹⁴ Roberts notes that the advent of the printing press coincided with France’s ‘law of silence’ of 1764: advances in technologies for the reproduction and dissemination of information generated a legislative response, which is a pattern that repeats in the United Kingdom across generations of

²¹¹ Roberts, A., *Blacked Out: Government Secrecy in the Information Age* (New York: Cambridge University Press, 2008) p.10.

²¹² Quill, L., *Secrets and Democracy: From Arcana Imperii to WikiLeaks* (New York: Palgrave Macmillan, 2014) p.68.

²¹³ *Ibid.*, p.69.

²¹⁴ *Ibid.*, p.9.

technology. Government secrecy, which Richard Crossman called ‘the British disease’, has old roots in Britain,²¹⁵ which Ben Worthy traces to the Privy Councillor’s Oath instituted in 1250.²¹⁶ However, it was in the Victorian era that a culture of government secrecy was codified.²¹⁷ Following Robert’s observation of the coincidence of printing and the French ‘law of silence’, the first Official Secrets Act in Britain was passed in 1889, at the end of the decade in which the typewriter was introduced into office work and gave a new application to carbon copying.²¹⁸

In the British context, the history of state secrecy has also been tied to a history of class. David Vincent discusses secrecy as a common value of governance before the reforms that followed the Northcote-Trevelyan Report of 1854, which opened public sector employment to lower socio-economic classes of people than had previously been involved in public administration.²¹⁹ Vincent uses terms such as ‘gentlemanly self-restraint’ and ‘honourable secrecy’ to describe expectations of official conduct in this period.²²⁰ He wrote that the ‘issue of official secrecy legislation stemmed in large part from the growing requirement to recruit to government offices those whose breeding, education, and pay excluded them from the rank of gentlemen’.²²¹ Christopher Moran acknowledges both the technologies of information production and dissemination and the changing class composition of the British public service as two related challenges to secrecy in the Victorian period.²²² The growth in the volume of information committed to paper, and the concomitant need for staff to manage the papers, ‘at the secretarial and menial levels’ resulted in a perceived need for tighter controls over official information.²²³ ‘Unlike their superiors, these individuals had not been brought up in country homes or been educated at public schools and then

²¹⁵ Michael, J., *The Politics of Secrecy: Confidential Government and the Public Right to Know* (London: Penguin, 1982) p.12.

²¹⁶ Worthy, B., *The Politics of Freedom of Information: How and Why Governments Pass Laws that Threaten their Power* (Manchester: Manchester University Press, 2017) p.17.

²¹⁷ *Ibid.*, p.17.

²¹⁸ Proudfoot, W. B., *The origin of stencil duplicating* (London: Hutchinson & Co. Ltd., 1972).

²¹⁹ Vincent, D., *The Culture of Secrecy: Britain, 1832-1998* (Oxford: Oxford University Press, 1998).

²²⁰ *Ibid.*

²²¹ *Ibid.*, p.91.

²²² Moran, C., *Classified: Secrecy and the State in Modern Britain* (Cambridge: Cambridge University Press, 2013).

²²³ *Ibid.*, p.31.

Oxbridge. For this reason, serious doubts existed about whether they could be trusted to maintain a judicious silence about their work'.²²⁴ A different view is taken by K.G. Robertson, who sees secrecy as one of the values comprising the neutral civil service envisaged by the Northcote-Trevelyan report.²²⁵ According to Robertson, the 1873 Treasury Minute on the 'Premature Disclosure of Official Documents', and subsequent Minutes in 1875 and 1884 show the 'civil service mentality as in the process of creation, in the appeal to the ethics appropriate to a profession of anonymous, silent, dedicated men who know their position and know that publicity has no part in it'.²²⁶

Moran argues that by the late 1880s, 'incidents of unauthorised disclosure by penumbral classes in Whitehall had greatly increased, despite a further round of circulars calling on officials to be honourable and disciplined'.²²⁷ The cases cited by Moran suggest that the motivation for these disclosures was often financial advantage rather than political allegiance or public-minded whistleblowing. Moran and Vincent agree that these disclosures drove the passage of the 1889 Official Secrets Act, though in considering the true role of class in the evolution of official secrets legislation, it is worth noting that many histories of British government secrecy begin with the 1833 injunction the Foreign Office obtained against the sale of Lord Hanley's papers, and Robertson describes a series of sales or attempted sales of the memoirs, letters and diaries of diplomats between 1830 and 1870.²²⁸

Moving into the twentieth century, the codification of government secrecy continued in parallel with a growth in the volume of information gathered and stored by government. Vincent describes the increase in the collection of information about citizens' lives, particularly working-class people as the welfare state grew: 'A record office at Kew soon contained details of the employment histories of almost 12 million citizens.'²²⁹ It was the growing tension between

²²⁴ Ibid., pp.31-32.

²²⁵ Robertson, K. G., *Public Secrets: A Study in the Development of Government Secrecy*, (London: MacMillan Press, 1982).

²²⁶ Ibid., p.53.

²²⁷ Moran, *Classified*, p.35.

²²⁸ Robertson, *Public Secrets*, p.46.

²²⁹ Vincent, *The Culture of Secrecy*, p.141.

Britain and Germany, rather than a concern for the privacy of British citizens, that would provide the impetus for the passage of a new, stricter Official Secrets Act (1911), however.²³⁰ In 1909, the subcommittee on foreign espionage of the Committee of Imperial Defence had started researching the subject, and drafted a Bill in 1910.²³¹ Hooper suggests that it was the deployment of a German gunboat to Agadir, Morocco, in 1911 that prompted the introduction of the Bill to Parliament.²³² Vincent observed that the law 'represented both a major turning-point in the state's acquisition of knowledge about its population, and the consolidation of the Victorian approach to the policing of official information'.²³³ He describes the relationship between citizen and state at this time as an 'unstable mixture of reluctant trust and incipient paranoia'²³⁴ and that:

The state sought information covertly because it feared illegitimate protest, and overtly because it accepted the legitimacy of at least some of the needs of the working-class family economy. The balance between what the public had a right to keep secret from the state and vice versa had to be renegotiated. The subjects of the growing domestic archive, who increasingly encountered their lives as objectified knowledge, were prepared to confide in a state which displayed a new respect for their personal dignity, but were inclined to mistrust both the methods and the ends of the inspection. The only controls lay in the *ad hoc* administrative regulations which were introduced in accordance with the bureaucratic cast of mind which underpinned and was in turn confirmed by the Official Secrets Act.²³⁵

The effects of the legislation, according to Ronald Wraith, included 'an attitude of mind whereby the citizen in search of information is expected to prove that he

²³⁰ See Maer, L. & Gay, O., *Official Secrecy* (House of Commons Library, 2008) <https://fas.org/irp/world/uk/secrecy.pdf> [accessed 25 April 2018]

²³¹ Hooper, D., *Official Secrets: The Use and Abuse of the Act* (London, Secker and Warburg, 1987) p.28.

²³² *Ibid.*

²³³ Vincent, *The Culture of Secrecy*, p.141.

²³⁴ *Ibid.*, p.142.

²³⁵ *Ibid.*

ought to have it, rather than the official should prove that he ought *not* to have it.²³⁶

Efforts to control the flow of information extended beyond controlling the officials who formed the porous boundary of bureaucracy through secrecy legislation, to mechanisms for muting the press, in another example of the operation of the record-as-command. The Services, Press and Broadcasting Committee was established in 1912 to issue D-Notices to the media.²³⁷ These notices were letters circulated to newspaper editors requesting that they not publish material that concerned defence matters. This was a voluntary system with no legal penalties for non-compliance, which evokes notions of participation in power relations, valorising and control information and the co-option of subjects into the state supersubject.²³⁸ According to its official historian, Nicholas Wilkinson, the D-Notice System ‘evolved slowly, in response to a confluence of trends in the late nineteenth and early twentieth Centuries’ including trends in geopolitics, society, technology and economics.²³⁹ The system exists today as the Defence and Security Media Advisory System.²⁴⁰

In 1920, an amendment to the Official Secrets Act criminalised a number of information-related activities, including falsification of reports, forgery, retaining documents contrary to duty and communicating secret passwords.²⁴¹ These offences show a concern not only for secrecy but for the threat of misinformation within government and publicly. The D-Notice system was an attempt to constrain the movement of official information beyond the boundaries of the state; another

²³⁶ Wraith, R. E., ‘United Kingdom’, in Rowat, D.C. (ed.), *Administrative Secrecy in Developed Countries* (London: MacMillan, 1979) p.185.

²³⁷ Moran, *Classified*.

²³⁸ Robertson, *Public Secrets*, p.75.

²³⁹ Wilkinson, N., *Secrecy and the Media: The Official History of the United Kingdom’s D-Notice System* (London: Routledge, 2009) p.3.

²⁴⁰ In 2015, the D-Notice system was renamed the Defence and Security Media Advisory (DSMA) System. It is overseen by a committee of government and media members and has five standing notices, covering: DSMA-Notice 01: Military Operations, Plans & Capabilities, DSMA-Notice 02: Nuclear & Non-Nuclear Weapon Systems & Equipment, DSMA-Notice 03: Military Counter-Terrorist Forces, Special Forces and Intelligence Agency Operations, Activities and Communication Methods and Techniques, DSMA-Notice 04: Physical Property and Assets, DSMA-Notice 05: Personnel and their Families who work in Sensitive Positions. See <http://www.dsma.uk/danotices/index.htm> (last accessed 14 September 2017).

²⁴¹ Robertson, *Public Secrets*, p.70.

attempt to control information in the public space took the form of proactive disclosure through authorised histories. As Moran has argued, 'by the 1960s the state had concluded that maintaining absolute secrecy with respect to some of its work was not only impossible but counterproductive' and authorised official histories as part of "offensive" information management, putting "secrets" into the public domain on its own terms.²⁴² Nevertheless, secrecy legislation continued to be reviewed and refined. A 1939 revision had clarified powers of interrogation under the Act.²⁴³ Between 1971 and 1972, Lord Franks chaired the Committee on Official Secrets Act Section 2, which found 'authorisation' and 'interests of the State' were not defined.²⁴⁴ The problems of these definitions can be interpreted as a problem of sovereign power and the state of exception, because as Robertson notes, there is an assumption that authorisation 'flows from the nature of the particular post' and that 'Ultimately... if the Minister is self-authorising, then the reason for secrecy can only be that the Minister or cabinet have decided not to authorise the release of information and this is the reason for secrecy'.²⁴⁵ This circularity reinforces ministerial power and its ability to exclude others from information, power and political participation.

The Official Secrets Act of 1989 repealed section 2 of the 1911 Act, removing the public interest defence that allowed for the disclosure of official information in the public interest, which was to be dealt with under the Public Information Disclosure Act (1998), discussed in relation to whistleblowing, below.

4.3. Archival Access

Like the connection between writing and the law, the significance of archival access for understanding and characterising principal / agent power relations is well-established both in the archival studies and political science literature.

Francis Blouin and William Rosenberg, writing about changing perceptions of archives in the 1800s, argued that '...the concept of popular sovereignty had not

²⁴² Moran, *Classified*, p.5.

²⁴³ Maer & Gay, *Official Secrecy*.

²⁴⁴ White, P., 'Official Secrets and Government Openness in Britain', *The Australian Library Journal*, 29:1, 1980, p.21.

²⁴⁵ Robertson, *Public Secrets*, p.85.

only shifted the notion of privacy in the archives from the state to the public, validating secrecy only if it could be justified as in the public interest, but also expanded the legitimating role of the archives into one that embraced accountability.²⁴⁶ The French Revolution is widely cited as the moment that archives were reconceived as public resources for enacting democracy, though Jennifer Milligan points out that the reality was more complex than a reading of the law might suggest: 'The granting of citizens' access to the Archives in Revolutionary legislation suggests that possibilities for a democratic and transparent archives were at least theorized and even legislated if only to be ultimately unrealized'.²⁴⁷ Less has been made, by archival scholars, of an earlier English case: during the Interregnum, 'An Act for turning the Books of the Law, and all Proces [sic] and Proceedings in Courts of Justice, into English' was passed, prohibiting the use of Latin and law French in legal record-making.²⁴⁸ The motivation for creating records in the common tongue, and thereby opening up the legal process, finds its opposite number in the return to Latin record-making with the Restoration of Charles II.²⁴⁹ Kirsten Weld claims that this connection between access and democracy may now be taken for granted, and calls for closer analyses of the phenomenon:

Scholars of archival and political science often link the accessibility of state archives to that state's levels of accountability; the more archival access, the more democracy. This observation has some truth to it, but it is not particularly interesting. Such a simple correlative elides the richness of information about a state's character – and the fabric of the relationships between state and citizenry – to be gleaned, historically and ethnographically, from the attitudes that state fosters about archives.²⁵⁰

²⁴⁶ Blouin, F. & Rosenberg, W., *Processing the Past: Contesting Authority in History and the Archives* (New York: Oxford University Press, 2011) p.22.

²⁴⁷ Milligan, S., 'The Problem of *Publicite* in the Archives of Second Empire France', in Blouin, F. & Rosenberg, W. (eds.), *Archives, Documentation and Institutions of Social Memory: Essays from the Sawyer Seminar* (Ann Arbor: University of Michigan Press, 2010) p.21.

²⁴⁸ Duncan, G., *The High Court of Delegates* (Cambridge: Cambridge University Press, 1971) p.214.

²⁴⁹ Stryker-Rodda, H., *Understanding Colonial Handwriting*, (Baltimore: Genealogical Publishing Co. Inc., 1986) p.9.

²⁵⁰ Weld, *Paper Cadavers*, p.51.

Such a study in the UK might find parallels between departmental discretion and uncoordinated record-keeping. A significant impediment to access to UK government information, historically, has been the lack of information about what information is held by government and where and how it might be made available. As Elizabeth Shepherd has shown, 'At the end of the eighteenth century public records were scattered between sixty buildings in London and Westminster...'.²⁵¹ Shepherd charts the movement towards the centralisation of government records through the nineteenth century. Following the establishment of the Public Record Office (PRO) with the *Public Record Office Act* of 1838, Sir Francis Palgrave was appointed the first Deputy Keeper.²⁵² Public opening hours were established and in Palgrave's structuring of the PRO, the Archival Department was charged with making inventories, catalogues and calendars to facilitate access to the records, though there was no statutory right of access at that time.²⁵³ The Public Record Office building on Chancery Lane, London, opened in 1858 and public reading rooms were opened in 1866,²⁵⁴ providing access to transferred records in line with any access restrictions set by the transferring department.²⁵⁵ These arrangements concerned records transferred from government to the PRO for permanent preservation; records in active use in government continued to be inaccessible to the public, in principle.

In 1903, rules on public access to records were formalised after a cabinet committee found a wide diversity of access provisions across government: '...the Admiralty and the War Office refused access after 1830, the Colonial Office – for the most part – after 1802, the Foreign Office after 1780, the Home Office after 1778 and the Treasury winning the race with no documents released to the public after 1759'.²⁵⁶ Robertson observes that the motivations for these access policies can only be guessed at, but that 'it seems to be the case that the older departments, the ones with the closest links to the crown, most clearly exemplify

²⁵¹ Shepherd, E., *Archives and Archivists in 20th Century England* (Farnham: Ashgate, 2009) p.22.

²⁵² Ibid.

²⁵³ Ibid., p.66.

²⁵⁴ Moran, *Classified*, p.27.

²⁵⁵ Roper, M., 'Access to Public Records', in Chapman, R. A. & Hunt, M., *Open Government: A study of the prospects of open government within the limitations of the British political system* (Abingdon: Routledge, 2011) p.83.

²⁵⁶ Robertson, *Public Secrets*, p.61.

a pre-democratic attitude of denial of access'.²⁵⁷ The 1903 rules follow a logic similar to those of today's Freedom of Information exemptions, with a concern for personal privacy and state security, though given that the War Office was one of the most open, Robertson concludes that national security was not necessarily the motivation for the new rules.²⁵⁸ Nevertheless there was some concern for secrecy in relation to security: most importantly, the Prime Minister's establishment, in 1909, of the Sub-committee on Foreign Espionage of the Committee of Imperial Defence, that would draft a more punitive Official Secrets Act in anticipation of German espionage and aggression, as mentioned above.

In 1954, the Grigg Committee on Departmental Records recommended a 50 year rule, which appeared in the Public Records Act 1958; records deemed to have enduring value would be transferred to the National Archives when they were 30 years old, and made available to the public when they were 50 years old.²⁵⁹ The 50 year rule concerned archives only, and as Norman Marsh noted in the introduction to his *Public Access to Government-Held Information*, '...the right of access to historical records can exist with or without any right of public access to government-held information'.²⁶⁰ Writing prior to the enactment of Freedom of Information law in Britain, McDonald observed two approaches to the treatment of archives and access in law. In the first, 'each country has archival legislation but it is not concerned with access arrangements.'²⁶¹ He cites Sweden and Canada as examples where this approach is taken, and where the rules established by access laws apply equally to the archives and the agencies of government creating and using current records.²⁶² In the second approach, 'the age of documents determines which of two consecutive access provisions applies.'²⁶³ Britain, Australia and New Zealand are given as examples of countries that take this approach.²⁶⁴ The 1967 revision to the Public Records Act

²⁵⁷ Ibid.

²⁵⁸ Ibid.

²⁵⁹ Shepherd, *Archives and Archivists in 20th Century England*, p.44.

²⁶⁰ Marsh, N., 'Access to Government-Held Information: An Introduction', in Marsh, N. (ed.), *Public Access to Government-Held Information* (London: Stevens and Sons, 1987) p.26.

²⁶¹ McDonald, A., 'Archives and Open Government', in McDonald, A. & Terrill, G. (eds.), *Open Government: Freedom of Information and Privacy* (Basingstoke: Macmillan Press, 1998) p.29.

²⁶² Ibid., p.32.

²⁶³ Ibid., p.29.

²⁶⁴ Ibid.

reduced the 50 year closure period to 30 years (the 30 year rule), a change that Andrew McDonald attributes partly to the pressure brought by the Oxford and Cambridge Group since 1962.²⁶⁵ As discussed below, the 30 year rule has since become a 20 year rule. The Public Records Act allows some variations to these rules, such as retention in departments longer than 30 years, with the Lord Chancellor's approval, for example through the Lord Chancellor's Security and Intelligence Instrument.²⁶⁶ Early opening is also made possible through Lord Chancellor's Instruments.²⁶⁷

Certain Acts of Parliament forbid the disclosure of information collected under their provisions, which are known as 'statute-barred records'.²⁶⁸ Roper cites the Agricultural Statistics Act 1979, the Land Registration Act 1925, the Statistics of Trade Act 1947 and the Competition Act 1980 as examples of Acts that bar the disclosure of certain types of records.²⁶⁹ Other Acts require the publication or disclosure of certain records, such as the register of nurses maintained by the Nursing and Midwifery Council, which must be made accessible to members of the public under Part III.8 of the Nursing and Midwifery Order 2001.²⁷⁰ As discussed below, the advent of Freedom of Information legislation in the United Kingdom would reduce the significance of the various rules for access to records. The Freedom of Information law was interpreted to cut across the 30-year rule in the sense that FOI applications made to the National Archives could result in the release of closed information before the expiration of the 30 year closure period.²⁷¹

4.4. Whistleblowing

²⁶⁵ Ibid., p.28.

²⁶⁶ *Notes on the Lord Chancellor's Security and Intelligence Instrument*, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/219905/notes-security-intelligence-instrument.pdf [last accessed 12 March 2019].

²⁶⁷ Ibid.

²⁶⁸ Roper, 'Access to Public Records', p.87.

²⁶⁹ Ibid.

²⁷⁰ Nursing and Midwifery Order 2001, part III.8, <https://www.legislation.gov.uk/uksi/2002/253/contents/made> [last accessed 12 March 2019].

²⁷¹ The National Archives, *Making a Freedom of Information Request* <http://www.nationalarchives.gov.uk/about/freedom-of-information/making-a-freedom-of-information-request/> [last accessed 12 March 2019].

In 1998, in tandem with revisions to the Official Secrets Act, whistleblowing protection legislation was introduced in the UK.²⁷² Throughout the modern history of government, there have been instances of public servants leaking information to the press, either as a political manoeuvre or in order to reveal illegal or improper conduct. This section will provide a brief overview of the latter as it concerns the release of information as an impulse to monitor. Information theft and cybercrime will be excluded for the same reason: these are not concerned with monitoring, but have other motivations, such as extortion, as in the Banque Cantonale de Geneve case, malicious pranking, espionage, or geopolitical motivations, as in the Sony Pictures hack following the production of the film *The Interview*.²⁷³ So-called 'ethical hacking' is an approach to testing and strengthening cybersecurity is also outside the scope of this section. A large and growing body of literature on whistleblowing includes analyses of legal frameworks for defining and protecting whistleblowers and understanding the motivations of whistleblowers and the responses of institutions. This literature suggests various motivations for leaking government or corporate information, from a concern for the public interest to commercial sabotage and political dissidence. Whistleblowing, then, can be considered as a kind of impertinent dialectic in some contexts, because it disrupts the operation of the record- or data-as-command. This research is concerned here with whistleblowing involving insiders releasing information beyond the bounds of the body corporate, with a motivation rooted in the concept of 'the public interest', and will use Marcia Miceli and Janet Near's widely accepted definition of whistleblowing: 'the disclosure of organizational members (former or current) of the illegal, immoral or illegitimate practices under the control of their employers to persons or organizations that may be able to effect action.'²⁷⁴

²⁷² Public Interest Disclosure Act 1998 <https://www.legislation.gov.uk/ukpga/1998/23/contents> [last accessed 12 March 2019].

²⁷³ Reuters, 'Hacker posts client emails from Swiss bank BCGE' 9 January 2015 <https://uk.reuters.com/article/us-bc-geneve-hacker/hacker-posts-client-emails-from-swiss-bank-bcge-idUKKBNOKI1MK20150109> [last accessed 12 March 2019]; BBC News, 'The Interview: A guide to the cyberattack on Hollywood', 29 December 2014 <https://www.bbc.co.uk/news/entertainment-arts-30512032> [last accessed 12 March 2019].

²⁷⁴ Near, J. P. & Miceli, M. P., 'Organizational dissidence: The case of whistle-blowing', *Journal of Business Ethics*, 1.4 (1982) p.4.

Gregor Thüsing and Gerrit Forst state that international interest in whistleblowing 'was originally triggered by the global fight against corruption in the wake of the early 2000s recession and particularly the US-American *Sarbanes-Oxley Act*'.²⁷⁵ However, British responses to whistleblowing came earlier. In the UK, the Public Interest Disclosure Act (PIDA) was passed in 1998, following a need for legislation that 'was graphically demonstrated by the revelations that several maritime and rail disasters and the consequent loss of life could have been avoided had employees been allowed to voice their concerns earlier'; Gehan Gunasekara cites the capsizing of the Herald of Free Enterprise, the Piper Alpha oil platform disaster and the Clapham Junction train disaster as examples.²⁷⁶ Under common law, there had long been a public interest defence against charges of breach of confidence, but the Public Interest Disclosure Act offered protection with respect to unfair dismissal for 'qualifying disclosures', which are 'any disclosure of information which, in the reasonable belief of the worker making the disclosure, tends to show one or more of the following—

- (a) that a criminal offence has been committed, is being committed or is likely to be committed,
- (b) that a person has failed, is failing or is likely to fail to comply with any legal obligation to which he is subject,
- (c) that a miscarriage of justice has occurred, is occurring or is likely to occur,
- (d) that the health or safety of any individual has been, is being or is likely to be endangered,
- (e) that the environment has been, is being or is likely to be damaged, or
- (f) that information tending to show any matter falling within any one of the preceding paragraphs has been, is being or is likely to be deliberately concealed.²⁷⁷

²⁷⁵ Thüsing, G. & Forst, G., 'Whistleblowing Around the World: A Comparative Analysis of Whistleblowing in 23 Countries' in Thüsing, G. & Forst, G. (eds.), *Whistleblowing: A Comparative Study* (Cham: Springer, 2016) p.5.

²⁷⁶ Gunasekara, G., 'Whistle-blowing: New Zealand and UK solutions to a Common Problem', *Statute Law Review*, 24.1 (2003) p.39.

²⁷⁷ Public Interest Disclosure Act 1998, S 43B(1)

<https://www.legislation.gov.uk/ukpga/1998/23/section/1> [accessed 14 September 2018].

Public Concern at Work (PCAW), the UK's whistleblowing charity, compiled an analysis of the 1200 claims of unfair dismissal or mistreatment registered in the first three years of the Act. PCAW found that two thirds of cases were settled or withdrawn without a public hearing, and of those going to hearing, '54% of claimants lost, 23% won under other employment or discrimination law and 23% won under PIDA'.²⁷⁸ The cases won by claimants concern a cross-section of public, private and third sector bodies, and matters ranging across physical and sexual assault, discrimination in employment and promotion, public health and food safety and improper conduct in financial trading and the delivery of public services.²⁷⁹

Whistleblowing returned to international attention in 2010 following the publication by Julian Assange's Wikileaks website of information received from Chelsea Manning concerning military incursions in the Middle East and particularly the resulting civilian deaths.²⁸⁰ This was followed in 2013 by Edward Snowden's exposure of surveillance operations through the release of classified information from the US' National Security Agency.²⁸¹ These cases, and in particular the popular and official responses to them, illustrate that whistleblowing often occurs in legal and ethical grey areas, perhaps outside of the law but perhaps within a regime of ethics, whether or not enshrined in professional or organisational codes of conduct. This ambiguity in how information releases from whistleblowers should be read in the face of the intricacies of sometimes conflicting national laws and international concords on whistleblowing, freedom of expression or speech, confidentiality, official secrets and national security, is mirrored in the ambiguity in the informational processes that whistleblowing can unveil. Carmen Apaza and Yongjin Chang have suggested that as 'insiders, whistleblowers are the source of valuable information that neither the government nor the public can get from

²⁷⁸ Public Concern at Work, *Whistleblowing Case Studies* (2003) <http://www.pcaw.org.uk/law-policy/legal-updates/whistleblowing-case-summaries> [accessed 4 September 2018].

²⁷⁹ Ibid.

²⁸⁰ Manning, C. E., 'The years since I was jailed for releasing the 'war diaries' have been a rollercoaster', *The Guardian*, 27 May 2015 <https://www.theguardian.com/commentisfree/2015/may/27/anniversary-chelsea-manning-arrest-war-diaries> [last accessed 12 March 2019].

²⁸¹ Finn, P. & Horwitz, S., 'US charges Snowden with espionage', *The Washington Post*, 21 June 2015 https://www.washingtonpost.com/world/national-security/us-charges-snowden-with-espionage/2013/06/21/507497d8-dab1-11e2-a016-92547bf094cc_story.html?utm_term=.6e05592680a7 [last accessed 19 March 2019].

oversight systems.²⁸² As discussed below, Freedom of Information is process-oriented and is epistemologically aligned with process-bound record-making and record-keeping. Whistleblowing makes public information that might be concealed or destroyed, or it may include information that is not recorded. It may include observations of behaviour, the witnessing of verbal or physical events that happen outside of record-making activities and bureaucratic workflows and may involve, as its primary or secondary concern, whistleblowing on failures to meet the duty to document or breaches of documentary custodianship. In this way, whistleblowing is a mechanism that can reinforce and safeguard measures such as Freedom of Information. Whistleblowing legislation such as the Public Interest Disclosure Act offers a level of protection that is a monitoring mechanism in itself and which enriches other monitoring mechanisms. It is a form of dissensuality that engenders the dissensual.

4.5. Freedom of Information

Two years after PIDA was passed, Freedom of Information legislation was also introduced in the UK.²⁸³ Ben Worthy has written a comprehensive history of the progress of freedom of information legislation in the UK.²⁸⁴ He takes Colin Darch and Peter Underwood's premise that the movement towards openness has its roots in 'two revolutionary processes, one philosophical and one technological'²⁸⁵; a desire for knowledge in the European Renaissance and the means for distributing information (the press) and identifies antecedents to the modern Freedom of Information movement in the English pamphleteers of the period of the English Revolution and the True Levellers' ideas about social organisation.²⁸⁶ Moving forward to the early twentieth century, Worthy sees Woodrow Wilson's presidential campaign of 1912 and Leon Trotsky's publication of secret treaties with Allied powers in 1918 as significant turning points in the progress of

²⁸² Apaza, C. R. & Chang, Y., 'Effective Whistleblowing Conceptual Framework', in Apaza, C. R. & Chang, Y. (eds.), *Whistleblowing in the World: Government Policy, Mass Media, and the Law* (Cham: Palgrave Macmillan, 2017) p.1.

²⁸³ Freedom of Information Act 2000 <https://www.legislation.gov.uk/ukpga/2000/36/contents> [last accessed 19 March 2019].

²⁸⁴ Worthy, B., *The Politics of Freedom of Information: How and Why Governments Pass Laws that Threaten their Power* (Manchester: Manchester University Press, 2017).

²⁸⁵ Worthy, *The Politics of Freedom of Information*, p.4.

²⁸⁶ *Ibid.*, p.6.

government openness; progress that would pick up speed in the 1960s and 1970s.²⁸⁷

Steps towards openness around this period, as Ronald Wraith has identified, include public inquiries into planning proposals; the 1967 establishment of the Parliamentary Commissioner for Administration which could order the release of information if a citizen could 'establish a *prima facie* grievance on grounds of maladministration'; legislation making similar provisions for the national health service and local government in 1973 and 1974; and the ruling in *Conway v. Rimmer* (1968) that modified the doctrine of Crown privilege 'whereby, on the sole authority of a minister, official information could be withheld, on the ground that its disclosure would not be in the public interest, even though its discovery was important for the defence',²⁸⁸ doing something to break the loop of ministerial self-authorisation described by Robertson.²⁸⁹ That the late 1960s and the 1970s were a particularly eventful period in the history of openness in the UK is borne out by Worthy's tabulation of the mentions of 'freedom of information' and 'open government' in the UK parliament from 1880 to 2000. His table shows a jump in mentions of FOI from two in the 1960s to 256 in the 1970s, and a jump in mentions of open government from four in the 1960s to 746 in the 1970s.²⁹⁰ Both show upwards trends through the 1980s and 1990s before dropping in the 2000s.²⁹¹ Edward Higgs attributes widening public concern for official information from the 1960s onwards in part to the 'sheer proliferation of contacts between citizens and civil servants that took place in the twentieth century. Given the amount of form filling that now became part of everyday life, it was impossible for citizens to be unaware of the amount of material held by government officials'.²⁹²

Worthy wrote that the 'first concrete reforms began in 1968 when the Fulton Report into reform of the Civil Service' proposed reducing secrecy.²⁹³ The 1968

²⁸⁷ Ibid., pp.7-8.

²⁸⁸ Wraith, 'United Kingdom', p.186.

²⁸⁹ Robertson, *Public Secrets*, p.85.

²⁹⁰ Worthy, *The Politics of Freedom of Information*, p.24.

²⁹¹ Ibid.

²⁹² Higgs, E., *The Information State in England: The Central Collection of Information on Citizens since 1500* (Basingstoke: Palgrave Macmillan, 2004) p.168.

²⁹³ Worthy, *The Politics of Freedom of Information*, p.24.

Fulton Report on the structure and operation of the Civil Service called for greater openness and more community consultation.²⁹⁴ As discussed above, the previous year saw the revision of the Public Records Act that changed the fifty year rule to a thirty year rule. Importantly, for thinking about secrecy and openness in the Commonwealth context, this move towards openness coincided with the end of the colonial period, the transfer of the 'British disease' having been effected in the preceding one hundred years. At the end of the colonial period, a culture of secrecy native to the UK civil service had been inculcated in the (now former) colonies. From this period onwards, the pace and the route of the journey towards greater openness would vary widely across the Commonwealth, but would always take place in the context of an administrative tradition with cultural and personal dimensions: civil servants 'are motivated... by the traditions of their service, and by thought for their own careers.'²⁹⁵ The need for cultural as well as legislative change is a recurrent theme in public and third sector statements about openness, as Wraith points out in connection with the Franks Report, the result of an Inquiry into the Official Secrets Act.²⁹⁶

In 1969, the *Information and the Public Interest* White Paper was published. It stated the value of secrecy for policy-making and claimed that Official Secrets legislation was not a barrier to openness, since 'authorised' disclosure could be extended without altering the law.²⁹⁷ The Franks Report of 1972, however, recommended reform to the Official Secrets Act.²⁹⁸ The Labour Party manifesto of 1974 pledged to reform the Official Secrets Act and pass FOI legislation, but despite returning to government, the Party was unable to make progress on openness in the face of institutional inertia in government.²⁹⁹ 1977 was a turning point, when 'Backbench pressure manifested itself', and on 6 July 1977 the Head of the Civil Service issued a memorandum called 'Disclosure of Official

²⁹⁴ Fulton Committee, *The Civil Service, Vol. 1: Report of the Committee 1966-1968, Cmnd. 3638*, (London: Her Majesty's Stationery Office, 1968)

<http://www.civilservant.org.uk/library/fulton/fulton8.pdf> [accessed 4 September 2018] p.91.

²⁹⁵ Wraith, 'United Kingdom', p.211.

²⁹⁶ *Ibid.*, p.215.

²⁹⁷ *Ibid.*, p.188.

²⁹⁸ Worthy, *The Politics of Freedom of Information*, pp.25-26.

²⁹⁹ *Ibid.*

Information' (the Croham Directive), calling for the release of information upon which policy is based.³⁰⁰

In June 1978, Justice, a branch of the International Commission of Jurists, published its *Freedom of Information* report, calling for new guidelines on information disclosure. In July 1978, a White Paper on the Official Secrets Act followed the recommendations of the Franks Committee proposing an Official Information Act that would criminalise the unauthorised release of specific types of information, being more explicit than the Official Secrets Act. March 1979 saw the publication of a Green Paper on open government and a report on access to information in other countries. The Green Paper proposed a Code of Practice on what ministers should authorise for release, in line with the Justice report of 1978. The Green Paper argued that FOI would be burdensome and expensive. As the 1970s came to a close, Wraith observed that

By training and tradition British civil servants are instinctively reticent. This reticence is reinforced by the influence of the Official Secrets Acts... But in recent years much more information has been disclosed at all levels than was once customary; and the predilection for government to be regarded by its initiates as a mystery whose secrets it is their function to guard is far less pronounced than it was a decade ago.³⁰¹

Nevertheless, Worthy has observed that there was, over all, a lack of progress on FOI in the 1970s despite the promises in Labour manifestos and agitation from the backbenches.³⁰² In 1982, access legislation was introduced in Australia (Freedom of Information Act), Canada (Access to Information Act) and New Zealand (Official Information Act).³⁰³ However, in the UK, Prime Minister Margaret Thatcher opposed FOI, was interested in reform to the Official Secrets Act only insofar as its restrictions could be made tighter, and prosecutions under

³⁰⁰ Nairne, P., 'Policy Making in Public', in Chapman, R. A. & Hunt, M. (eds.), *Open Government*, p.42.

³⁰¹ Wraith, 'United Kingdom', p.188.

³⁰² Worthy, *The Politics of Freedom of Information*, pp.32-33.

³⁰³ Marsh, N., 'Public Access to Government-Held Information: A Select Comparative Survey' in Marsh, N. (ed.) *Public Access to Government-Held Information* (London: Stevens and Sons, 1987) pp.259-6.

the Act became more frequent during her time in power.³⁰⁴ Worthy has suggested that Thatcher's opposition to openness created secrecy scandals and hardened Labour support for FOI.³⁰⁵

There were, nevertheless, further attempts to introduce FOI legislation: '...four MPs attempted to introduce Private Member's Bills between 1978 and 1984'.³⁰⁶ In 1984 the Outer Circle Policy Unit, which had been agitating for FOI, became the Campaign for Freedom of Information (CFOI).³⁰⁷ In addition to lobbying for FOI legislation, CFOI efforts were directed at smaller scale change, drafting and/or supporting legislation such as the 1985 Local Government Act, which extended public access to local government meetings, the 1987 Access to Personal Files Act and the 1990 Access to Health Records Act.³⁰⁸ In 1989, the House of Commons began to be televised (later than its USSR counterpart).³⁰⁹ While David Wilkinson has suggested that a shift was marked by the publication, in 1991, of the Citizen's Charter, which '...made openness one of the six key principles of government...'³¹⁰ – also noting that while it promoted openness, openness was still on the government's terms - Worthy points out that, because of the public sector's ingrained secrecy culture, the UK resisted FOI into the 1990s, after other Westminster countries and the UK's neighbours had passed access to information laws.³¹¹

Andrew McDonald states that liberalisation continued 'through administrative action', such as the beginning, in 1992, of a declassification programme known as the Waldegrave Initiative, after William Waldegrave, John Major's Minister for Open Government.³¹² As Nicholas Wilkinson has suggested, government transparency gained a high profile after the Scott Report into British arms sales to

³⁰⁴ Ibid.

³⁰⁵ Ibid., p.34.

³⁰⁶ Chapman, R. A., 'Introduction', in Chapman, R. A. & Hunt, M. (eds.), *Open Government*, p.18.

³⁰⁷ Worthy, *The Politics of Freedom of Information*, p.33.

³⁰⁸ Ibid.

³⁰⁹ Ibid., p.34.

³¹⁰ Wilkinson, D., 'Open Government: The Development of Policy in the United Kingdom in the 1990s' in McDonald, A. & Terrill, G. (eds.), *Open Government* (Basingstoke: Macmillan, 1998) p.14.

³¹¹ Worthy, *The Politics of Freedom of Information*, p.16.

³¹² McDonald, A., 'Archives and Open Government', in McDonald, A. & Terrill, G. (eds.), *Open Government: Freedom of Information and Privacy* (Basingstoke: Macmillan Press, 1998) p.26.

Iraq raised questions about access to documents.³¹³ A 1993 White Paper rejected FOI and laid the foundations for the 1994 voluntary Code of Access, which was praised as the first real step towards openness but also criticised as a stop gap against the flood of liberal opinion in favour of a move towards openness.³¹⁴ The White Paper ‘tightened the criteria documents had to meet if they were to be closed for longer than 30 years’.³¹⁵ In 1997, the Labour Government’s *Your Right to Know* White Paper was positive about FOI, though the draft Bill that followed was criticised by CFOI and parliamentary committees.³¹⁶ As Worthy explains: ‘The story of FOI from the radical White Paper of December 1997 to the draft Bill of May 1999 is that of a ‘closed door’ institutional struggle isolated from electoral pressure... It is the story of mobilisation and counter-mobilisation as factions supporting first a radical and then a lesser policy gained the upper hand in different contexts’.³¹⁷

This was happening in an international context that has been characterised as the heyday of neoliberalism. Kregg Hetherington considers ‘transparency’ and ‘accountability’ to be catchwords popular in the 1990s, part of a post-Cold War discourse in which democracy was seen as a consequence of making more information available.³¹⁸ Hetherington has argued that information economics informed the ‘information-for-development’ agenda of international development organisations such as the World Bank, which he claims promoted neoliberal concepts, though he also notes that ‘information-for-development’ had bipartisan support, citing conservatives like de Soto and North and critics of neoliberalism, such as Stiglitz.³¹⁹ Hetherington argues that:

Even as neoliberalism came under increasingly harsh condemnation in the late 1990s from people within the development apparatus, the argument

³¹³ Wilkinson, N., *Secrecy and the Media: The Official History of the United Kingdom’s D-Notice System* (London: Routledge, 2009) p.21.

³¹⁴ Worthy, *The Politics of Freedom of Information*, p.35.

³¹⁵ McDonald, ‘Archives and Open Government’, in McDonald and Terrill (eds.), *Open Government*, p.26.

³¹⁶ Campaign for Freedom of Information, *About Us*, <https://www.cfoi.org.uk/about/> [accessed 8 September 2018].

³¹⁷ Worthy, *The Politics of Freedom of Information*, p.57.

³¹⁸ Hetherington, K., *Guerilla Auditors: The Politics of Transparency in Neoliberal Paraguay* (London: Duke University Press, 2011) p.1.

³¹⁹ Hetherington, *Guerilla Auditors*, p.4.

that information would solve both political and economic ills of all sorts actually increased its influence after hard-edged neoliberalism waned.³²⁰

Robertson observes that ‘Criticisms of government secrecy can be found at all points of the ideological spectrum’³²¹ and shows how a desire for openness manifests in the left-wing, the anti-statist, the rationalist and the liberal.³²² In the UK setting, Hazell, Worthy and Glover point to the erosion of belief in government’s abilities, the desire for collaborative policy-making, growing expectations of information access following technological developments and the ‘information revolution’, and the need to address social issues by working across portfolios,³²³ as preconditions for the passage of FOI law. The argument in favour of broader information access would win out in the UK in 2000, with the passage of the Freedom of Information Act, concluding a ‘pre-FOI history of openness in the UK [that was] a series of compromise reforms forced upon uninterested, if not hostile, politicians and officials seeking to do as little as possible to ‘head off’ Parliamentary and media criticism.’³²⁴ The FOI Act passed into law with a five year grace period for implementation. The provisions of the law cut across the thirty year rule. The power of the law to open information is not solely derived from its request processes; Worthy points out that it can be used to build pressure or attract attention.³²⁵

Looking at the impact of FOI legislation, Hazell, Worthy and Glover found that FOI ‘has met the objective of increasing the transparency of government and has clearly allowed the public to see through the glass panes and know “what is going on inside”. Freedom of Information has led to more proactive disclosure of information across a range of issues.’³²⁶ However, they note that cultural change has not been, and is unlikely to be, total.³²⁷ Worthy has stated that the law has ‘reversed deeply ingrained secrecy habits and cultures’ and that it has ‘led to a

³²⁰ Ibid.

³²¹ Robertson, *Public Secrets*, p.15.

³²² Ibid.

³²³ Hazell, R., Worthy, B. & Glover, M., *The Impact of the Freedom of Information Act on Central Government in the UK: Does FOI Work?* (London: Palgrave MacMillan, 2010) p.108.

³²⁴ Worthy, *The Politics of Freedom of Information*, p.38.

³²⁵ Ibid., p.120.

³²⁶ Hazell, Worthy, and Glover, *Does FOI Work?*, p.103.

³²⁷ Ibid., p.115.

greater willingness to publish proactively',³²⁸ though the 'effect of transparency is not uniform' as may be expected from a complex bureaucracy like the UK central government.³²⁹

Although established in law and implemented across government, the UK's Freedom of Information regime continues to be contested. Just as the boundary between what is disclosable and non-disclosable will continue to be tested, the boundaries of FOI regimes are sites of tension. Greg Treadwell has outlined a number of threats to FOI outside of the UK context, including its superficial enactment as a pre-requisite for international finance or trade deals, the use of FOI as a misnomer for secrecy laws, the increased secrecy around state information following the September 11 attacks, and what H. Brinton Milward and Keith Provan call 'the hollow state' – the privatisation of public services, which moves information out of public hands and outside the scope of many FOI laws.³³⁰ Well before the Freedom of Information law was enacted in the UK, Patrick Birkinshaw foresaw this issue, remarking of privatisation: '...such developments are rarely accompanied by any appropriate accountability mechanisms and display the worst features of secrecy in the tradition of British government.'³³¹ Even within the public sector, the subversion of the spirit of the law has been evident in places. Historians and others have complained of the 'chilling effect' of FOI legislation, whereby public servants avoid documenting conversations and decisions for fear that the records will be released.³³² Roberts pointed out that technology is 'causing millions of undocumented conversations to be transformed into documents – all at risk for public disclosure',³³³ and while we see this in emails, which document conversations that might once have happened by phone, technologies such as mobile phones and intra-organisational social media have been problematic for full records capture.

³²⁸ Worthy, *The Politics of Freedom of Information*, p.119.

³²⁹ *Ibid.*, p.120.

³³⁰ Treadwell, G., 'FOI scholarship reflects a return to secrecy', *Pacific Journalism Review*, 22.1 (2016) p.125, and; Milward, H. B. & Provan, K. G., 'Governing the Hollow State', *Journal of Public Administration Research and Theory*, 10.2 (2000).

³³¹ Birkinshaw, P., *Reforming the Secret State* (Milton Keynes: Open University Press, 1990) p.48.

³³² Worthy, B., 'Does FOI Create a Chilling Effect?', *Open Data Study*, 28 September 2017 <https://opendatastudy.wordpress.com/2017/09/28/does-foi-create-a-chilling-effect-evidence-from-the-uk/> [last accessed 15 March 2019].

³³³ Roberts, *Blacked Out*, p.214.

Avoidance of record-making, through technological choices or otherwise, has caused a number of scandals. For example, in the Republic of Ireland, the chilling effect on documenting decisions at the ministerial level led the Think-tank for Action on Social Change to issue a policy brief calling for, amongst other things, a legal 'duty to document'.³³⁴ In the UK, this duty is enshrined in the Civil Service Code of Conduct, though compliance cannot be total. Though there have been some high profile cases of non-compliance in the UK, Worthy concludes that it is 'impossible to say if there is or is not an effect. It is likely there is some chilling at the margins and obviously much more skulduggery from politicians than we know of but it is not as widespread as some claim'.³³⁵ Given the size and range of government activities and the related difficulties for comprehensive scrutiny by the legislature and courts, through FOI principals - the press, citizen journalists and citizens in their own rights - have some agent monitoring capability.

4.6. Open Government Data and the Open Government Partnership

Under the Cameron government, open data eclipsed Freedom of Information as the focus of the openness movement, moving data into the spotlight in UK political discourse. Open data is any data that conforms to the Open Definition: 'Open means anyone can freely access, use, modify, and share for any purpose (subject, at most, to requirements that preserve provenance and openness)'.³³⁶ Open government data is public-sector information that has been made open in the sense of the Open Definition. In contrast to Worthy's history of the progress of Freedom of Information in the UK, the history of the open data movement has yet to be written. Writing in 2014, Rob Kitchin stated that the open data movement 'has been developing for a couple of decades, in tandem with, but largely separate from, the right to information (RTI) movement... and the open source

³³⁴ Think-tank for Action on Social Change (TASC), *Setting the Record Straight: Record Keeping for Good Governance* (TASC, 2016) https://www.tasc.ie/download/pdf/policybrief_settingtherecordstraightfinal.pdf [accessed 14 September 2018].

³³⁵ Worthy, B., 'Does FOI Create a Chilling Effect? Evidence from the UK, *Open Data Study: Research on Open Data and Transparency*, <https://opendatastudy.wordpress.com/2017/09/28/does-foi-create-a-chilling-effect-evidence-from-the-uk/> [accessed 14 September 2018].

³³⁶ Open Knowledge International, *The Open Definition*, <https://opendefinition.org> [accessed 14 September 2018].

and open science movements...'. For Kitchin, a watershed moment in the British open data movement was the Guardian newspaper's 'Free Our Data' campaign³³⁷ that started in 2006 with Charles Arthur and Michael Cross' article 'Give us Back our Crown Jewels', in which they argued 'Our taxes fund the collection of public data - yet we pay again to access it. Make the data freely available to stimulate innovation'.³³⁸

Internationally, the Organization for Economic Cooperation and Development encouraged data openness in 2008 and the following year the US government launched its open data portal <http://data.gov>.³³⁹ There has been an explosion of similar platforms since 2010, and it was in that year that the UK government launched data.gov.uk.³⁴⁰ Kitchin suggests the rapid development of open government data

has been facilitated by influential international and national lobby groups such as the Open Knowledge Foundation and the Sunlight Foundation, accompanied by the lobbying of knowledge-economy industry groups and companies, as well senior civil servants convinced by the arguments used, and dozens of local groups seeking to leverage municipal data.³⁴¹

As I will show below, the role of the third sector has been significant in open data promotion and use, but as Robertson notes, successive governments have been more liberal on the question of information access, and the Cameron government's enthusiasm for open data was critical in its ascendancy in British political life. This enthusiasm did coincide with questions about the efficacy of FOI, but as we have seen, the FOI regime survived that period intact. Published as the British Freedom of Information law was being implemented, James

³³⁷ The Guardian, *Free Our Data*, <http://www.theguardian.com/technology/free-our-data> [accessed 14 September 2018].

³³⁸ Arthur, C. & Cross, M., 'Give us back our crown Jewels,' *The Guardian* (9 March 2006) <https://www.theguardian.com/technology/2006/mar/09/education.epublic> [accessed 14 September 2018].

³³⁹ Kitchin, R., *The Data Revolution: Big Data, Open Data, Data Infrastructures & their Consequences* (London: Sage, 2014) p.49.

³⁴⁰ Government Digital Service, *data.gov.uk*, <https://data.gov.uk/about> [last accessed 15 March 2019].

³⁴¹ Ibid.

Michael's essay on openness offered some thoughts on the future of the openness movement, anticipating a trend towards openness that would continue through the spread and strengthening of FOI laws. Although he could not have foreseen the advent of the open data movement, he did anticipate an internationalisation of the structures and systems for implementing and enforcing openness measures.³⁴² Perhaps one of the greatest spurs to the growth of the open data movement was the establishment of the Open Government Partnership (OGP) in 2011, when eight national governments signed the Open Government Declaration (OGD), committing themselves to 'Increase the availability of information about governmental activities... Support civic participation... Implement the highest standards of professional integrity throughout our administrations [and]... Increase access to new technologies for openness and accountability...'.³⁴³ As at September 2018, membership of the OGP stands at 79 national governments and 20 subnational governments.³⁴⁴ To join the OGP, governments are required to sign the OGD, publish an action plan that sets out commitments to action and commit to independent reporting on progress.³⁴⁵ The United Kingdom was one of the eight founding countries of the OGP, and its first action plan (2011-2013) described commitments relating to the formalisation of open data mechanisms and, in contrast to the concerns of civil society that the government was preparing for a rollback of FOI, commitments to strengthening and improving FOI, such as higher cost caps on processing requests for information in digital systems.³⁴⁶

The Open Government Declaration characterises openness as concerned with increasing the availability of information about governmental activities, supporting civic participation, promoting professional integrity and increasing access to 'new

³⁴² Michael, J., 'Fol in Other Countries' in Wilson, D. (ed.) *The Secrets File: The case for freedom of Information in Britain today* (London: Heinemann Education Books, 1984) p.108.

³⁴³ Open Government Partnership, *Open Government Declaration*, <https://www.opengovpartnership.org/open-government-declaration> [accessed 14 September 2018].

³⁴⁴ Open Government Partnership, *About OGP*, <https://www.opengovpartnership.org/about/about-ogp> [accessed 14 September 2018]

³⁴⁵ Open Government Partnership, *How To Join*, <https://www.opengovpartnership.org/how-join> [accessed 14 September 2018]

³⁴⁶ United Kingdom Cabinet Office, *UK Open Government National Action Plan 2011 to 2013*, <https://www.gov.uk/government/publications/uk-open-government-national-action-plan-2011-to-2013> [accessed 14 September 2018].

technologies for openness and accountability’, but information access and civic participation have been the focus of much of the work connected with the OGP in the UK. In 2012, Tim Berners-Lee and Nigel Shadbolt established the Open Data Institute (ODI) in London, ‘as a world-leading centre to innovate, exploit and research the opportunities for the UK created by the Government’s Open Data policy’.³⁴⁷ The ODI rapidly became a hub of OGP and open data activity, serving as a base for work on policy, technology, research and business innovation. It became the meeting place for the UK’s OGP civil society network as it drafted proposals for National Action Plan (NAP) commitments in an example of participatory policy-making. The civil society network is a group of non-government organisations (NGOs) and individuals concerned with a range of issues from transparency in particular fields (extractive industries, healthcare, etc) to crosscutting issues such as Freedom of Information (the CFI), privacy (Privacy International) and records management (the International Records Management Trust).³⁴⁸

In the same year as the ODI was founded, Anne Thurston’s ‘Trustworthy Records and Open Data’ was published in the *Journal of Community Informatics*.³⁴⁹ The article called for the recognition of the limitations of open data in the absence of robust information quality controls of the type developed by record-keeping professionals. This was the first attempt to link the open data and records communities, though the message was not well-received by the proponents of open data, who saw the publication of data as the priority, with quality control being done by users after publication. However, in 2013, the Transparency and Accountability Initiative (TAI), an NGO, developed the Open Gov Guide, a source of draft commitments that governments and civil society could use to develop their own action plans. The Open Gov Guide identified a number of cross-cutting and sector-specific issues, and for each, listed recommended commitments at

³⁴⁷ Open Data Institute, *Business Plan 2012-2017*, <http://e642e8368e3bf8d5526e-464b4b70b4554c1a79566214d402739e.r6.cf3.rackcdn.com/odi-business-plan-may-release.pdf> [accessed 14 September 2018].

³⁴⁸ UK Open Government Civil Society Network, <https://www.opengovernment.org.uk/> [last accessed 15 March 2019].

³⁴⁹ Thurston, A., ‘Trustworthy Records and Open Data,’ *The Journal of Community Informatics*, 8:2, 2012.

‘initial’, ‘intermediate’ and ‘advanced’ stages.³⁵⁰ The International Records Management Trust (IRMT) developed text for commitments on records management as a cross-cutting issue, each with links to guidance and standards, in order to bring information quality concerns to the fore.³⁵¹

The second cycle of NAP development, resulting in the 2013-2015 action plan, was lauded as an exemplary exercise in participatory policy-making. The UK process saw the network of civil society organisations and individual activists draft commitments that were proposed to government, many of which were taken forward into the action plan.³⁵² The commitment development process required civil society partners to work with government agencies to draft text, and the IRMT proposed a commitment on records management, of the sort seen in the first US NAP. The IRMT was partnered with The National Archives (TNA), as the government body leading on records management policy at that time. TNA failed to recognise the significance of the OGP process, and delegated the liaison role to a junior member of staff who was not able to direct or commit TNA to a policy direction. As the deadline for commitment submission drew closer, a decision was finally taken to develop a commitment that closely aligned with TNA’s own strategic plan, for the sake of ensuring that records management was featured in the action plan, raising the profile of the issues around data authenticity. Even if the commitment was unambitious, the issue was, at least, reflected in the NAP. The TNA’s blog post about the development of the commitment takes a contrasting view of the process.³⁵³ In any case, the records management commitment would become commitment 5 in the 2013-2015 action plan:

³⁵⁰ Transparency and Accountability Initiative, *Open Gov Guide*, now at <https://www.opengovpartnership.org/resources/open-government-guide> [accessed 15 March 2019].

³⁵¹ *Ibid.*, p.154.

³⁵² Worthy, B., *Independent Reporting Mechanism (IRM): United Kingdom End of Term Report 2013-2015* (Open Government Partnership, 2016). https://www.opengovpartnership.org/sites/default/files/UK_EOTR_2013-2015_0.pdf [accessed 14 September 2018]; The Democratic Society, *Developing a Model for Open Government in the EU*, http://www.demsoc.org/wp-content/uploads/2016/12/Demsoc.OSEPI_EUOpenGovModel.pdf [accessed 14 September 2018].

³⁵³ Merifield, M., *Loving it when a plan come together*, <https://blog.nationalarchives.gov.uk/blog/loving-plan-comes-together/> [accessed 14 September 2018].

The UK government will manage and capture digital records and there will be a comprehensive, accessible and timely paper and digital record of UK government available to the citizen.³⁵⁴

The Independent Reporting Mechanism (IRM) report on this commitment noted that TNA had made substantial progress on the commitment, but that the commitment's effects on openness were minimal, being instead 'an important first step and essential background to future transparency efforts'.³⁵⁵ During the development of the third action plan, records and archives professionals met at University College London's Department of Information Studies to develop a draft commitment on records management. The draft text was centred around the idea of a strengthening of the Public Records Act, with a view to empowering TNA to lead on the standardisation of records creation, capture, management and preservation across government, seen as foundational to the advanced kind of integrated openness regime seen in Norway.³⁵⁶ This was taken to the OGP Civil Society Forum and endorsed. Meetings with TNA staff revealed that there was 'no appetite' for another commitment, particularly as the draft commitment proposed a revision to the Act. This position was adopted at a time when Sir Alex Allen's review of government record-keeping resulted in the recommendation that the records management policy responsibility be shifted from TNA to the Cabinet Office, at odds with technical and professional thinking about the policy-setting and system design responsibilities of archives in the digital environment, and when TNA was moved from the Ministry of Justice to the Department for Culture, Media and Sport, which was widely seen as a demotion.

The third UK OGP national action plan (2016-18) was published in May 2016 and contained commitments on specific types of data, such as election, grants and company ownership data, and a discernible turn towards considerations of data infrastructure and participation, with commitments on social and technological

³⁵⁴ Ibid.

³⁵⁵ Worthy, *IRM Report*, https://www.opengovpartnership.org/sites/default/files/UK_EOTR_2013-2015_0.pdf [accessed 14 September 2018] p.17.

³⁵⁶ Lowry, J., 'Opening Government: Open Data and Access to Information', in Lowry, J. & Wamukoya, J., (eds.), *Integrity in Government through Records Management: Essays in Honour of Anne Thurston* (Farnham: Ashgate, 2014) pp.164-165.

innovation against corruption, improved use of web-based tools to encourage data use and civic participation. The National Archives was not listed as a lead implementing organisation for any of the commitments. Instead, the Government Digital Service was identified as the lead agency on digital infrastructure, and commitments in data use in policy formulation, the economy and civil society fell to National Statistics. While the TNA's lack of engagement with openness has seen the potential contributions of the record-keeping profession excluded from the UK's OGP work, the plan reaffirms a commitment to FOI, despite initial concerns amongst civil society actors that open data was intended to replace an expensive and sometimes cumbersome FOI regime:

Commitment 8: Enhanced transparency requirements and revised Freedom of Information Act Code of Practice - To increase transparency and improve the operation of the Freedom of Information (FOI) Act in the public interest.³⁵⁷

As I observed in 2014, 'By 2013, enthusiasm for open data had largely overshadowed access to information [Freedom of Information], which tends increasingly to be seen by government as expensive and laborious, and by some civil society groups as too limited in scope to enable true transparency and accountability'.³⁵⁸ Through the OGP, successive governments have reaffirmed their commitments to FOI, though this is partly attributable to the defence and promotion of FOI by activists and scholars, often working through the OGP process, notably Maurice Frankel and Ben Worthy.

In parallel with the opening of government data through OGP, there has been a growth in the use of data through civic technologies. These resources take government data and present them to users in visual or interactive ways for the purposes of effecting social change. Broadly defined, civic technology is technology that intersects with public life, enabling engagement or participation of the public with the government for a variety of purposes, from making

³⁵⁷ United Kingdom Cabinet Office, *UK Open Government National Access Plan 2016-2018*, <https://www.gov.uk/government/publications/uk-open-government-national-action-plan-2016-18/uk-open-government-national-action-plan-2016-18> [accessed 14 September 2018].

³⁵⁸ Lowry, 'Opening Government', p.164.

government more transparent and accountable, to enhancing civic communities, and impacting policy decisions. The Knight Foundation has identified two specific strands of civic technology: open government and community action. From this, civic technology can be further divided into eleven streams: data access and transparency; data utility; public decision making; resident feedback; visualization and mapping; voting; civic crowdfunding; community organising; information crowdsourcing; neighbourhood forums; peer-to-peer sharing.³⁵⁹ Civic technology, therefore, can take various forms: an application that enables residents to share goods and skills and a website allowing resident feedback on environmental issues posited by local government are both examples of civic technology. However, a common thread is a foregrounding of the user-citizen, with a focus on their empowerment and engagement within the community, and therefore the political process, through technological means. Popular examples in the UK include Fix My Street (<https://www.fixmystreet.com>), which allows residents to report problems in their area, such as damaged roads, streetlights or graffiti to local authorities, but publicly, and responses are also documented publicly. Another example is They Work For You (<http://www.theyworkforyou.com>), which aggregates data about Members of Parliament so that constituents can see their attendance records, voting records, contributions to committees and so on. Through these technologies, civic participation is increased. Participation is a form of the dialogic exchange of information discussed in Chapter 2; an exchange that binds individuals to the *dispositif* of bureaucracy and the social contract, realised in records. Therefore, like records, civic technologies and the open data that fuels them may be dissensual technologies or technologies that foreclose dissensus. I will develop this idea in Chapter 10, turning next, instead, to the question of the quality of information that is being opened through the monitoring mechanisms described above.

4.7. Conclusion

³⁵⁹ Knight Foundation, *The Emergence of Civic Tech: Investments in a Growing Field*, December 2013 https://www.knightfoundation.org/media/uploads/publication_pdfs/knight-civic-tech.pdf [last accessed 20 September 2018].

This brief history of secrecy and openness in British government has demonstrated a general trend towards openness, from Victorian and Edwardian legislation for secrecy in a growing state apparatus, to a rapid and arguably accelerating move towards openness from the 1960s onwards, taking first legislative forms and more recently the modes of policy and technology. Much of the literature suggests that there is a correlation between authoritarian, feudal and monarchic forms of government and secrecy, on one hand, and democratic and participatory forms of government and openness on the other. The common assumption that openness correlates with democratic governance is borne out by the history provided by Alasdair Roberts, which cites evidence in the right to free speech, the growth of the free press, open lawmaking and the gradually widening circle of suffrage following the revolutions in England and France.³⁶⁰ As new approaches to openness are developed, a question arises whether openness is the result of the changing face of sovereign power, or if a threshold has been crossed, and openness is now the cause of a changed or changing and increasingly devolved political dynamic. This is an idea I will return to in Chapter 10. The popular view of the relationship between political power and information asymmetry is expressed in Hetherington's observation of guerilla auditors in Paraguay; 'The secrecy that made documents capable of channeling sovereign power now [opened] made them capable of channeling the new democracy'.³⁶¹ Robertson, however, is skeptical that secrecy and democracy are in opposition to each other, arguing that official secrecy was codified during a period of reform that sought to democratise the civil service.³⁶² In any case, some of the political implications of openness are clear: 'The ecosystem forms part of a new political information cycle that is gradually supplementing the mainstream media, offering new participants avenues into influencing politics...'³⁶³

Support for government openness has often been bipartisan, both across the state/civil society divide and across political divides, as Chapman points out: 'In this environment [characterised by secrecy], with its imprecision or lack of

³⁶⁰ Roberts, A., *Blacked Out: Government Secrecy in the Information Age* (New York: Cambridge University Press, 2008) p.10.

³⁶¹ Hetherington, *Guerrilla Auditors*, p.45.

³⁶² Robertson, *Public Secrets*, p.41.

³⁶³ Worthy, *The Politics of Freedom of Information*, p.184.

agreement about basic values and about the practical applications of those values in the system of government, some individuals and groups with various beliefs and interests have found common ground and expressed pragmatic demands to make government more 'open'.³⁶⁴ In the UK, steps towards openness have been taken by Labour and Conservative governments, and have been driven forward by pressures from inside and outside of the state's bureaucracy. Although hindsight allows us to see a progression, the present state of affairs is less easy to define. Roberts observed that by 'the end of the 1990s there were many people who believed that the "right to know" – and the presumption of openness – had finally become entrenched as a basic principle of democratic governance',³⁶⁵ but information access continues to be a political battleground. For instance, the high court recently ruled that the Data Retention and Investigatory Power Act of 2014, which required 'internet and phone companies to keep their communications data for a year and regulates how police and intelligence agencies gain access to it,' is illegal and inconsistent with European Union law.³⁶⁶ The legal contestation reflects the multifarious perspectives on 'appropriate' openness, and we should acknowledge that different perspectives will continue to proliferate around the question of appropriateness. And as Robertson has asked, would 'those who see conspiracies... be satisfied by any institutional or legal structure by which 'the will of the people' was to be decided.'³⁶⁷

Roberts wrote in 2006 that as 'the power of... new monitoring techniques becomes apparent, we can expect the actors that are subject to monitoring (businesses, citizens, governments) to react by seeking to restrict the inflow of information to government databases, or the outflow of that information from public agencies.'³⁶⁸ Over ten years later, it is not possible to state categorically that this expectation has been fulfilled. The chilling effect of FOI, the subversion of formal communication channels (private email servers used for public

³⁶⁴ Chapman, 'Introduction', in Chapman and Hunt, (eds.), *Open Government*, p.24.

³⁶⁵ Roberts, *Blacked out*, p.9.

³⁶⁶ Bowcott, O., 'High court rules data retention and surveillance legislation unlawful', *Guardian* (17 July 2015) <http://www.theguardian.com/world/2015/jul/17/data-retention-and-surveillance-legislation-ruled-unlawful> [accessed 14 September 2018].

³⁶⁷ Robertson, *Public Secrets*, p.15.

³⁶⁸ Roberts, *Blacked Out*, p.225.

business, for instance), the barriers to information access that arise when public services are outsourced to companies beyond the reach of access legislation (taken to its extreme in America's 'shadow government' as outlined by Stacy Wood³⁶⁹), these are parts of an information ecology in which there also exist powerful government actors who subscribe to participatory governance models on ideological grounds, where counterintelligence technologies are employed by civil society and individuals to engage in record-making about government (another ecosystem of 'public' records), and where community data projects wrest functions from the public sector or perform them in parallel. These serve to remind us that information asymmetry is always in flux – it exists under constant pressure from diverse forces that conflict, compete or collaborate in a wider socio-political and economic field. Information, then, is a site of dissensus and potentially a mechanism for impertinent dialectics. As with Ranciere's conceptualisation of viable democracies as always in a state of self-correction, the limits on government openness and the corpus of public sector information are always contested and, according to perspective, correcting. The diversity of information creation, capture, use and sharing practices in this environment raises questions about the trustworthiness of all information, and specifically in relation to sovereign power's exercise, the trustworthiness of government information. This history reveals a general trajectory towards openness, but can we trust the information being released? This question is central to the informational contestations and corrections of the social contract.

³⁶⁹ Wood, S., 'Making Secret(s): The Infrastructure of Classified Information', (doctoral thesis, University of California, Los Angeles, 2017) p.57.

Chapter 5

Trustworthy Information

No technology without rules, without signatures, without bureaucracies and stamps. Law itself is no different from the world of technologies: it is the set of the modest technologies of writing, registering, verifying, authenticating that makes it possible to line up people and statements... A signature on a contract, an endorsement, an agreement stabilizes the relative size of the actors by lending to the provisional definition of alliances the assistance of the law, a law whose weight is enormous because it is entirely formal and because it applies equally to everyone. Mr. Lagardère may vary in size, the ministry will change hands ten times - it would be unwise to count on stability there; but the signatures and stamps remain, offering the alliances a relative durability. *Scripta manent*. That will never be enough, for signed documents can turn back into scraps of paper.³⁷⁰

Bruno Latour, *Aramis, or the Love of Technology*

³⁷⁰ Latour, B., *Aramis, or the Love of Technology*, trans. Porter, C. (Cambridge: Harvard University Press, 1996) p.45.

5.1. Introduction

Once government data can be accessed, a question arises about its trustworthiness. In order for monitoring mechanisms to be effective in countering information asymmetry, the information to which they provide access must be trustworthy. It is not enough to have quantities of information. Hervé Stolowy, Charles Baker, Thomas Jeanjean and Martin Messner wrote that ‘Roberts... has suggested that ‘increased information may simply create an “illusion of transparency”... rather than actually improving control and accountability’.³⁷¹

Reflecting on the importance of trust in investment decision-making, they observe that the role of trust ‘in everyday (economic) life is easily overlooked, due to its fundamental and implicit nature... Trust is needed because it would be impossible or at least overly costly to obtain a sufficient level of information that would allow us to say that we “have checked it all by ourselves”’.³⁷² The need to be able to trust information – to avoid the expense and labour of each user checking it, which is, in any case, impossible in the increasingly complex and data saturated information ecology – has produced approaches to protecting and authenticating information as trustworthy. This chapter will look at informational trustworthiness in historical perspective, explaining the introduction of diplomatics into thinking about trustworthiness in digital records, which forms the dominant conceptualisation of trustworthiness in the record-keeping field.

5.2. Informational Trustworthiness in Historical Perspective

Interest in the trustworthiness of information has been entwined in the legal system from the time that writing was introduced into the execution of transactions. The significance of the issue is apparent in ancient Roman law, under which punishment for the destruction or falsification of records ‘could range in various cases from corporal punishment to death, depending on the rank of the

³⁷¹ Stolowy, H., Baker, R., Jeanjean, T. & Messners, M., ‘Information, trust and the limits of “intelligent accountability” in investment decision making: Insights from the Madoff case,’ *Les Cahiers de Recherche*, 956, (2011) pp.3-4.

³⁷² *Ibid.*, p.3.

accused and the nature of the forgery'.³⁷³ In the origins of the common law legal systems that exist today in the Commonwealth of Nations and the United States, there is evidence that the veracity of information was a central concern in legal transactions. Heather MacNeil has explained the emergence of records as evidence in England in the transfer of freehold interest in land under the feudal system – a transaction known as a *feoffment* - which was given effect by 'livery of seisin', a ceremony in which the grantor would, on or near the land being transferred, give the grantee a sod of turf and a twig from a tree growing on that land.³⁷⁴ Of principal importance in the performance of the ceremony was the presence of witnesses. The act was done in public so that the transfer of property rights was well known and could be testified to. It was under the influence of the civil law tradition that the written charter was introduced as a replacement for the symbols of the transferred property - the turf and twig - allowing the livery of seisin to happen away from the land being transacted.³⁷⁵ The function of the charter was initially primarily symbolic, though importantly it recorded the names of the witnesses to the ceremony.³⁷⁶ MacNeil writes that between 'the twelfth and fifteenth centuries, charters were gradually transformed from an aide-memoire of an oral act to the indisputable proof of that act', but notes that people 'only began to trust the written memorial and prefer it to the living memory of witnesses when a generally accepted method of authenticating documents emerged'.³⁷⁷ This illustrates the way in which authentication mechanisms support the function of records-as-commands that transmit authority over time and space. The means of authentication become part of that dialogical mechanism of information and power exchange. MacNeil explains that the early form of this method was the use of seals, and that it 'gave rise, in turn, to the legal doctrine of estoppel by deed, which prohibited an individual from denying any facts stated in a writing to which he has set his seal'.³⁷⁸ The seal performed a 'non-repudiation function'. Seals were intimately linked with persons, tying character to the written record and

³⁷³ Ferguson, M. C. quoted in MacNeil, H., *Trusting Records: Legal, Historical and Diplomatic Perspectives* (Dordrecht; London: Kluwer Academic, 2000) p.3.

³⁷⁴ MacNeil, *Trusting Records*, p.315.

³⁷⁵ Clanchy, M. T., *From Memory to Written Record: England 1066-1307* (London: Edward Arnold Publishers, 1979) p.36.

³⁷⁶ MacNeil, *Trusting Records*, p.316.

³⁷⁷ *Ibid.*, pp.316-317.

³⁷⁸ *Ibid.*, p.318.

assigning personal responsibility for the content of the record. Brigitte Bedos Rezak has described how the matrices used to make the seals' impressions in wax were modified to reflect changes in the status or function their owners and, on the owner's death, destroyed or buried with them.³⁷⁹ This close link between the character of the person and the promise of the seal is discussed by Arthur Giry, who wrote that the seals:

...of sovereigns, barons, prelates, churches, and municipalities were from the very beginning used to guarantee the authenticity, not only of those instruments in which the owner of the seal bound himself or was otherwise a party; but also of all documents to which it was desired to give (in legal phraseology) an "authentic" character...³⁸⁰

Over time, the non-repudiation function performed by medieval seals would come to be performed by the signature,³⁸¹ which linked the author to the document:

'...a signature is not just any mark. It is commonly understood to be a representation of a name, and it is the link between the name and a person that gives the signature value'.³⁸² Johanna Drucker has observed the same connection, noting that 'many basic ideas about authenticity and identity are linked to writing as a personal gesture understood as a personal act'.³⁸³ MacNeil wrote that the 'purporting seal or signature is sufficient proof that the document has been genuinely executed by the purporting official, whose official character is assumed without evidence'.³⁸⁴ She sees, in this use of sealed or signed documents as evidence, the roots of the modern idea of documentary evidence, and indeed the current US *Rules of Evidence* consider sealed documents to be 'self-authenticating'.³⁸⁵ Present day legal norms in the treatment of documentary

³⁷⁹ Bedos-Rezak, B., 'Seals and Sigillography, Western European,' in Strayer, J. (ed.) *Dictionary of the Middle Ages*, vol. 11, (New York: Charles Scribners Sons, 1988) p.127.

³⁸⁰ Giry, A., *Manuel de Diplomatique*, trans. and quoted in Wigmore, J. H., *Evidence in Trials at Common Law, Volume 7*, ed. and rev. by Chadbourn, J., (Boston: Little, Brown, 1978) para. 2161.

³⁸¹ MacNeil, H., 'Providing Grounds for Trust: Developing Conceptual Requirements for the Long-Term Preservation of Authentic Electronic Records' *Archivaria* 50.1 (2000) p.61.

³⁸² Robertson, C., *The Passport in America: The History of a Document* (Oxford: Oxford University Press, 2012) p.54.

³⁸³ Drucker, J., *What Is? Nine Epistemological Essays* (Berkeley: Cuneiform Press, 2013) p.17.

³⁸⁴ MacNeil, *Trusting Records*, p.48.

³⁸⁵ Rule 902 'Evidence That Is Self-Authenticating' states: 'The following items of evidence are self-authenticating; they require no extrinsic evidence of authenticity in order to be admitted: (1) *Domestic Public Documents That Are Sealed and Signed*. A document that bears:

evidence are also rooted in the history of documentary forms. Looking to the history of the assessment of written evidence, in particular, we see the emergence of a method for interrogating the trustworthiness of charters in 1681, with the publication of Jean Mabillon's *De re diplomatica*, as noted in Chapter 2. David Knowles has described how, in the seventeenth century, historians began to use charters and diplomas from the early Middle Ages, and that

It was soon common knowledge that many of them were wholly or in part unreliable; some were falsifications of the medieval centuries, others were copies, more or less inaccurate and interpolated, of genuine originals now lost, while some *soi-disant* medieval pieces were in fact forgeries of the modern world. The last class owed its origin to the use made of medieval documents in disputes over property and still more to the efforts of noble families in France, and ancient abbeys in Germany, to establish precedence and prestige and rights that were due to extreme antiquity.³⁸⁶

Following an attempt by the Flemish hagiographer Daniel Papebroch, in the second volume of the *Acta Sanctorum* (1675), to develop rules for the criticism of these documents, Mabillon saw that, in assessments of these documents,

...a decision could be reached only after the expert had examined a whole series of different indications – the material used, the seal, the signature, the grammar and orthography, the modes of address, the plausibility of the dating, the intrinsic consistency of the whole document.³⁸⁷

(A) a seal purporting to be that of the United States; any state, district, commonwealth, territory, or insular possession of the United States; the former Panama Canal Zone; the Trust Territory of the Pacific Islands; a political subdivision of any of these entities; or a department, agency, or officer of any entity named above; and

(B) a signature purporting to be an execution or attestation.' Supreme Court of the United States, *Federal Rules of Evidence* (2016), https://www.law.cornell.edu/rules/fre/rule_902 [accessed 17 August 2018].

³⁸⁶ Knowles, D. D., *The Historian and Character: And Other Essays* (Cambridge: Cambridge University Press, 2008) p.221.

³⁸⁷ *Ibid*, p.222.

He therefore wrote a more comprehensive guide, accompanied by facsimiles of charters to illustrate the principles of diplomatics. Importantly, Mabillon also situated the charters and diplomas as the products of bureaucratic procedures that produced norms. As MacNeil put it, by ‘focusing attention on the importance of chancery procedures in the establishment of norms for documentary creation, [Mabillon] provided future diplomatists with the tools for assessing the conformity of the document’s formal elements with those established by chancery procedures’.³⁸⁸

Concerns for the trustworthiness of information also manifest in other legal traditions and ideas. The ‘best evidence rule’ privileged original records over copies or facsimiles because of the potential for the introduction of errors or fraudulent elements during copying, particularly copying by hand. In the United Kingdom, the best evidence rule has fallen out of effect, and today all relevant evidence may be admitted, though the trustworthiness of documents will be a consideration in assessing their evidential weight.³⁸⁹ The ‘business entry rule’ or ‘business records exception to the hearsay rule’ in US evidence law is another manifestation of a concern for authenticity. That rule holds that records created in the normal course of business that meet certain criteria are not to be treated as hearsay.³⁹⁰ The logic at work in the rule is that records naturally occurring from business processes can be presumed authentic because they are used by the originating organisation to effect transactions or as the basis of decisions – they are acted upon by their creators in such a way as to suggest that the creator has faith in their authenticity, suggesting they have not tampered with them or allowed them to be tampered with. The rule implicates people and systems as constituent elements of an environment of trustworthiness. Although legal standards for evidence are not necessarily the standards by which the authenticity of government information is judged by the individual or civil society, the consideration given to authenticity in those legal standards is helpful in defining conditions for trustworthiness.

³⁸⁸ MacNeil, *Trusting Records*, p.48.

³⁸⁹ Ormerod, D., Murphy, P. & Phillips, J. (eds.), *Blackstone's Criminal Practice* (Oxford: Oxford University Press, 2008) p.2285.

³⁹⁰ Supreme Court of the United States, *Federal Rules of Evidence* (2016), https://www.law.cornell.edu/rules/fre/rule_803 [accessed 15 March 2019]

5.3. Archival Approaches to Trustworthiness

The law also provides the epistemological foundations of the discipline of archival science, since it is from legal necessities that record-making and record-keeping have emerged in government and business.³⁹¹ Much of the language of modern archival science is grounded in work conducted by Duranti, MacNeil, Mak and others at the University of British Columbia in looking to law and diplomatics for models or methods for assessing and protecting authenticity in digital records. The modern interest in new applications of diplomatics can be traced to Christopher Brooke's notion of a modern diplomatic in 1970, but really begins in earnest in the 1980s, with work by Tom Nesmith, Hugh Taylor, Don Skemer and Francis Blouin.³⁹² Blouin's contribution is particularly noteworthy because his paper, given at the 1989 International Conference of the Round Table on Archives, prompted the delegates, who were senior government archivists and leaders in the field of archival science, to recommend 'that the development of the discipline of modern diplomatics be promoted through research in the typology of contemporary records and in the records-creating procedures of contemporary institutions'.³⁹³ Also in 1989, Luciana Duranti published the first of six articles on diplomatics. These essays would later (in 1998) be compiled and published as a monograph - *Diplomatics: New Uses for an Old Science*. They set out ideas on the forms and uses of records and their place in law, rights and business, forming the foundations of the research agenda in this area. The term 'archival-diplomatics' has been used to describe this application of diplomatics to digital records.

In 1992, David Bearman stated that electronic records systems would require

³⁹¹ MacNeil, H., 'Trusting Records in a Postmodern World.', *Archivaria*, 51 (2001) p.39.

³⁹² Brooke, C. N. L., 'The Teaching of Diplomatic,' *Journal of the Society of Archivists*, 4.1(1970) pp.1-9; Skemer, D. C., 'Diplomatics and Archives,' *The American Archivist*, 52.3 (1989) pp.376–82; Taylor, H., "My Very Act and Deed": Some Reflections on the Role of Textual Records in the Conduct of Affairs', *The American Archivist*, 51.4 (1988) pp.456–69; Blouin, F. X., 'Convergences and Divergences in Archival Tradition: A North American Perspective,' *Second European Conference on Archives. Proceedings*, (Paris: International Council on Archives, 1989) pp.28-9; Nesmith, T., 'Archives from the Bottom Up: Social History and Archival Scholarship', *Archivaria* 14 (Summer 1982) pp.5–26.

³⁹³ Koucky, J. (ed.), *Second European Conference on Archives: Proceedings* (Paris: International Council on Archives, 1989) p.113.

archivists to define new diplomatic norms. Between 1994 and 1997, Luciana Duranti, with co-investigator Terry Eastwood and assistance from Heather MacNeil, enquired into the essential characteristics of reliable and authentic records in the 'UBC MAS Research Project'.³⁹⁴ The research found:

1. the reliability and authenticity of electronic records are best ensured by embedding procedural rules in the overall records system and by integrating business and documentary procedures;
2. the reliability and authenticity of electronic records are best guaranteed by emphasizing their documentary context; and
3. the reliability and authenticity of electronic records can only be preserved if they are managed together with all the other records belonging in the same fonds.³⁹⁵

In effect, these findings underscore the crucial importance of systems, context and provenance. The research also found:

1. the life-cycle of the managerial activity directed to the preservation of the integrity of electronic records can be neatly divided into two phases: one phase directed to the control of the creation and maintenance of reliable and authentic active and semiactive records, and the other phase directed to the preservation of authentic inactive records; and
2. the integrity of electronic records is best preserved by entrusting the creating body with responsibility for their reliability and the preserving body with responsibility for their authenticity.³⁹⁶

These two findings essentially draw a distinction between reliability, as a consequence of record-making, and authenticity, as a consequence of record-keeping.

5.4. Defining Trustworthiness

³⁹⁴ UBC MAS stands for University of British Columbia Masters of Archival Science.

³⁹⁵ Duranti, L. & MacNeil, H., 'The Protection of the Integrity of Electronic Records: An Overview of the UBC-MAS Research Project', *Archivaria*, 42 (Fall 1996) p.58.

³⁹⁶ *Ibid.*, p.59.

The findings of UBC MAS laid the foundations for the International Research on Permanent Authentic Records in Electronic Systems (InterPARES) programme, led by Duranti, which began in 1998 and continues to the present day. An important contribution that this work has made to the pursuit of authenticity in digital records is the definition of relevant concepts and their relationships with each other. In archival studies and archival science today, the prevailing usages of the terms ‘trustworthiness’, ‘reliability’ and ‘authenticity’ follow the definitions arising from the UBC interest in law and diplomatics.

The second phase of the InterPARES project (2002 – 2007) produced a Terminology Database to define key terms pertinent to the research, to ensure that the research teams, working around the globe, were working with commonly understood terminology. The InterPARES Terminology Database defines ‘trustworthiness’ as the ‘accuracy, reliability and authenticity of a record’.³⁹⁷ In turn, these component parts of trustworthiness are defined:

Accuracy – ‘The degree to which data, information, documents or records are precise, correct, truthful, free of error or distortion, or pertinent to the matter’.

Reliability – ‘The trustworthiness of a record as a statement of fact. It exists when a record can stand for the fact it is about, and is established by examining the completeness of the record's form and the amount of control exercised on the process of its creation.’

Authenticity – ‘The trustworthiness of a record as a record; i.e., the quality of a record that is what it purports to be and that is free from tampering or corruption’.³⁹⁸

5.4.1. Reliability and Accuracy

³⁹⁷ InterPARES 2, *Terminology Database*, http://www.interpares.org/ip2/ip2_terminology_db.cfm [last accessed 15 March 2019].

³⁹⁸ Ibid.

According to the work of InterPARES 2, the reliability of a record should be assessed by gauging its completeness and the documentary procedure generating it: it is 'the sole responsibility of the physical or juridical person making the record'.³⁹⁹ Duranti has considered the bureaucratic controls that can be used to assure reliability:

If the writer is reliable, the user can identify him/herself with the writer, that is, with the witness to the fact. To be able to rely on record-writers requires controlling them in a number of different ways... Raffel identifies four ways: 1) restricting the privilege of record-writing to professionals (of course, record-writers are those who sign records and/or are responsible for them); 2) imposing sanctions on record-writers by requiring signatures, so that the bureaucrat has a record anyway, either a record of the fact or a record of who failed to report the fact; 3) instituting procedures, that is, giving responsibility to each writer for reporting only a portion of a fact, and/or increasing the number of those who report the same fact, so that what their records will have in common will be the true fact; 4) making different purposes concurrent, that is, making the same record serve a number of different users: instead of the number of writers, the size of the audience is increased, so that the writer cannot tailor the message to the audience.⁴⁰⁰

Of these, Duranti sees the signature as the most relevant control because it 'gives responsibility to the writer for his/her words; therefore the user does not need to check the record against the fact, because the signature shows and legally establishes where the responsibility lies. The signature is the fact'.⁴⁰¹ There are two dynamics at work in this view. One concerns the trustworthiness of the signatory – the character of the signatory comes to bear on the reliability of

³⁹⁹ Duranti, L., Eastwood, T. & MacNeil, H., *Preservation of the Integrity of Electronic Records* (Dordrecht: Kluwer Academic Publishing, 2002) p.27.

⁴⁰⁰ Duranti, L., *Diplomatics: New Uses for an Old Science* (Lanham; London: Scarecrow Press, Inc., 1998) p.71.

⁴⁰¹ *Ibid.*, p.72.

the record – and character is something established in public, which concerns reputation. The other dynamic concerns legal responsibility – it absolves the user of the record from ascertaining the veracity of the records in that any actions that follow on the basis of information provided by the signatory are traceable to the signatory. Robertson observes that this link persisted through the move to mass-produced documents: ‘The tension between replication and authenticity that characterized the emergence of mass-produced documents is evident in the concern that the person who signed the passport, who gave the document the semblance of personal authority, was turned into a machine by the excessive need to repeat his signature’.⁴⁰²

The non-repudiation function of the signature also relates to the issue of accuracy, which is determined at the point of the creation of the record and concerns the character and ability of the creator. In their report on their InterPARES study *Authenticity, Reliability and Accuracy of Digital Records in the Artistic, Scientific and Governmental Sectors*, John Roeder, Philip Eppard, William Underwood and Tracey Lauriault note that there ‘is concern with accuracy of information, but it is likely to be limited to the accuracy of data at the point of creation, after which accuracy (like authenticity) is presumed to be protected by procedural controls’.⁴⁰³ In discussing the definitions in the United States E-Government Act of 2002, they observe that ‘...accuracy is regarded as important, but it is assumed as a given requirement for the content of records and thus is not given much direct attention here or elsewhere in the literature’.⁴⁰⁴ This may explain why accuracy very quickly disappears from the InterPARES research agenda. Indeed, later discussions of trustworthiness often only mention two components - reliability and authenticity.

Returning to the view of signatures as admissions of responsibility for reliability, and by extension, for accuracy, we find that it accords with Vismann’s argument

⁴⁰² Ibid., p.61.

⁴⁰³ Roeder et al, ‘Part Three—Authenticity, Reliability and Accuracy of Digital Records in the Artistic, Scientific and Governmental Sectors: Domain 2 Task Force Report’ in Duranti, L. & Preston, R. (eds.), *International Research on Permanent Authentic Records in Electronic Systems (InterPARES 2): Experiential, Interactive and Dynamic Records* (Padova: Associazione Nazionale Archivistica Italiana, 2008) p.9.

⁴⁰⁴ Roeder et al., ‘Authenticity, Reliability, and Accuracy’, pp.8-9.

that the law operates ‘not *in mundo* but in the medium of literality; it believes only what is written – more precisely, what it has itself written down’.⁴⁰⁵ In effect, the truthfulness of all records is suspect, but records function as trustworthy or not in particular contexts, such as legal regimes. MacNeil acknowledged this when she wrote that the ‘methods for assessing a record’s truth-value are underpinned specifically by an assumption that a unitary and stable relationship exists between a representation (that is, a record) and its referent (i.e., a pre-existent reality)...’.⁴⁰⁶ With government records, there is no unitary and stable relationship between the record and a pre-existent reality, but one between the record and the reality described by the bureaucrat, which becomes ‘the fact’ for the system of law. There will be gaps in these records - archival silences – and as Ann Stoler has written, reading ‘along the archival grain’, that is, reading records with a view to what is missing and how their language constructs a narrative, reveals the role of the record in creating or reinforcing a narrative and a set of power relations.⁴⁰⁷ But ultimately, the truth of the record is the truth of the system of government that produces and uses the record, and it is on the basis of that ‘truth’ that the system will operate and the individual will necessarily proceed in interactions with the system.⁴⁰⁸ That is, unless mechanisms can be built in for correcting or augmenting the record. This notion is at the heart of new work emerging from the University of California Los Angeles (UCLA) Department of Information Studies and the Records Continuum Research Group (RCRG), based in the Monash University Faculty of Information Technology. The UCLA and RCRG work is interrogating accepted principles of archival theory and looking at how counter-narratives can be introduced into archives. Notions emerging from this discourse include co-creation (the idea that ‘subjects’ of records might be considered ‘co-creators’ of records), ‘rights in records’ (the idea that co-creators or others might have their rights reflected in the management of records, such as the rights to

⁴⁰⁵ Vismann, C., *Files: Law and Media Technology*, trans. by Winthrop-Young, G., (Stanford: Stanford University Press, 2008) pp.56-57.

⁴⁰⁶ MacNeil, ‘Trusting Records in a Postmodern World’, p.42.

⁴⁰⁷ Stoler, A. L., *Along the Archival Grain: Epistemic Anxieties and Colonial Common Sense* (Princeton; Oxford: Princeton University Press, 2009).

⁴⁰⁸ “...the integrity of a record cannot be made in any absolute sense but, rather, in relation to the purpose the record serves in the environment in which it has been created, maintained, and used”, MacNeil, H., ‘Providing Grounds for Trust’, p.70.

challenge, annotate or correct archival descriptions of records, proposed by Anne J. Gilliland and Sue McKemmish)⁴⁰⁹, and participatory archives, a concept which

acknowledges that multiple parties have rights, responsibilities, needs and perspectives with regard to the archives. Participatory archives consequently become a negotiated space built around critical reflection in which these different communities share stewardship and expertise - they are created by, for and with multiple communities, according to and respectful of community values, practices, beliefs and needs. They may also work in partnership with institutional archives, including government archives and collecting institutions.⁴¹⁰

These ideas are changing archival theory and practice, and may change our understanding of the reliability of records. However, under traditional government record-making and record-keeping practices, the truth of the record is the truth for the system that made and uses the record.

At the time of signing, responsibility for the reliability of the record is assigned, but the record then passes into circulation within the bureaucracy, into use by various officers and perhaps eventually into an archive for preservation. It is during use and preservation that authenticity becomes the predominant concern.

5.4.2. Authenticity

The work of the UBC scholars characterised authenticity as being reliant on the integrity of the record, which is the 'quality of being complete and unaltered in all essential respects', and the identity of the record, which is

⁴⁰⁹ Gilliland, A. J. & McKemmish, S., "Rights in Records as a Platform for Participative Archiving" in Cox, R. J., Langmead, A. & Mattern, E. (eds.) *Archival Research and Education: Selected Papers from the 2014 AERI Conference* (Sacramento: Litwin Press, 2015) pp.355-385; Gilliland, A. J. & McKemmish, S., "The Role of Participatory Archives in Furthering Human Rights, Reconciliation and Recovery," *Atlanti: Review for Modern Archival Theory and Practice*, 24.1 (2014).

⁴¹⁰ Gilliland, A. J. & McKemmish, S., 'The Role of Participatory Archives in Furthering Human Rights, Reconciliation and Recovery', *Atlanti: Review for Modern Archival Theory and Practice*, 24.1 (2014) p.82.

defined as the ‘whole of the characteristics of a document or a record that uniquely identify it and distinguish it from any other document or record’.⁴¹¹ These two elements of authenticity require closer analysis.

The InterPARES Terminology Database defines ‘integrity’ as follows.

The integrity of a record refers to its wholeness and soundness: a record has integrity when it is complete and uncorrupted in all its essential respects... its physical integrity, such as the proper number of bit strings, may be compromised, provided that the articulation of the content and any required annotations and elements of documentary form remain the same.⁴¹²

The UK National Archives report *Generic Requirements for Sustaining Electronic Information Over Time* reflects the privileged place of documentary form and content over media or carrier in the InterPARES regime, stating ‘a record is considered to be essentially complete and uncorrupted if the message that it is meant to communicate in order to achieve its purpose in unaltered’.⁴¹³ Paper may deteriorate or bit strings may alter, but the integrity of a record is intact if the content and documentary form remain. By extension, threats to integrity are threats to authenticity, and any controls that protect the content and form of the record offer some protection to the record’s authenticity.

In the InterPARES work, the identity of a record is conceived of as those attributes that ‘uniquely characterize it and distinguish it from other records’.⁴¹⁴

The InterPARES Authenticity Task Force stated that, from an archival-diplomatic perspective, those attributes include:

⁴¹¹ InterPARES 2, *Terminology Database*, http://www.interpares.org/ip2/ip2_terminology_db.cfm [last accessed 15 March 2019].

⁴¹² Ibid.

⁴¹³ The National Archives (UK), *Generic requirements for sustaining electronic information over time: 1 Defining the characteristics for authentic records*, (Kew: The National Archives, 2003) p.8.

⁴¹⁴ InterPARES 1 Authenticity Task Force, ‘Appendix 2 - Requirements for Assessing and Maintaining the Authenticity of Electronic Records’, in InterPARES 1 Project, *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project* (InterPARES Project, 2002) pp.1-2.

the names of the persons concurring in its formation (that is, its author, addressee, writer, and originator); its date(s) of creation (that is, the date it was made, received, and set aside) and its date(s) of transmission; an indication of the action or matter in which it participates; the expression of its archival bond, which links it to other records participating in the same action (for example, a classification code or other unique identifier); as well as an indication of any attachment(s) since an attachment is considered an integral part of a record.⁴¹⁵

These attributes are essentially metadata that give the record its uniqueness. Metadata, then, is integral to the identity of a record and, by extension, integral to ensuring its authenticity.

5.5. Accuracy, Reliability or Authenticity?

In this schema of values, then, *trustworthiness* requires *accuracy*, *reliability* and *authenticity*. Accuracy and reliability concern the truth-value of records, and are at issue in the creation of a record. Authenticity concerns the quality of a record being what it purports to be, and is at issue in the use and preservation of records. Authenticity requires *identity* and *integrity*.

That a record could be reliable without being authentic is explained by Caroline Williams: '[m]edieval forgeries were often seen as a legitimate attempt to "put the record straight", and although spurious in that they were not produced "authentically" could be accurate and "reliable" in the essentials of the information recorded'.⁴¹⁶ The InterPARES distinction between the ideas of reliability and authenticity follows the same lines drawn by Jeremy Bentham when he wrote:

⁴¹⁵ InterPARES 1 Authenticity Task Force, 'Appendix 2 - Requirements for Assessing and Maintaining the Authenticity of Electronic Records', in InterPARES 1 Project, *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project* (InterPARES Project, 2002) pp.1-2.

⁴¹⁶ Williams, C., 'Diplomatic Attitudes: From Mabillon to Metadata,' *Journal of the Society of Archivists*, 26.1 (2005) p.11.

A distinction must here be observed, between evidence of authenticity, and evidence of fairness. Authenticity may be proved by similitude of hands: it may be proved, provisionally at least, *ex tenore*, with or without the other presumption *ex custodia*. To the question of fairness, none of these media of proof, it is evident, can apply. A bond is produced in evidence: the obligor may have been in a state of insanity or intoxication when he executed it: he may have executed it with the fear of a pistol or a dagger before his eyes, or in a state of illegal imprisonment, to which he had been subjected for that purpose. Of none of these modifications will the signature, or the custody of the instrument, or the tenor of it, afford any sort of warning.⁴¹⁷

Bentham asserts that authenticity may be adduced from the signature, the custody of the instrument and its 'tenor', in the legal sense of that word, which is 'the wording of the document' (in other words, that it follows certain norms of official language), but that reliability – what he calls 'fairness' – cannot be judged by the same means.

In monitoring government activity, authenticity is key, because, following the logic of the business exception to the hearsay rule and Vismann's argument that the law operates in the medium of literality, monitoring is concerned with what bureaucracy has said, as the basis upon which it operates. Discussing the distinction between reliability and authenticity, Williams observed that issues of fact and truth are difficult for those concerned with keeping and preserving records, 'areas which some might argue are not our concern', and indeed the controls developed in archival science are primarily concerned with the protection of authenticity. Williams writes: 'The archivist, in facilitating the creation of authentic records which are to some degree capable of standing for the facts to which they attest, must also acknowledge that their content will remain open to interpretation'.⁴¹⁸ Potentially, the current UCLA / RCRG work will reposition the concept of reliability as a matter for record-keepers, but as noted above, it is too

⁴¹⁷ Jeremy Bentham quoted in MacNeil, *Trusting Records*, p.25.

⁴¹⁸ Williams, 'Diplomatic Attitudes', p.19.

soon to gauge the impact of that work on record-keeping practice. Similarly, it would be interesting to consider metering mechanisms as methods for assessing reliability, but that is outside of the scope of this work.

5.6. Establishing and Maintaining Authenticity

The work of InterPARES 1 was divided between task forces. An Authenticity Task Force defined *Requirements for Assessing and Maintaining the Authenticity of Electronic Records* (2002), which comprised ‘requirements that support the presumption of the authenticity of electronic records before they are transferred to the custody of the preserver, and requirements that support the production of authentic copies of electronic records that have been transferred to the custody of the preserver’.⁴¹⁹ Using two sets of requirements (Requirement Set A - *Benchmark Requirements Supporting the Presumption of Authenticity of Electronic Records* and Requirement Set B - *Baseline Requirements Supporting the Production of Authentic Copies of Electronic Records*), the report attempts to articulate a calculable approach to determining authenticity, reflecting the civil law mindset of diplomatics and Duranti, the project’s leader. The InterPARES Authenticity Task Force posited that presumptions of authenticity could be based on the number of requirements that had been met and the degree to which they had been met, saying that the ‘requirements are, therefore, cumulative: the higher the number of satisfied requirements, and the greater the degree to which an individual requirement has been satisfied, the stronger the presumption of authenticity’.⁴²⁰ Epistemologically, this approach is rooted in the civil law tradition in which the value of evidence can be calculated.⁴²¹ I have not been able to find an example of this approach being practiced in assessments of the authenticity of digital records as evidence, but nevertheless the requirements articulate principles that may be helpful in assessing the consideration given to authenticity in open government data.

5.6.1. Metadata

⁴¹⁹ InterPARES 1 Authenticity Task Force, ‘Appendix 2’, p.1.

⁴²⁰ InterPARES 1 Authenticity Task Force, ‘Appendix 2’, p.3.

⁴²¹ Duranti, L., *Diplomatics: New Uses for an Old Science*.

'Requirement A.1: Expression of Record Attributes and Linkage to Record' details attributes (metadata) of records in two categories – the first concerning the identity of the records and the second concerning the integrity of the record. The identity attributes or metadata include names of authors and addressees, dates of transmission and receipt, classification codes, etc. The integrity attributes or metadata include the name of the handling office or office of primary responsibility and indications of annotations and technical modifications.⁴²² These requirements are essentially a basic set of metadata concerned with documenting the record's provenance and the context of the record's creation and use. The presence of metadata for identity establishes the uniqueness of the record, while the presence of metadata for integrity allows an examination of changes to the record after creation. The stronger or more complete these two categories of metadata, the greater the trustworthiness of the record.

5.6.2. Systems

There are significant limitations to the application of diplomatics to modern records, not least because modern bureaucracies adhere less strictly to formal conventions of document production. The absence of a baseline against which to gauge the authenticity of records informed early criticisms of Mabillon's approach. Barthélémy Germon, in his *De veteribus regum Francorum diplomatibus* of 1703, observed that Mabillon's samples, intended as benchmarks against which documents could be assessed for adherence to conventions, could not be established as authentic themselves.⁴²³ Even when procedural and documentary norms can be established, they may be of limited use. MacNeil acknowledged this when, in reporting the findings of the Authenticity Task Force of InterPARES, she wrote that the Task Force found that 'the determination of documentary forms in general and the establishment of required elements of form and annotations in particular were deeply embedded within specific institutional and

⁴²² InterPARES 1 Authenticity Task Force, 'Appendix 2', pp.5-4.

⁴²³ Head, R., 'Documents, Archives, and Proof around 1700', *The Historical Journal*, 56.4 (2013) p.926.

procedural contexts and were resistant to any easy generalizations'.⁴²⁴ Modern office work, being much freer of procedure than earlier expressions, produces a wide variety of record types. Similarly, the diversification of technological systems has resulted in a wide variety of formats. Faced with a diversity of practices, systems and formats, true diplomatics can have only limited use as an analytic approach. This fact caused MacNeil to consider 'a reorientation of its concepts and principles'.⁴²⁵ She suggested that this might mean 'focusing less attention on establishing whether the record is complete, stable, and unchangeable, and more attention on determining whether and to what extent the system is capable of tracking changes and how that tracking function might be managed over time'.⁴²⁶

Requirement Set A includes requirements for systemic controls that support a presumption of authenticity. These include defined and implemented access privileges (Requirement A.2.), protective procedures related to the loss and corruption of records (A.3) and media deterioration and technological change (A.4), the establishment of documentary forms (A.5), rules governing the authentication of records (A.6), methods for distinguishing authoritative versions from copies (A.7) and procedures for the removal or transfer of electronic records from systems (A.8). This turn towards the capabilities of the system as an indication of the authenticity of the records held within it was acknowledged in the early InterPARES work: applications of the InterPARES 1 *Template for Analysis* (a template of record characteristics) were case studies of systems, not records.⁴²⁷ InterPARES research has found that an important component of vouchsafing authenticity is 'to embed procedural rules for creating, handling, and maintaining... records in an agency-wide records system, and to integrate documentary procedures with business processes'.⁴²⁸ This is what more recent standards for record-making and record-keeping system functionality attempt to do. For instance, the international standard *ISO 16175:2012 - Information and documentation - Principles and functional requirements for records in electronic*

⁴²⁴ MacNeil, H., 'Providing Grounds for Trust II: The Findings of the Authenticity Task Force of InterPARES', *Archivaria*, 54 (2002) p.31.

⁴²⁵ Ibid. p.32.

⁴²⁶ Ibid.

⁴²⁷ MacNeil, 'Providing Grounds for Trust', p.56.

⁴²⁸ MacNeil, *Trusting Records*, p.98.

office environments promotes the incorporation of records management functionality into the business systems that create records, such as payroll or personnel management systems, so that record-making and record-keeping are more closely connected, and blockchains arguably unify the transaction and the record into one event.⁴²⁹ Closing the gap between record-making and record-keeping has a significant effect on trustworthiness, in that it moves the record and the system closer to affording an assurance of reliability. If reliability is a consequence of the procedures and the thoroughness of record-making, and responsibility for record-making is documented in automatically ascribed metadata that can be scrutinised, the author of the record is subject to the non-repudiation function of the signature and the legal and reputational consequences that may flow from that. In this way, the system operates on the author in the same way that the author operates on the system, bringing to mind Pasquinelli's concept of valorising and control information. While systems are susceptible to manipulation by people, the reverse may also be true. Panoptically, the system can promote trustworthiness by promoting ethical behaviour. This is significant for monitoring in principal-agent problems, so long as the system and its metadata are open to scrutiny.

It is important to note that this 'increased attention to the system' was not new attention. In archival theory, the consistent application of procedure and adherence to systems has been equated with authenticity, at least since the time of Leon de Laborde, Director General of the Imperial Archives during Napoleon III's Second Empire between 1857 and 1868,⁴³⁰ who saw processing archives with *respect des fonds* as an important underpinning of a presumption of authenticity.

⁴²⁹ Diplomatics distinguishes the 'moment of action' (*actio*) from the 'moment of documentation' (*conscriptio*). In diplomatic thinking, the closer together these things are, the more reliable the record. This is something that digital systems facilitate. Citing Charles Babbage's self-registering apparatus for trains and Etienne-Jules Marey's various medical-notation machines, Greg Siegel writes '...the graphic methodologists of the nineteenth century thought it imperative that the phenomenon under investigation author its own inscriptions... The requirement that the motional body originate its own tracings – a requirement that represents the cornerstone of the graphic method – assumes and implies both the *automaticity* of the recording process and the *authenticity* of its products'. Siegel, G., *Forensic Media: Reconstructing Accidents in Accelerated Modernity*, (Durham: Duke University Press, 2014) p.71.

⁴³⁰ MacNeil, H., 'Trusting Description: Authenticity, Accountability, and Archival Descriptive Standards', *Journal of Archival Organization*, 7.3 (2009) p.91.

5.6.3. Custodianship

In his 1922 manual of archives administration, Sir Hilary Jenkinson states that archivists should not take into archival custody any records that do not have an 'Archive quality'⁴³¹ unless they fill gaps in existing series, noting that these records should be identified by an appropriate mark. In a footnote to this point, he stated: 'At the Public Record Office such recovered strays are stamped with a special inscription "some time out of custody"'.⁴³²

The notion of custodianship is important in both law and archival science. MacNeil has observed that '[f]or public documents, the fact that the records have been found in official custody has long been accepted as sufficient external evidence of their authenticity'.⁴³³ For Jenkinson, custody concerned the two interrelated concepts of *trusted custodians* and *chains of custody*. His *Manual* describes the need for an unbroken chain of custody of records to ensure their authenticity, and talks about the professional obligations of archivists as trusted custodians.⁴³⁴

Turning first to chains of custody, the concept accords with the logic of the business entry exception to the hearsay rule in that it assumes authenticity from creation through official use, before the record is taken into archival custody. The concept of the chain of custody is therefore concerned with the maintenance of authenticity through time, which, according to the InterPARES Preservation Task Force, is complicated in the digital environment:

⁴³¹ Jenkinson states that documents have an 'archive quality' when they have 'impartiality' – that they have been created for the use of their creator, rather than to tell a story (here Jenkinson echoes the logic of the business entry rule in US evidence law) - and 'authenticity' – that they have not been tampered with. (Jenkinson, H., *A Manual of Archive Administration*, (London: Percy Lund Humphries, & Co LTD., 1937)) p.12.

⁴³² Jenkinson, *Archive Administration*, p.43.

⁴³³ MacNeil, *Trusting Records*, p.47.

⁴³⁴ An anecdote widely known amongst British archivists holds that during Sir Hilary's time as Deputy Keeper of the Public Records (1947 - 1954) the requirement for an unbroken chain of custody led to keepers taking it in turns to spend nights in the Public Record Office. Presumably nightwatchmen did not constitute 'trusted custodians' because they were not bound by the same professional ethics as archivists.

Given that the storage and retrieval processes for electronic records inevitably entail physical and representational transformations, the traditional concept of an unbroken chain of custody need to be expanded to encompass processes that are necessary to ensure that an electronic record is transmitted over time without inappropriate alteration.⁴³⁵

This expanded concept was called the ‘chain of preservation’:

In the digital world, a meaningful presumption of authenticity depends on what Ken Thibodeau and others term “an unbroken chain of preservation,” which entails the implementation of processes and procedures that protect and preserve the identity and integrity of records from the moment of their creation and throughout their life.⁴³⁶

While the chain of custody describes the general notion of custody throughout the life of records, the InterPARES Chain of Preservation (COP) Model acknowledges that digital records require not only custody but intervention to ensure that the identity and integrity of records remain intact over time. The COP Model attempts to identify all likely actions on or with records with the measures the InterPARES researchers deem necessary for safeguarding authenticity. The COP model is an elaboration of the controls identified elsewhere in the InterPARES work; metadata, custodianship, system controls.

Turning to the notion of trusted custodians, the role of custodian has been the subject of professional codes of ethics. At its General Assembly in Beijing in September 1996, the International Council on Archives adopted its *Code of Ethics*, which states:

The primary duty of archivists is to maintain the integrity of the records in their care and custody. In the accomplishment of this duty they must

⁴³⁵ InterPARES 1 Preservation Task Force, ‘Appendix 6 – How to Preserve Authentic Electronic Records’, in InterPARES 1 Project, *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project* (InterPARES Project, 2002).

⁴³⁶ MacNeil, H., ‘Trust and professional identity: narratives, counter-narratives and lingering ambiguities’, *Archival Science*, 11.3-4 (2011) p.183.

have regard to the legitimate, but sometimes conflicting, rights and interests of employers, owners, data subjects and users, past, present and future. The objectivity and impartiality of archivists is the measure of their professionalism. They should resist pressure from any source to manipulate evidence so as to conceal or distort facts.⁴³⁷

The role of custodian has also been acknowledged in technical standards. It is recognised in guidance on digital records and archives, such as the Open Archival Information Systems (OAIS) standard, *ISO 14721:2003 - Space data and information transfer systems – Open archival information systems – Reference model*, which defines the archive's management function, and of *ISO16175:2011 - Information and documentation - Principles and functional requirements for records in electronic office environments*; the system administrator's responsibilities implicit in its requirements make ethical custodianship a key component of a trusted environment.

As with the idea of a chain of custody, the idea of the custodian has also developed since Jenkinson's time. In 1981, F. Gerald Ham saw that the proliferation of records and the decentralisation of control over records (together characterising what he called the 'post-custodial era') would require a more proactive and networked approach to archival management.⁴³⁸ Ten years later, David Bearman would radically expand the networked approach by suggesting distributed custody as an approach to managing digital records.⁴³⁹ In a rejection of this post-custodial thinking, in 1996, Duranti elaborated the much older idea of an 'archival threshold', which divides the place of a document's creation from that of its preservation. Having argued that archives have, since ancient times, been places that offer an assurance of the trustworthiness of records, she wrote:

⁴³⁷ International Council on Archives, *Code of Ethics*, (1996) p.1. Note that the use of the term 'integrity' in this quote, and 'reliability' at other places in the *Code of Ethics*, does not necessarily conform with the InterPARES definitions of these terms, which were developed at least six years after the adoption of the *Code*, during InterPARES II (2002 – 2007).

⁴³⁸ Ham, F. G., 'Archival Strategies for the Post-Custodial Era', *The American Archivist*, 44:3 (1981) pp.207-216.

⁴³⁹ Bearman, D., 'An Indefensible Bastion: Archives as a Repository in an Electronic Age,' *Archives and Museum Informatics Technical Report 13* (1991) pp.14-24.

Somewhere between the outside and the inside of the archival building, the documents must unfold into evidence and memory, prior to being ensconced within the building as testimony of past actions. There must be a space, an inbetween space, where this happens, a space bound by two limits, one bordering the documents and the other bordering the evidence: the *archii limes* or “archival threshold.”⁴⁴⁰

For Duranti, the archival threshold is the ‘space where the officer of the public authority takes charge of the documents, identifies them by their provenance and class, associates them intellectually with those that belong in the same aggregation, and forwards them to the inside space’ and through these activities, authenticated records.⁴⁴¹ ‘At the archival threshold, and beyond it, the authenticating function took place’.⁴⁴²

Like the signature, the acceptance of records into the archives was a statement that performs a non-repudiation function. Duranti was invoking the *ius archivi*, a body of writing that emerged in Germany in the 17th century that employed the Roman notion of *fides publica* (public faith) to define public archives as trustworthy repositories. According to the *ius archivi* theorists, documents from public repositories that met certain criteria could be presumed authentic. Duranti elaborates the idea of the threshold by reference to the *ius archivi* theorist Ahasver Fritsch:

...in 1664, [he] commented that archival documents did not acquire authenticity by the simple fact of crossing the archival threshold, but by the fact that 1) the place to which they were destined belonged to a public sovereign authority, as opposed to its agents or delegates, that 2) the officer forwarding them to such a place was a public officer, that 3) the documents were placed both physically (i.e., by location) and intellectually (i.e., by description) among authentic documents, and that 4) this association was not meant to be broken.⁴⁴³

⁴⁴⁰ Duranti, L., ‘Archives as Place’, *Archives and Manuscripts* 24.2 (1996) pp.243-244.

⁴⁴¹ Ibid.

⁴⁴² Ibid.

⁴⁴³ Ibid., p.244.

The *ius archivi* was an alternative to diplomatics, one that privileged the custodial institution's role in authentication over the significance of documents' adherence to established documentary forms. Assessments were made of the archival institution rather than the records themselves. Randolph Head demonstrates that the two discourses – diplomatics and the *ius archivi* - were in contact, and were at least to a limited extent, in opposition. Head wrote that, for Germon, 'the lack of repositories characterized by public faith from the earliest periods represented a critical problem that Mabillon's method could not overcome. "Bring me charters from public *tabularia*, or at least from those that have been constantly protected from the perfidy of forgers", Germon demanded'.⁴⁴⁴

Although public faith in public institutions can no longer be assumed, the *ius archivi* and its influence on the more recent debate over archival custody is helpful in thinking about authenticity insofar as the quality of custodianship may be examinable, offering another assurance of trustworthiness. While the two schools of thought may have been in opposition, their influences in the present day do not have to be treated as contradictory. Indeed, Duranti has used both in combination, and Mak, in her entry for 'Authenticity' in the *Encyclopedia of Archival Science*, wrote that contemporary archival science's conceptualisation of authenticity 'blends the perspective of early modern diplomatists and historians, in which material form, formulae and format were paramount, with that of the *ius archivi* ramified through Hilary Jenkinson's emphasis on continuous control'.⁴⁴⁵

Notwithstanding the professional debates about custody and post-custodialism that have continued to the present day, in practice, archives are still taking custody of records, whether hardcopy or digital. For instance, Anne Thurston has noted the transfer and ingest of digital records into digital repositories in Estonia, Finland and Norway as part of a larger programme of record-keeping and

⁴⁴⁴ Head, R., 'Documents, Archives, and Proof around 1700', p.926.

⁴⁴⁵ Duranti, L. & Franks, P. C. (eds.), *Encyclopedia of Archival Science* (Lanham: Rowman and Littlefield, 2015) p.120.

government accountability.⁴⁴⁶ So, we may say that an ‘archival threshold’ exists in the digital environment and custodianship continues to be a consideration in assessing the authenticity of records; to some extent, trust in archival records is connected with the trust in archival repositories and workers.

Archives have had to be more proactive and interventionist than they once were because of the nature of digital records and systems. The changeable nature of digital systems and the fragility of digital information requires preservation controls to be introduced at least from the point of creation. This has been reflected in an increased integration between systems for record-making and systems for record-keeping during the use of records. Following the appraisal of records as warranting permanent preservation in archives, their transfer across the archival threshold and subsequent preservation in digital repositories must be managed in such a way that the identity and integrity of the records remains intact.

Custodianship and chains of custody are becoming increasingly automated, which implies that they could also be, or become, more closely auditable, so that the implicit guarantee of authenticity afforded by the archives ‘as a place’, that is, as an institution, could become an explicit guarantee, via audit trails. In effect, these audit trails are the digital equivalents of the public witnesses to the livery of seisin. Through technology, it is possible for custodianship to be done ‘in public’.

5.7. A Note on Originality and Fixed Forms

The InterPARES 1 Authenticity Task Force’s *Requirement Set B*, which concerns the production of authentic copies of records, comprises Requirement B.1:

Controls over Records Transfer, Maintenance, and Reproduction, Requirement B.2: Documentation of Reproduction Process and its Effects and Requirement

⁴⁴⁶ Thurston, A., *Managing Records and Information for Transparent, Accountable, and Inclusive Governance in the Digital Environment: Lessons from Nordic Countries (English)*, (Washington: World Bank Group, 2015)
<http://documents.worldbank.org/curated/en/222041468189536195/Managing-records-and-information-for-transparent-accountable-and-inclusive-governance-in-the-digital-environment-lessons-from-Nordic-Countries> [accessed 22 August 2018].

B.3 – Archival Description.⁴⁴⁷ Requirement Set B’s treatment of originals and copies foregrounds the principle of ‘best evidence’, in which the original is considered better evidence than a copy, but as the interpretation or adherence to the best evidence rule has changed over time, so has the interpretation or adherence to ideas of originals and copies. While the procedural and legal imperatives for the authentication of copies persists in many jurisdictions, thinking about the nature of digital originals has shifted in line with a deepening understanding of the implications of the digital environment, and the introduction of new perspectives from postmodern positions, and particularly the work of the Records Continuum Research Group. In a move in this direction, MacNeil and Mak wrote in 2007:

The preservation of “authentic” digital materials is perceived to have been made more difficult by the technology itself, which promotes the proliferation of multiple and simultaneous instantiations. For which of these instantiations falls under the responsibility of the librarian or archivist, and what constitutes its authenticity? Because these questions cannot be answered with any specificity, the management of digital materials may involve shifting the prevailing attitude from an approach that is characterized by constriction to one of expansion.⁴⁴⁸

This is indicative of the trajectory of the archival discourse on authenticity – a move from a presumption that a digital original exists and can be identified, including over time, to an acceptance that digital technology has fundamentally changed the nature of records. ‘Although there may be a moment in which a digital record is original, in the sense of being the first complete and effective instantiation, such moment [sic] is ephemeral; for all intents and purposes, then, digital records live their lives as reproductions’.⁴⁴⁹

The Authenticity Task Force’s *Draft Requirements for Authenticity* stated that ‘any electronic copy of an authentic electronic record is authentic if declared to be so

⁴⁴⁷ InterPARES 1 Authenticity Task Force, ‘Appendix 2’, pp.7-8.

⁴⁴⁸ MacNeil, H. & Mak, B., ‘Constructions of Authenticity,’ *Library Trends*, 56.1 (2007) p.45.

⁴⁴⁹ MacNeil, ‘Trust and professional identity’, p.182.

by an officer entrusted with such function'.⁴⁵⁰ This reflects the InterPARES thinking about the archival threshold and the authenticating power of the archive, and is telling about the nature of government records – they are authentic if the archive (as an apparatus of the law) states that they are authentic. As Vismann wrote, the law believes what it has, itself, written down.

Whereas paper records may have fixed and final forms, such forms do not naturally manifest in digital records. This has been problematic for adducing authenticity, leading to increased attention to the system within which the record resides, and its ability to be trustworthy by controlling and recording interactions with the record, as discussed above. The approach of the National Archives of Australia (NAA) is markedly different from the traditional view. NAA has

...adopted a performance model of a digital record in which digital records are understood as “a series of performances across time” rather than as fixed and stable artefacts. In the performance model, the “essence” of a digital record comprises those characteristics of its context, rendition and structure that need to be preserved together with the content to ensure the record’s authenticity.⁴⁵¹

This further removes the burden of authenticity from the record to the system that will preserve and render the record. Nevertheless, fixed forms are essential to some types of record, and attempts have been made to capture the elements of those records as they appear at creation or on the designation of a document as a record. For instance, the Portable Document Format (PDF) was developed in the early 1990s to enable the various elements of documents to be fixed, allowing consistency of rendition across platforms. The Portable Document Format / Archival (PDF/A) format, described in ISO standards in 2005, is ‘intended to address the areas the PDF format is lacking in terms of long-term digital preservation, for example, by prohibiting encryption and requiring embedding of

⁴⁵⁰ MacNeil, H., ‘Providing Grounds for Trust: Developing Conceptual Requirements for the Long-Term Preservation of Authentic Electronic Records,’ *Archivaria* 50 (2000) p.68.

⁴⁵¹ MacNeil, ‘Trust and professional identity’, p.182.

fonts necessary for rendering, so as to eliminate external dependencies'.⁴⁵² The Library of Congress guidance on the sustainability of digital formats notes that the PDF/A standards 'attempt to maximize: Device independence, Self-containment, Self-documentation'.⁴⁵³ Migration to PDF/A, as with all other file format normalisation, underscores the move away from the idea of an original, emphasizing the essential elements of the records as their content and documentary form, which may be best protected by moving them out of their original format. Intervention by a custodian is increasingly likely to be necessary if the identity and integrity of records are to be preserved.

There is an international standard for digital evidence security techniques, *SIO/IEC 27037:2012 Information technology – Security techniques – Guidelines for identification, collection, acquisition, and preservation of digital evidence*. This standard defines reliable evidence using the term known to archival-diplomatics as authenticity – that it 'is what it purports to be'.⁴⁵⁴ The standard prioritises trusted persons establishing physical and intellectual control over the evidence as quickly as possible. The standard relies heavily on metadata about context events, personnel, etc. for instance, acquiring 'a digital evidence copy' requires 'documenting the methods used and activities performed'.⁴⁵⁵ The standard has a focus on documented procedures, and with respect to custodianship, Annex A to the standard is a table of core skills and competencies for a DEFR (Digital Evidence First Responder). The trustworthiness of the custodian becomes tied to the trustworthiness of the record.

5.8. Conclusion

The InterPARES Project is now in its fourth iteration⁴⁵⁶ (InterPARES Trust 2013 –

⁴⁵² Koo, J. & Chou, C. H., 'PDF to PDF/A: Evaluation of Converter Software in Digital Repository Workflow', *New Review of Information Networking*, 18.1 (2013) p.2.

⁴⁵³ Library of Congress, *Sustainability of Digital Formats: Planning for Library of Congress Collections*. <https://www.loc.gov/preservation/digital/formats/> [last accessed 15 March 2019].

⁴⁵⁴ International Organization for Standardization, *SIO/IEC 27037:2012 Information technology – Security techniques – Guidelines for identification, collection, acquisition, and preservation of digital evidence* (Geneva: International Organization for Standardization, 2012) p.6.

⁴⁵⁵ *Ibid.*, p.9.

⁴⁵⁶ InterPARES 1 (1999-2000), InterPARES 2 (2002-2007), InterPARES 3 (2007-2012) and InterPARES 4 (2013-2018), <http://www.interpares.org> [accessed 22 August 2018].

2018) and its scope has widened considerably, with seven international research teams looking at diverse questions from trustworthiness in cloud environments to archival education provision in Africa. The conversation about what constitutes informational authenticity in the digital environment is still evolving and new techniques for gauging and protecting authenticity will develop. Advances in digital technologies and techniques, the expansion of postmodern theory into archival science, and convergences between disciplines have sustained and expanded this conversation. Furthermore, new circumstances cause new groups to engage with the issue of informational authenticity. For instance, on 8 June 2015, Thomson Reuters Foundation reported on a new mobile phone app 'enabling civilians in conflict-torn countries to capture and share verifiable footage of war crimes'. They reported:

Mobile phone footage of human rights abuses, mainly shared on social media in recent years, is often fake, impossible to verify or lacking the information necessary to be used as evidence in court, said the International Bar Association (IBA). The "EyeWitness to Atrocities" app records the user's location, date and time, and nearby Wi-Fi networks to verify that the footage has not been edited or manipulated, before sending it to a database monitored by a team of legal experts.⁴⁵⁷

In this example, we see an interplay between concepts familiar to archival science, law and diplomacy: metadata that helps to establish the identity of the footage as a record, the transmission of the footage across an archival threshold and management of a repository by custodians, all systematised. The concepts that have been developed in archival science, law and diplomacy offer ways of thinking about and articulating approaches to the protection and preservation of the authenticity of information. From the history of the idea of the authenticity of information, it is possible to draw out some threads:

- Metadata: Metadata can be used to establish identity and integrity. Metadata

⁴⁵⁷ Guilbert, K., 'Smartphone-wielding civilians to combat war crimes – with an app', *Thomson Reuters Foundation* (7 June 2015) <http://news.trust.org/item/20180607151711-ymjan/> [accessed 20 May 2018].

can be used to describe the context and provenance of records and, as it accumulates around records, it builds up the data that constitutes audit trails.

- **Systems of Control:** All these types of metadata must be captured by a system that can store the record and its metadata, and maintain the links between the record and its metadata, while controlling and tracking access and actions on the records and their metadata.
- **Custodians:** These systems must be managed in such a way that a chain of custody is established. Like the metadata and the systems, the custodians must be trustworthy. In assessing the trustworthiness of government information, it is necessary to have some assurance of the character of custodians and to be able to interrogate the actions of those with the ability to manipulate or protect that information.

A common thread through all of these controls is 'publicity'. As MacNeil observed, '...methods for assessing and ensuring record trustworthiness... are rooted in observational principles'.⁴⁵⁸ We must be able to see what has been done to information. Is it possible to see the metadata? Is it possible to see that the system has provided uninterrupted control? Is it possible to see who has done what with the records and the system in which they are maintained? In other words, can the metadata, systems and custodians be audited or interrogated? The monitoring of government requires information that, in order to be trustworthy, can itself be monitored.

It is important to acknowledge the subjective nature of trustworthiness and its constituent parts. MacNeil has written that '[t]he authenticity of digital materials... cannot be defined in any monolithic sense'⁴⁵⁹ and that '...the methods for assessing and ensuring record trustworthiness are more accurately characterised as a hedge against uncertainty concerning our knowledge about past events, than as a guarantor of certainty about it'.⁴⁶⁰ Trustworthiness is always necessarily

⁴⁵⁸ MacNeil, 'Trust and professional identity', p.115.

⁴⁵⁹ MacNeil & Mak, 'Constructions of Authenticity', p.17.

⁴⁶⁰ MacNeil, *Trusting Records*, p.113.

defined according to considerations that are personal, professional, cultural and systemic, according to the needs of the actors, processes and frameworks where the interrogation of information is taking place. The broadly conceived controls for authenticity here derived from archival-diplomatics are hedges, but importantly, they are derived from the same tradition of legal and archival thinking as the record-as-command that animates the relations between the state, juridical persons and rights. However contingent, some assurance of authenticity is possible when information identity and integrity are protected through metadata, systems and documented custodianship, all of which must be transparent. Currently, the focus of civil society monitoring efforts is open government data, and as established in Chapter 1, data authenticity has not been closely considered by the open government data community. The next chapter will look at how controls in the three broad categories of metadata, auditable systems and documented custodianship have been used to support assumptions of authenticity in information, before these controls are applied to open government data in the subsequent case studies.

Chapter 6

Practices for Establishing and Maintaining Authenticity in Information

Recorders that record thisnesses become thisnesses themselves. That makes every instance of archiving into a discursive event.⁴⁶¹

Friedrich Kittler, *Discourse Networks 1800/1900*

⁴⁶¹ Kittler, F., *Discourse Networks 1800/1900* (Stanford: Stanford University Press, 1990) p.339.

6.1. Introduction

In Commonwealth countries such as those featured in the following case studies, government record-keeping is often rooted in British administrative traditions. Goody has suggested that record-keeping was central to the imperial project: ‘...bureaucracy took a major leap under British imperial rule, which Smith has described in terms of complementary aspects of rule-by-records and rule-by-reports’.⁴⁶² With the changes to the British public sector that followed the Fulton Report of 1968 – coinciding with the end of the colonial period – current government practices in some Commonwealth countries more closely resemble pre-Fulton British practices than do current British practices.⁴⁶³ These administrative traditions emerge from the same culture as the archival theory outlined in the previous chapter, a theory that has drawn on common law legal theory and, more recently, diplomatics to identify metadata for identity and integrity, systems of control and custodianship (continuous control and ethical action) as core components of establishing and maintaining the authenticity of records. Record-keeping standards and guidance have evolved to help to ensure authenticity in records as the basis for decision-making and accountability. This chapter will demonstrate how these elements were manifested in the management of paperwork and how they have been transposed to the digital environment. To do this, it will examine examples of guidance on record-keeping; one example of guidance on managing paper records (The Bengal Records Manual 1943) and one of managing hybrid digital / paper records (ICA-Req).⁴⁶⁴ The chapter includes sometimes substantial quotes from the two sources. These are included to give a sense of the detail and prescriptiveness of the documents, which is necessary to understand the systems of control that they create. As discussed later in this thesis, it is at a similarly granular level that controls are needed in open government data curation practices if the data is to be trustworthy. The Bengal Manual and ICA-Req both provide evidence that the

⁴⁶² Goody, J., *The Logic of Writing and the Organization of Society*, (Cambridge: Cambridge University Press, 1986) p.115.

⁴⁶³ Subramaniam, V., *Transplanted Indo-British Administration*, (New Delhi: Ashish House Publishing, 1977) p.26.

⁴⁶⁴ International Council on Archives, *ICA-Req: Principles and functional requirements for records in electronic office environments*, (Paris: ICA, 2008).

record-keeping profession has recognised the significance of metadata, systems and custodians for authenticity and developed ways to implement them.

6.2. The Bengal Records Manual 1943

In 1944, the Government of Bengal Board of Revenue issued The Bengal Records Manual 1943. I came into possession of a copy of the Manual by chance. I will use it here as an example of guidance on government record-keeping because it was created in a time and place where government record-keeping was heavily prescribed.

Timothy Lovering, Alistair Tough and Paul Lihoma, Abiola Abioye, and Charles Magaya and I have shown that British colonial record-keeping was not systematically controlled from the metropole, with the effect that while sharing common bureaucratic terminology, concepts and structures, the colonies were sites of varied administrative practice.⁴⁶⁵ A common feature was the organisation of registries for controlling the registration, circulation and managed storage of records, a system first instituted by the Foreign Office in 1810.⁴⁶⁶ These offices would accept incoming correspondence, register it to a file and circulate the file for action. They frequently also served as records stores, or they would separate the record store from the larger office. Attempts, in the twentieth-century, both to standardise practices across the Empire and to reform colonial registries to improve the efficiency of colonial governments, did not achieve uniformity but did introduce practices and tools, such as classification schemes, such that it is

⁴⁶⁵ Lovering, T., 'British Colonial Administrations' registry systems: a comparative study of Northern Rhodesia and Nyasaland', *Archival Science* 10.1 (2010) pp.1-23; Tough, A. & Lihoma, P., 'The development of recordkeeping systems in the British Empire and Commonwealth, 1870s–1960s', *Archives and Manuscripts*, 40.3 (2012); Abioye, A., 'Milestones in archives administration in Nigeria', *Comma*, 2012.1 (2013) pp.15-26; and Magaya, C. & Lowry, J., 'The National Archives of Tanzania Fifty Years After Wright's Report on the Government Records and Public Archives of Tanganyika', *Comma: Journal of the International Council on Archives*, 2012.1 (2013). It could be argued that the ad hoc character of colonial record-keeping was a reflection of the ad hoc character of colonial governance, following my argument in Chapter 1 that government records are the material embodiment of government authority and Jeffrey Herbst's observation that 'so much of "colonial science" was made up in the face of particular exigencies and often by the man on the spot rather than in the colonial capital, much less in Europe' (Herbst, J., *States and Power in Africa: Comparative Lessons in Authority and Control* (Princeton: Princeton University Press, 2000) p.82).

⁴⁶⁶ Vincent, D., *The Culture of Secrecy: Britain, 1832-1998* (Oxford: Oxford University Press, 1998) p.32.

possible, today, to find traces of the same practices and tools in the archives of countries in the Caribbean, sub-Saharan Africa, the Pacific, Australasia, and the Indian subcontinent. Indeed, some such practices and tools are still employed in government registries in these regions, as I found in the judiciaries of Kenya, Uganda and Tanzania during earlier research.⁴⁶⁷

Tough and Lihoma's discussion of colonial record-keeping practices excludes discussion of India, noting 'it is a subject in itself' and while administrative histories of each colony would each have a unique character, the case of the countries of the Indian subcontinent is made markedly different by the fact that colonisation was earlier, larger in territory and its government more complex than most other British colonies.⁴⁶⁸ As V. Subramaniam has observed: 'Governmental administration in India is essentially a creation of the Indo-British contact of two centuries and its main characteristics and problems took shape during this period'.⁴⁶⁹ Lovering and Tough and Lihoma note that record-keeping practices were not transplanted wholesale from the metropole to sub-Saharan African colonies, but Hull has argued that in the territory now constituting Pakistan, government record-keeping evolved from the highly developed systems of the British East India Company, which were designed to provide management control of trade in India from London. Hull wrote that the '...practices of written accountability designed for the management of far-flung and unreliable commercial agents were carried over into the operations of territorial rule as the Company gradually assumed the form of the colonial government of India'.⁴⁷⁰ He describes the significance of written records for the Company, stating that '...only through a connection with a piece of paper (a bill, warrant, note, book, and so forth) could an action be construed as an action of the Company'.⁴⁷¹ Given the British East India Company's management of British India, Hull's argument – that the Company's practices provide the foundation for government practices in Pakistan - might be equally valid for other parts of the Indian subcontinent. His

⁴⁶⁷ Lowry, J., 'Correlations between ICT and records policy integration and court case management system functionality', *Records Management Journal*, 23.1 (2013) pp.51-60.

⁴⁶⁸ Tough & Lihoma, 'The development of Recordkeeping Systems', p.191.

⁴⁶⁹ Subramaniam, *Transplanted Indo-British Administration*, p.126.

⁴⁷⁰ Hull, M. S., *Government of Paper: The Materiality of Bureaucracy in Urban Pakistan*, (Berkeley; Los Angeles; London: University of California Press) p.7.

⁴⁷¹ *Ibid.*, p.8.

observations of the city administration of Islamabad certainly describe a bureaucratic machinery of the sort that is intimated in the Bengal Records Manual 1943.

In 1943, Bengal was a province of British India ruled by the Government of Bengal.⁴⁷² The preamble to the Manual notes that the Manual's instructions apply primarily to 'the offices of Collectors and their Subordinates, but they should be followed as far as possible, in all offices subordinate to Government'.⁴⁷³ The Manual reproduces the Destruction of Records Act, 1917 (Act V of 1917) and the associated Statutory Rules, before Chapter I sets out procedures for the management of 'English records'. Chapter II sets out procedures for the management of 'Vernacular records'. Subsequent chapters deal with subdivisional records, birth and death registers, rules regarding record-room plans, rules relating to the destruction of various classes of records, record grant and record-room receipts, the management of official libraries, and rules for the supply of copies or information about records. Sample forms and registers are provided as appendices. I will restrict my study to Chapter I 'English records', since it is the most extensive and comprehensive guidance provided in the Manual and should therefore suffice to demonstrate the presence of requirements for metadata, systems and custodians. The Manual is silent on whether 'English records' refers to records emanating from Britain or all records written in the English language. A note at the head of Chapter II 'Vernacular records' states that vernacular records include all case records whether in English or the vernacular, which may suggest that Chapter I concerns records in the English language with the exception of case records.

The Manual is a late specimen of a genre that Hull asserts flourished in the late nineteenth century, stating that '[n]ormative procedures were laid down in hundreds of manuals produced for every sphere of administration in the late nineteenth century'.⁴⁷⁴ He claims that record-keeping procedures in Pakistan

⁴⁷² Travers, R., *Ideology and Empire in Eighteenth-Century India: the British in Bengal* (Cambridge: Cambridge University Press, 2007).

⁴⁷³ Board of Revenue, Government of Bengal, *The Bengal Records Manual 1943*, (Alipore: Bengal Government Press, 1943) p.3 (hereafter, *The Bengal Manual*).

⁴⁷⁴ Hull, *Government of Paper*, p.10.

have changed little over the last hundred years,⁴⁷⁵ noting that ‘...though they are part of new projects and repurposed in novel ways, many of the bureaucratic inscriptional practices from the colonial period have remained vital in the contemporary period’.⁴⁷⁶ This may also have been the case in Bengal, and the Manual is still the primary source of records management guidance for the government of Bangladesh.⁴⁷⁷ In the Manual’s preface, P.R. Das Gupta, Secretary to the Board of Revenue, outlines the genealogy of rules and laws enacted in Bengal for the management of records beginning with the ‘Records Rules’ of 1854. Gupta situates the 1943 manual as the latest iteration of a manual first published in 1888 and revised in 1902, 1911, 1917 and 1928, and notes that the 1943 manual is ‘practically a reprint’ of the 1928 manual with corrections. Though the Manual is an example of the highly developed administrative procedures that have their roots in the late nineteenth century, as I will argue, many of the controls set out in the Manual manifest in the modern standard for digital records management system functionality, ICA-Req.

6.3. ICA-REQ

With the introduction of computers into office work, the methods of record-making changed. Michael Moss has argued that computerisation and access to the internet resulted in the ‘disappearance of secretaries (who typed letters), managers now typing their own communications (largely emails), the closure of registries where files were maintained and stored, the closure of libraries and the takeover of information systems by computer scientists’.⁴⁷⁸ This was the decentralisation of the control of corporate records, a move from creation by secretaries and storage and circulation by registries, to creation by all officers, storage on their personal computers or shared network drives and circulation by email. There arose a need to find ways to manage these dispersed digital records

⁴⁷⁵ Ibid., p.118.

⁴⁷⁶ Ibid., p.11.

⁴⁷⁷ While teaching at the University of Dhaka and East West University in Bangladesh in October 2018, I met with staff of the National Archives, the Information Commissioner’s Office and several government departments, who stated that the 1943 Manual was still in use across the Bangladesh public service.

⁴⁷⁸ Moss, M., ‘What is the same and what is different’, in Moss, M., Endicott-Popovsky, B. & Depuis, M. J. (eds.), *Is Digital Different?* (London: Facet Publishing, 2015) p.6.

so that they could be found quickly, shared appropriately, and relied upon as authentic evidence for decision-making and accountability, which led to the development of the Electronic Document and Records Management System (EDRMS). These systems sought to replicate registry practices in the digital environment and re-centralise control of records. James Lappin has described the emergence of what he calls the 'EDRMS model' in the decade beginning in 2000 and its decline by 2010, noting that, at 2010:

The various electronic records management system statement of requirements has [*sic*] lost their market power. One product (Microsoft's SharePoint) has captured a huge percentage of the market that these standards were supposed to govern, without being certified against those standards.

The EDRMS approach has been found not to be a universal panacea. Some organizations made it work, but others failed to complete costly and complex EDRMS projects. The pipeline of new EDRMS projects has slowed to a trickle.⁴⁷⁹

Five years later, Moss was more bold, stating that 'EDRMS have been a failure'.⁴⁸⁰ Asking 'What will be the next records management orthodoxy [following the decline of the EDRMS model]?', Lappin questions whether it might be the 'records repository model', which involves a 'business classification housed in a back office repository, but used to govern content held in the various native applications of the organization'.⁴⁸¹ Although the records repository model has not risen to the prominence attained by the EDRMS model (Lappin rightly anticipated that it would not), one of its central concepts – engagement with native applications – has characterised other initiatives from the record-keeping profession aimed at enabling control and management of digital records. The most notable of these initiatives is the development of an international standard (ICA-Req) setting out functional requirements for digital records management that

⁴⁷⁹ Lappin, J., 'What will be the next records management orthodoxy', *Records Management Journal*, 20.3 (2010) p.253.

⁴⁸⁰ Moss, 'What is the same and what is different', p.6.

⁴⁸¹ Lappin, 'What will be the next records management orthodoxy', p.260.

contemplates the incorporation of the records management functionality into native applications (in Module 3 of the standard). Jean Dryden discussed ICA-Req in her column on standards in the *Journal of Archival Organization*, and claimed that its authors

broke new ground when they made an important distinction between ERM systems and business systems. ERM systems include record-keeping functionality... Business systems, on the other hand, are designed to conduct business, and rarely include record-keeping functionality...⁴⁸²

Standards for functional requirements for digital records management have a long history: the first version of Norway's NOARK standard appeared in 1984. Other national standards have been developed by Australia, Brazil, Canada, Finland, Germany, Malaysia, The Netherlands, New Zealand, Romania, Russia, South Africa, Switzerland and the United Kingdom.⁴⁸³ The first English language standard was Canada's RDIM (Records/Document/Information Management (RDIM): Integrated Document Management System for the Government of Canada), which began life as the software requirements accompanying a request for proposals from software vendors to supply the government with digital systems for records management, but it was then issued by the National Archives of Canada (now Library and Archives of Canada) as a standard in 1996.⁴⁸⁴

Emerging out of a collaboration between the UBC MAS Project and the US Department of Defense, DoD 5015.2-STD 'Design Criteria Standards for Electronic Records Management Software Applications' was published in 1997 and last revised and reissued in 2007 as DoD 5015.02-STD 'Electronic Records Management Software Applications Design Criteria Standard'. DoD 5015.2-STD became a touchstone for records managers across the English-speaking world, partly because of the awareness raised in the professional community by the

⁴⁸² Dryden, J., 'Standards: News, Progress Reports and Reviews', *Journal of Archival Organization*, 8.3 (2010) p.261.

⁴⁸³ DLM Forum, *Other Specifications*, <http://moreq2.eu/other-specifications> [accessed 22 August 2018].

⁴⁸⁴ *Ibid.*

connection to the University of British Columbia (UBC) programme and partly because software vendors were required to have their products certified against the standard in order to sell to the US government, which led to companies marketing their products as DoD 5015.2-STD certified.⁴⁸⁵

In 2001, the 'Model Requirements for the Management of Electronic Records' (MoReq) was published.⁴⁸⁶ Developed by the DLM Forum – a European network of government archives and information professionals - with funding from the European Commission, the standard was intended to be applicable in different countries and industries, and it has been widely used and has informed the development of similar national standards, particularly in Europe.⁴⁸⁷ The UK's Public Record Office (now The National Archives) issued its PRO2002 Specification in 2002.⁴⁸⁸

Given that these standards often addressed specific jurisdictional needs, and that many jurisdictions had not yet developed their own standards, the International Council on Archives (ICA) undertook a project to produce a generic set of requirements.⁴⁸⁹ It is not clear how the generic nature of MoReq figured in the ICA's decision to develop new requirements, though ICA-Req refers users to DoD 5015.2 and MoReq2 for more detailed requirements.⁴⁹⁰ Nevertheless, in 2008, the ICA published its *Principles and Functional Requirements for Records in Electronic Environments* (known as ICA-Req), which was developed by an international project team led by the National Archives of Scotland and coordinated by the National Archives of Australia, with a secretariat provided by Archives New Zealand.⁴⁹¹ The project team, which consisted of members from the Cayman Islands, France, Germany, Malaysia, The Netherlands, South Africa,

⁴⁸⁵ Laserfiche, *Why You Need to Care About DoD 5015.2*, <https://www.laserfiche.com/ecmblog/why-you-need-to-care-about-dod-5015-2/> [last accessed 15 March 2019].

⁴⁸⁶ DLM Forum <http://moreq.info> [last accessed 15 March 2019]. The current version of MoReq (MoReq2010) was published in 2011.

⁴⁸⁷ Ibid. The Russian standard is partly based on MoReq and the Romanian standard is a translation of MoReq2.

⁴⁸⁸ The National Archives, *PRO2002*

<http://www.nationalarchives.gov.uk/documents/referencefinal.pdf> [last accessed 15 March 2015].

⁴⁸⁹ International Council on Archives, *ICA-Req: Principles and functional requirements for records in electronic office environments*, (Paris: ICA, 2008).

⁴⁹⁰ Ibid.

⁴⁹¹ Ibid.

Sweden, the United Kingdom and the United States, reviewed existing guidance and functional requirements to produce the three modules that constitute ICA-Req:

- Module 1: Overview and Statement of Principles: background information, organisation, fundamental principles and additional context;
- Module 2: Guidelines and Functional Requirements for Records in Electronic Offices: a global high-level statement of core and optional requirements, including application guidelines and a compliance checklist; and
- Module 3: Guidelines and Functional Requirements for Records in business Systems: guidelines and generic core and optional functional requirements for records in business systems.

In 2010, ICA-Req was adopted by the International Standards Organisation (ISO) as ISO16175.⁴⁹² Since it is the ISO approved standard for functional requirements for digital records systems, I will use ICA-Req Module 2 as an example for comparison with the Bengal Records Manual. I intend to show how controls in traditional record-keeping systems, such as the one elaborated by the Bengal Records Manual, manifest in guidance for digital record-keeping systems. I will refer to the standard as ICA-Req, rather than ISO16175, because I am making use of the version published by the ICA, which may differ from the ISO version.

Systems, whether EDRMS, 'repository model' systems, or native applications that incorporate ICA-Req style functionality, are significant in that they provide the means by which metadata can be captured and linked to records, and because they enable the documentation of custodianship. In discussing Canada's Uniform Electronic Evidence Act, MacNeil has cited Section 5 (a) as showing that the integrity of electronic records is presumed 'by evidence that supports a finding that at all times the computer system or other similar device was operating properly or, if it was not, the fact of its not operating properly did not affect the

⁴⁹² International Council on Archives, *Principles and Functional Requirements for Records in Electronic Office Environments: Overview of Implementation Guidance and Training Products*, (Paris: ICA, 2013) p.5.

integrity of the electronic record, and there are no other reasonable grounds to doubt the integrity of the electronic records system'.⁴⁹³ I intend to show that analogue systems, such as the one described by the Bengal Records Manual, and digital records management systems, such as ICA-Req compliant systems, work with metadata and custodianship to create a 'trusted environment', in which users can be reasonably assured that the authenticity of records has been preserved, noting that, as MacNeil has argued, '[t]he authenticity of digital materials... cannot be defined in any monolithic sense'.⁴⁹⁴

6.4. Controls for Authenticity: The Bengal Manual and ICA-Req in Comparison

It is impractical to divide the requirements of the Manual and ICA-Req between metadata, systems and custodians, since any one of the requirements might involve any of these three in combination. Instead, this chapter will proceed through Chapter I of the Manual to identify requirements that relate to any of the three categories and then look to ICA-Req for their digital equivalents. These findings are arranged below according to the record-keeping activities to which they relate, for instance, 'registration' or 'foliation'. By this process, the chapter shows that the three elements that I have distilled from the UBC-MAS / InterPARES work can be operationalised to establish and maintain authenticity in information.

6.4.1. Registration

The InterPARES Terminology Database does not define registration, but defines 'registered record' as a 'record that has been assigned a registration number and for which all the data necessary to identify the persons and acts involved and the documentary context of the record are recorded within a protocol register'.⁴⁹⁵ Registration, then, is the practice of recording data about a record in a register

⁴⁹³ MacNeil, H., *Trusting Records: Legal, Historical and Diplomatic Perspectives* (Dordrecht; London: Kluwer Academic, 2000) p.53.

⁴⁹⁴ MacNeil, H. & Mak, B., 'Constructions of Authenticity', *Library Trends* 56.1 (2007) p.17.

⁴⁹⁵ InterPARES 2, *Terminology Database*, http://www.interpares.org/ip2/ip2_terminology_db.cfm [last accessed 15 March 2019].

and linking the record and the register through the use of a number. Registration is the primary method of inducting the record into the greater system that will support its trustworthiness, as defined in the InterPARES Terminology Database. It is through registration that the first trace of the record appears in the system, in the form of metadata. The Bengal Manual states that:

1.A.8. The smooth working of the system will depend much on the correct and careful registration of the letters received and issued.-

Three annual registers will be maintained, viz. :-

I.-Register of letters received. (No. 60.) Form No. 9.

II.-Register of letters issued. (No. 61.)-Form No. 10.

III-Index Register of English correspondence. (No. 62.) Form No. 11.⁴⁹⁶

These three registers are vital control mechanisms in that they provide evidence of the receipt or creation of records. They can be used as the starting point in tracing records following queries or complaints about a given matter and provide a trace of the existence of records in the event that a record has been removed from the system or destroyed. The registers also capture metadata that provides a point of comparison with the records themselves in the event that records need to be examined for signs of tampering. Finally, the metadata in the registers relate records to files, which effectively assigns them their place within the record-keeping system and begins to build contextual relations around the record. The more elaborate the processes of registration, the more opportunity there is to add metadata, creating a more complex contextual web of relations between folios, files and registers. I will return to the functions of the specific metadata elements set out in the registers, but first it is helpful to understand something of the process of handling incoming letters. The Manual states:

1.B.13. All correspondence received will be opened by the Collector or some other responsible officer to be deputed by him for the purpose.

When the letter or telegram is opened the officer opening it should

⁴⁹⁶ *The Bengal Manual*, p.5.

mark on it by stamp or in writing on the first page the date of receipt...⁴⁹⁷

In this stipulation, we see responsibility for custodianship during this activity assigned to the Collector or his deputy and the first assignation of metadata to the record, in the form of the date of receipt. In ICA-Req, the same function is envisaged in requirement 15: systems must '[a]utomatically capture the date and time of capture of each record as metadata elements linked to each record'.⁴⁹⁸

In the Bengal Manual, custodianship then passed to the head assistant or his deputy and further metadata was added:

1.B.13... When the document is received in the office the head assistant or an assistant specially deputed by him will see that the document bears the two docket stamps mentioned below. If not, he will affix the two stamps, if possible, on the left and right-hand top corners at the first page of the document as shown below. The head assistant will then fill in the spaces provided for the department and class of papers to which the document belongs and note in pencil for the guidance of the diarist the collection under which it falls. The head assistant is responsible for the correct classification of the papers...⁴⁹⁹

The Manual attempts to standardise the placement of the two docket stamps, which provides a visual sign of conformance with system controls and thereby, an indication that the record has 1) been accepted into the system as a record and 2) been handled according to the procedures of the system, which, according to the archival-diplomatic view, is suggestive of an authentic nature. The process continues with the act of registration itself:

1.B.13... The letter will then be made over to the diarist who will enter it

⁴⁹⁷ Ibid., p.6.

⁴⁹⁸ International Council on Archives, *Principles and Functional Requirements for Records in Electronic Environments - Module 2: Guidelines and Functional Requirements for Electronic Records Management Systems* (Paris: ICA, 2008) p.25 (hereafter, *ICA-Req 2*).

⁴⁹⁹ *The Bengal Manual*, p.6.

in the “Register of letters received”, columns 1 to 4 and 6 being at once filled in, but column 7 being left blank till the letter is disposed of. The date of receipt entered in this register will be the date marked by the opening officer and not the date of receipt in the office. The necessary entries will then be made in the space provided in the docket stamp.⁵⁰⁰

The two docket stamps mentioned above perform an important function in that they set out the mandatory metadata elements and, importantly, serve as a template for metadata capture, so that it is easy to see if any metadata has not been captured, which would indicate the improper functioning of the system. The stamps themselves indicate the relationships that are established through metadata:

1.B.15. The following forms of docket stamps, which may be obtained from the Controller of Stationery, should be used in cases where the printed forms prescribed by the Board have not been used:-

Left-hand docket stamp.

Enclosures.
Maps or plans.
Spare copies.
Class of papers.
Reply issued: No. , date .

Figure 2. Left-hand docket stamp, the Bengal Manual, 1948.

The left-hand docket stamp calls for the recording of enclosures, which performs two functions: 1) it relates the primary folio to all accompanying folios, making clear their relationship, which helps readers to understand how they should be read, and 2) it acts as an additional trace of the presence of the enclosures – in

⁵⁰⁰ *The Bengal Manual*, p.6.

other words, provides information that helps the reader establish whether all folios are present on the file, which helps to establish both the identity and integrity of the record. This is the same function foliation performs for records registered to files, as I will explore below. In ICA-Req compliant systems, these links are established because the system must:

30 – Capture compound electronic records (records comprising more than one component) so that:

- the relationship between the constituent components of each compound record is retained;
- the structural integrity of each compound record is retained; and
- each compound record is retrieved, displayed and managed as a single unit.⁵⁰¹

The left-hand docket stamp also requires an indication of accompanying maps or plans, which alerts the reader of the need to refer to the system of management for outsize documents of these types, thereby creating a link between systems and elaborating the contextual web of relations around the folio. In asking for an indication of spare copies, the stamp helps the reader to know whether they might expect to find other instances of the record in other files. ‘Class of papers’ refers to classification of records into one of three classes that inform their retention and destruction (Class A papers (permanent retention), Class B papers (twelve year retention) and Class C papers (two year retention)). The field for ‘Reply Issued’, with space for the number of the reply and its date, creates a link between the record and the related reply, and both the incoming and outgoing letters registers.

These connections created by metadata are essential controls because they leave traces that are auditable, or as Hull posits, ‘[t]hrough the taciturn medium of stamps, clerks create a metatext that emplots the official career of the artifact in time, space, organizational order, and the order of other graphic artifacts’.⁵⁰²

Following the loss of a record or the identification of anomalies in its appearance,

⁵⁰¹ *ICA-Req 2*, p.28.

⁵⁰² Hull, *Government of Paper*, p.121.

the existence of the record and/or the details of its creation and receipt will be attested to by the registers and the incoming or outgoing record to which it relates. Importantly, the correct attribution of metadata is assigned to particular officers, which creates an incentive for officers to properly give effect to the system. Furthermore, responsibility for custody is assigned at each step, establishing the kind of continuous control envisaged by Jenkinson.

1.B.15... Right-hand docket stamp.

Diary or Register No,	
Department.	
Branch.	
Collection No.	
No. and year of File.	
Serial No. in File.	
No. and date of orders issued. No. reply.	

Figure 3. Right-hand docket stamp, the Bengal Manual, 1948.

The right-hand docket stamp is concerned with situating the record within the system by linking the record to the diary or register in which it is recorded, the sections of the organisation that are assigned to deal with it, the collection in which it will be kept, the file to which it relates and in which it should be found, its serial number in the file (see *Foliation*, below) and any orders or replies connected to it. The significance of these links is that they create relationships that enable tracing and auditing, which offers some protection to the record and underpins a presumption of authenticity in records kept within the system. The process of registration continues by relating the folio to the wider record-keeping system via the 'Register of letters received'. The diarist was expected to populate

columns 1 to 4 and 6 of the register, which are:

1. Consecutive number.
2. Date of receipt in office.
3. Letters, etc., subdivided into –
 - (a) From whom received.
 - (b) Number.
 - (c) Date.
4. Short subject.

6. Where the letter is placed, subdivided into –
 - (a) File number.
 - (b) Serial number.
 - (c) Number and title of the collection.
 - (d) Number of the file within the collection.⁵⁰³

Column 1 asks for a consecutive number, which underpins the integrity of the register as a record of incoming correspondence. The progress of the sequence of numbers down the column is a visual clue to the unaltered condition of the register. The date of receipt in the office required by Column 2, in combination with the date of receipt noted on the letter by the Collector or his deputy at receipt, provides a measure of the duration of the registration process, which is information that may be used in auditing the custodianship of the letter and the functioning of the system. Columns 3 and 4 capture the metadata of the record itself, which is the central activity and purpose of registration and provides the primary trace of the presence of the record. The information required in Column 6 is key to placing the record within the larger system, but goes beyond that to support the integrity of the system itself by providing not only a file number, but a secondary number - the number of the file within the collection (a sequential number recorded in Index Register No.62, discussed at *File Registration*, below). Such a number relates the file to its position within the collection, allowing users to see from an inspection of the surrounding files whether or not a file is missing.

⁵⁰³ *The Bengal Manual*, p.102.

The presence of this number throughout the registers, whenever there appears a record related to the matter the file pertains to, is evidence of the existence of the file beyond any file lists or catalogues that may exist in the registry. In other words, there is a second layer of traces of the existence of the file: even if the catalogue is doctored or destroyed to cover the removal of a file, the file's existence is discernible from the registers. Column 5 relates to reminders that may be issued to officers throughout the use of the record, which hints at the management function of the registry and the significance of written records for the performance of the functions of the organisation, and Column 7 calls for a 'Brief note of action' after the letter is disposed of.

Section 3.1.2 of ICA-Req provides functional requirements for 'point of capture metadata', including the requirements to:

11 – Acquire metadata elements for each record and persistently link them to the record over time.

18 – Allow entry of additional metadata by users during record capture and/or a later stage of processing by the user.⁵⁰⁴

The Manual's guidance on the attribution of each record to a file in column 6 of the register is replicated in ICA-Req: '21 – Ensure that all records captured within the electronic records management system are associated with at least one aggregation'.⁵⁰⁵

The registration of outgoing letters according to the rules of the Bengal Manual similarly created metadata: '1.C.24. Enter the issuing letter in Register No. 61 of letters issued, filling in columns 1 to 6 and entering the remainder, if required, in columns 8 and 9'.⁵⁰⁶ Again, two docket stamps were used to collect metadata. These were used on final drafts, which were retained on file⁵⁰⁷, and on 'fair

⁵⁰⁴ *ICA-Req 2*, pp.25-26.

⁵⁰⁵ *Ibid.*, p.26.

⁵⁰⁶ *The Bengal Manual*, p.9.

⁵⁰⁷ Rule 1.C.25 states: 'Enter the register number and the date and give the number of the collection and file to which the letter belongs, on the draft which is to remain in the office file'.

copies' for transmission. In the latter case, certain fields were left blank for use in the receiving office. I will not analyse the instructions for handling outgoing letters or the elements of the relevant registers or stamps since they perform the same functions as those for incoming letters – they create and capture metadata, introduce the record into the larger system of management, and assign custodial responsibility throughout the process of preparation, registration and dispatch.

In ICA-Req, there is a requirement that systems must: '35 – Ensure the capture of email transmission data as metadata persistently linked to the email record'.⁵⁰⁸ In the Bengal system, similar metadata elements are captured in the register and in the fields of the right-hand docket stamp on the letter, creating a link between the folio and the register, or the record and the greater system for its control. These stamps and registers are concerned with the creation of metadata, a concern that has persisted into present day records management standards. Whereas the Bengal Records Manual provides the metadata elements to be populated, in the fields of the sample forms provided in its appendices, ICA-Req refers users to the ISO metadata standard.⁵⁰⁹

6.4.2. File Registration

Just as records must be registered so that their presence in the system is documented, files must be registered. In the Bengal Manual, file registration is done via Index Register No. 62, a simple register of four columns for recording 1) number and title of collection (the highest level of organisation), 2) number of the file within the collection (the sequential number for the file, recorded in column 6 of the 'Register of letters received' as discussed above), 3) the file subject, and 4) remarks.⁵¹⁰

⁵⁰⁸ *ICA-Req 2*, p.29.

⁵⁰⁹ ICA-Req states: 'The reference metadata standard for these requirements is ISO 23081 – 1: 2006, Information and Documentation – Records Management Processes – Metadata for Records, Part 1 – Principles. The high-level metadata element set found in ISO/TS 23081 – 2: 2007, Information and Documentation – Records Management Processes – Metadata for Records, Part 2 – Conceptual and Implementation Issues provides the basis for the requirements.', *ICA-Req 2*, p.6.

⁵¹⁰ *The Bengal Manual*, p.104.

New Index Registers were started at the beginning of each year. Files related to unfinished business were carried forward into the new year by re-registering them in the new year's Index Register, which would require the assignation of a new number. This new number was noted in the 'remarks' column of the previous year's Index Register, and the old number noted in the same column of the new register. Rule 70 explicitly states that this is '[f]or the purpose of tracing them in the future'. By this method, the Index Register served as list of all current files. It also served as a guide to the location of the files within storage, since files were arranged on shelves according to collection and then number, mirroring their arrangement in the register.

Concerns for the registration of information about files are also present in ICA-Req. The registration of new files (or aggregations or volumes) takes the form of these three requirements:

61 – Record the date of opening of a new aggregation within its associated records management metadata.

83 – Record the date of opening of a new volume in the volume's records management metadata.

73 – Record the date of closing of a volume in the volume's records management metadata.⁵¹¹

The Index Register's function as a list of all open records is performed in ICA-Req-2 compliant systems through reporting functionality. The Manual states that new files should be opened for new matters arising, but omits guidance on when a file is considered 'closed' in the sense that no more records should be registered to it. However, it requires that, at the end of each year, files 'which it is no longer necessary to keep' should be appraised and either destroyed (according to exhaustive rules for each class of records) or transferred to the record room at the District headquarters.⁵¹² The opening or closing of files is

⁵¹¹ *ICA-Req 2*, pp.31-33.

⁵¹² *The Bengal Manual*, p.14.

mirrored in ICA-Req compliant systems, which must:

79 – Be able to close a volume of an electronic aggregation automatically on fulfilment of specified criteria to be defined at configuration, including at least:

- volumes delineated by an annual cut-off date (for example, end of the calendar year, financial year or other defined annual cycle);
- the passage of time since a specified event (for example, the most recent addition of an electronic record to that volume); and
- the number of electronic records within a volume.

85 – Support the concept of open and closed volumes for electronic aggregations, as follows:

- only the most recently created volume within an aggregation can be open; and
- all other volumes within that aggregation must be closed (subject to temporary exceptions required by Requirement 68).

86 – Prevent the user from adding electronic records to a closed volume (subject to the exceptions required by Requirement 68).⁵¹³

The measures for opening and closing files, both in the Manual and in ICA-Req, help to ensure that all records relating to a particular matter are kept together, building a history of the subject of the file that provides context and enables review in support of accountability and informed decision-making.

6.4.3. Forward Diaries

A particular form of register is constituted by what the Bengal Records Manual calls a ‘Forward Diary’, and which in other instances of the registry system is called a ‘Bring Up’ or ‘Bring Forward’ diary.⁵¹⁴ The diary ‘does not supersede any

⁵¹³ *ICA-Req 2*, pp.33-34.

⁵¹⁴ For example, the *Manual of Registry Practices and Procedures*, May 1989, states ‘52. Markings [of “B.U.” (for “Bring Up”)] on folios by officers requiring access to them at a future date] should be noted by the officer-in-charge of the branch registry in his diary, and the movement

existing prescribed register' (1.E.46, p.12), but alerts clerks of the need to retrieve files from the registry and circulate them to officers for review or further action, usually on a date specified by the officer at the return of the file to the registry. In the Bengal Records Manual, the prescribed format of the Forward Diary is given at 1.E.47.(a):

The dates are printed on the edge of the diary. For each date there are 12 compartments, facing the date. Ordinarily the top left-hand compartment on the page would be for January, the next for February, the next for March, and so on. For each day's entries one of the 12 compartments should be used...⁵¹⁵

Registers of incoming and outgoing correspondence were to be kept permanently but the Manual is silent on the fate of Forward Diaries. However, while they exist, Forward Diaries constitute another trace of files and records: '1.E.47.(b) In the space below the name of the month, the numbers of the cases, files or letters required to be put up on the date against the page should be noted'.⁵¹⁶

Although there is no direct correlation between Forward Diaries and the requirements of ICA-Req because access does not require (though may involve) the movement of records in the digital environment, the appearance of files and records in the diaries constitutes metadata of the sort accumulated in ICA-Req compliant systems when officers access or take action on records, such as that discussed under *Tracking*, below.

6.4.4. Naming and Numbering Conventions

ladder [on the file cover] entry initialed by him...'. This manual does not indicate in which country it was published, but notes that it was published by the 'O.H.C.S.', which suggests that, like the Bengal Manual, it is from one of the Commonwealth countries, where O.H.C.S. usually stands for Office of the Head of Civil Service. The year of publication together with the discussion, in the introduction, of the importance of records management for the Civil Service Reform, suggest that it was published in sub-Saharan Africa. Civil service reform programmes occurred across that region in the 1980s and 1990s. Another use of the term 'Bring Up Diary' occurs in the 'Procedures for Bring Up Action' in the *Registry Procedures Manual* published by Uganda's Ministry of Public Service and Cabinet Affairs in 1988.

⁵¹⁵ *The Bengal Manual*, p.12.

⁵¹⁶ *Ibid.*

In the Bengal Manual's description of registering outgoing letters, we see an example of the coding used to identify individual records:

1.C.28. In the fair copy [of the reply], fill in the number of the letter, *e.g.*, No. 707 R.S./XII-3-19 (which means letter No. 707 in register No. 61 of the Settlement branch of the Revenue Department, Collection No. XII, file No. 3, serial No. 19).⁵¹⁷

ICA-Req's requirement 20 requires systems to automatically allocate unique identifiers to records at the point of capture and requirement 38 requires systems to associate every record with a unique identifier.

A marked difference between the analogue and digital appears in the need to name digital records. While metadata about a paper record (for instance, its author, recipient or subject) might be recorded in a register, it is at no point given a name or title in the register, though the record itself may bear a title, as reports usually do, for instance. Digital records must have a name or title in the form of a filename. Filenames are covered in ICA-Req:

5 – Support manual naming of electronic records, and allow this name to be different from the existing file name (including email subject lines used to construct record titles). If the existing filename is taken by default, the electronic records management system must allow this name to be amended at the time of capture.⁵¹⁸

As for titling and numbering files, or what ICA-Req calls volumes or aggregations, the Manual only asks for a *'suitable* file "subject or "title" (1.B.16, p.7) and the assignation of the next consecutive number in the relevant collection in the Index Register. ICA-Req-2 similarly requires systems to:

59 – Provide at least two naming mechanisms for records in the classification scheme:

⁵¹⁷ *The Bengal Manual*, p.9.

⁵¹⁸ *ICA-Req 2*, p.25.

- a mechanism for allocating a structured alpha, numeric or alphanumeric reference code (that is, an identifier which is unique within the classification scheme) to each classification level; and
- a mechanism to allocate a textual title for each electronic aggregation.

It must be possible to apply both identifiers separately and together.⁵¹⁹

While record names or titles are only significant in ICA-Req (other forms of metadata are significant for identifying records in the Bengal system, not least their unique numbers), in both the Manual and ICA-Req, unique numbers and free text file titles form the main identifying attributes of files / aggregations. In all cases, these are metadata that are essential to identifying records and files and enabling their retrieval, and relating them to their place within the larger system.

6.4.5. Foliation and Fly-Leafs

Foliation is the application of sequential numbers to all records on a given file. Its primary function is to enable referencing and information retrieval, but it also provides an assurance of the integrity of the file, which contributes to a presumption of a well-functioning system through which the authenticity of records can be more readily presumed. Disguising the removal of a record from a foliated file would require the renumbering of all subsequent records on the file. The Bengal Manual calls for foliation in rule 1.A.3.(1):

Each letter will have assigned to it a consecutive number within the file to which it belongs, called the *serial number* which should be inserted in a *big* figure in the space provided for the purpose in the standard forms.⁵²⁰

The Manual enlarges on this process later:

1.B.17 If the letter is in continuation of a file which has already been

⁵¹⁹ ICA-Req 2, p.31.

⁵²⁰ The Bengal Manual, p.4.

formed, the reference clerk will get out the file to which it belongs, will assign to the fresh letter the next consecutive serial number of the series of the letters of that file, will place the fresh letter on the top of the file, and write up a fly-leaf, which will remain with the file. When a private document or exhibit is received with a letter, its sub-number should be separately entered in column 1 of the fly-leaf, its description in column 2 and the letter D in column 5. The serial number of the letter should be entered in the space provided for the purpose in the standard forms or in the docket stamp.⁵²¹

Rule 1.B.17 uses foliation to create a trace of the record on the fly-leaf, so that covering the removal of a foliated record would require not only renumbering subsequent foliated records, but altering or removing the fly-leaf, which would alert users to a breach of the integrity of the file. Fly-leaves were expected to be present in every file:

1.G.62. In every file should be kept a fly-leaf, *i.e.*, list of the letters which constitute it, an addition being made to the list, as before directed (*see* rule 17 and 32), as each letter is received or issued...⁵²²

These fly-leaves will have acted as contents lists for the files, enabling information retrieval, but also as a measure against which to check the completeness of the file. As discussed above, the fly-leaves also worked in combination with foliation as a further safeguard against the illicit removal of records from files. In ICA-Req, it is the accumulation of metadata that fulfills this function. The standard requires systems to collect and keep metadata about events that happen to records and all levels of aggregation:

119 - Track events, once the metadata functionality has been activated, without manual intervention, and store in the metadata information.

120 - Maintain the metadata for as long as required.

⁵²¹ *The Bengal Manual*, p.7.

⁵²² *Ibid.*, p.15.

121 - Provide metadata of all changes made to:

- electronic aggregations (including volumes); [this would include the addition and removal of records, performing a function similar to foliation in the Bengal Manual's system]
- individual electronic records; and
- records management metadata associated with any of the above.⁵²³

These requirements also perform the same tasks performed by removal slips, receipts for transfer and certificates of safe custody sanctioned by the Bengal Manual for file tracking, described below.

6.4.6. Tracking: Removal Slips, Receipts for Transfer and Assurances of Safe Custody

In each office I encountered the same thing: a young secretary with a long ledger book detailing the movements of *expedientes* [my italics] all over the sprawling institution, inscribed in the order in which they were received in each department... The only way to find a document was by following its trail and the traces it left on other documents in its wake.⁵²⁴

The scenario Hetherington describes is common in many large organisations where the registry system is still in use. While visiting the Kenyan Ministry of Lands, I observed the use of file movement registers in each office and was told that the secretaries were conscientious in recording the movement of files into and out of their offices, since they were responsible for the safekeeping of all files from receipt in their office until receipt in another office or return to the central registry. In some cases, the messengers who carried files between offices were made to sign receipts, freeing secretaries from responsibility after files left their offices.

⁵²³ ICA-Req 2, p.38.

⁵²⁴ Hetherington, K., *Guerrilla Auditors: The Politics of Transparency in Neoliberal Paraguay*, (London: Duke University Press, 2011) p.146.

The Bengal Manual does not describe a system of file movement registers in departmental offices, but provides a number of means for recording the movement of records and files. One of these means is the use of removal slips within files: '1.G.65. When a paper is removed from a file, a removal slip note is to be kept in it to indicate where the paper has been removed'.⁵²⁵ The same principle is in operation in rule 1.G.66., which calls for removal slip notes (attached to file boards) to be put in the place of any register or bound volume that is removed from the record-room.

Responsibility for the custody of records is assigned to various officers throughout the life of the records. This is seen, perhaps most explicitly, in the Manual's requirements for receipts to be collected on the transfer of records:

1.G.66A. When any valuable document is transferred from any officer to any other officer or to the Government Pleader, the officer transferring should insist on taking a proper receipt for the same and such receipt should be carefully preserved.⁵²⁶

And the periodic submission of certificates assuring safe custody of registers and documents:

1.G.66B. [referring to registers and documents taken out by the Government Pleader]...the Government Pleader should be asked to submit monthly certificates to the Collector certifying that the registers and the documents are in the personal custody of the Government Pleader and are being properly safeguarded.⁵²⁷

The receipts envisaged in 1.G.66A are records that effect a transfer of responsibility for the records being transferred, extending and documenting the chain of custody outside of the organisation. Both the receipts and the certificates

⁵²⁵ *The Bengal Manual*, p.15.

⁵²⁶ *Ibid.*

⁵²⁷ *Ibid.*

required by 1.G.66B provide testimony to the transfer of responsibility and perform the kind of non-repudiation function MacNeil describes as being performed by all signatures.

In addition to the accumulation of metadata for events involving records and aggregations outlined in the previous section, ICA-Req compliant systems must also:

70 – Keep a clear history of the location of reclassified aggregations (including volumes) prior to their reclassification, so that their entire history can be determined easily.

131 – Provide a tracking feature to monitor and record information about the location and movement of both electronic and non-electronic aggregations.

132 – Record information about movements including:

- unique identifier of the aggregation or record;
- current location as well as a user-defined number of previous locations (locations should be user-defined)
- date item sent/moved from location;
- date item received at location (for transfers); and
- user responsible for the move (where appropriate).⁵²⁸

6.4.7. Cross-Referencing and ‘Stitching’

In the ‘Procedure for putting up letters’ we see how cross-referencing letters creates yet-more complex relationships through metadata:

1.B.17... [The reference clerk] will enter in pencil on the margin of the letter the serial number or numbers of the previous letters to which reference is made, and will, when necessary, get out any other files required for the elucidation of the subject-matter of the new letter...⁵²⁹

⁵²⁸ *ICA-Req 2*, pp.32-40.

⁵²⁹ *The Bengal Manual*, p.7.

The use of pencil suggests that these references were conceived of as temporary aids to the officers handling the file, but while they persisted, the references constituted metadata that created links across files. As discussed above, this kind of cross-referencing was also used to link replies to the letters that prompted them, which remains standard administrative practice today. Cross-referencing appears in ICA-Req. To be compliant, systems must: ‘76 – Allow users to create cross-references between related aggregations or between aggregations and individual records’.⁵³⁰

Instructions for managing notes on and orders arising from letters call for the stitching of any additional folios to the original letter:

1.B.19. ...The draft reply should be on a separate paper of foolscap size, and not on the letter itself, or should be afterwards copied on a separate paper by the office. If the order be written on a separate piece of paper, the piece should be foolscap size, and it should be stitched to the letter as directed in rule 4 above.⁵³¹

Rule 4 calls for stitching ‘with cotton in the upper left-hand corner’. Documents attached in this way ‘form one unit under one serial number in the file’.⁵³² This is an example of a physical trace that carries information that can be used to understand the history of the record – it forms part of the audit trail. In this respect, we could consider its digital counterpart to be the accumulation of metadata about event history described above, under *Foliation and Fly-Leafs* and *Tracking*.

6.4.8. Copying

The bureaucratic processes employed in the management and circulation of information sometimes left other traces in the form of copies. The Manual provides guidance on the production and distribution of copies:

⁵³⁰ *ICA-Req 2*, p.33.

⁵³¹ *The Bengal Manual*, p.8.

⁵³² *The Bengal Manual*, p.4.

1.C.34. It happens, not unfrequently, that orders or reports, in the course of communication, through official channels, to the officers for whom they are ultimately destined, are copied and recopied in every successive office they pass through. As typewriting machines are now in general use, a good deal of time and labour would be saved, if the issuing officers prepared and transmitted, with the original typewritten orders or reports the necessary number of typed copies for the office which they will ultimately reach. Ordinarily, three copies can be prepared on foolscap paper by a single operation from a type-writer. And if more than three copies are required, they may be obtained by duplicators where they are in use. Where there is no such machine, or when less than seven copies are required, three copies only should be prepared by type-writers and transmitted, leaving the remaining copies to be prepared in the office through which the orders or reports would pass.⁵³³

The Manual is silent on how these copies are to be treated by registry staff. There is no mention of the use of signatures or stamps on copies, nor how copies should be distinguished from their original, either in themselves or in the registers. As they were importing and, effectively, testing established archival concepts in the digital environment, the InterPARES research team approached the question of copies with an assumption that a digital original would exist and should be distinguishable from its copies. The InterPARES Requirement Set B calls for the activity of reproduction to be documented in a manner that includes:

- B.2.a The date of the records' reproduction and the name of the responsible person;
- B.2.b The relationship between the records acquired from the creator and the copies produced by the preserver;
- B.2.c The impact of the reproduction process on their form, content, accessibility and use; and

⁵³³ Ibid., p.10.

B.2.d In those cases where a copy of a record is known not to fully and faithfully reproduce the elements expressing its identity and integrity, such information has been documented by the preserver, and this documentation is readily accessible to the user.⁵³⁴

These requirements were intended for use after records have been accepted across the archival threshold, and they presumed traditional notarial processes for duplication would persist in the digital environment. Nevertheless, the underlying principle in their articulation is the documentation of custodial interventions, a principle that is pivotal in establishing informational trustworthiness. They ask for transparency in curation. The ICA-Req standard allows some flexibility in the management of what it calls ‘manifestations’:

24 – Allow users to choose at least one of the following where an electronic object has more than one manifestation:

- register all manifestations of the object as one record
- register one manifestation of the object as a record; or
- register each manifestation of the object as a discrete record.⁵³⁵

It is noteworthy that the National Archives of Australia, which conceives of digital records as ‘performances across time’, had a large part to play in the development of the standard.⁵³⁶ This may explain the flexibility of ICA-Req’s requirement 24. Significantly, there is scope for reducing duplication in the ability to assign records to ‘multiple aggregations without their duplication’ (Requirement 25), setting records in multiple contexts without duplication.⁵³⁷

6.4.9. Security Classifications

The Bengal Manual’s ‘rules for the treatment of confidential correspondence’

⁵³⁴ InterPARES 1 Authenticity Task Force, ‘Appendix 2 - Requirements for Assessing and Maintaining the Authenticity of Electronic Records’, in InterPARES 1 Project, *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project* (InterPARES Project, 2002) pp.7-8.

⁵³⁵ *ICA-Req 2*, p.27.

⁵³⁶ *Ibid.*

⁵³⁷ *Ibid.*

state:

1.F.52. Confidential papers should not pass through the office in the usual course. Only the head of the office and one or two trustworthy clerks (whose names are to be noted) should deal with them...⁵³⁸

This rule brings to the fore the significance of custodianship and the assignment of responsibility, which must be documented. That process of documentation is a mechanism of the system for protecting authenticity. The Manual identifies certain types of confidential papers that must be kept under lock and key. These:

1.E.54... should be kept in the office in a box or almirah fitted with a Chubb's lock, the key of which should remain with the Personal Assistant when there is one, or with the head assistant, unless for any reason the head of the office thinks it necessary to keep they key himself. They should be entered in a special register, in which should be entered the date of receipt, the number and date of the letter, the subject to which it relates, and the nature of the action taken. When a case is taken out of, or returned to, the almirah or box, a note to that effect should be prepared in offices in which the cases are considerable in number. The register and index should also be under lock and key. When the officer in charge of these papers goes on leave he should hand over the key to his successor, after verifying the contents of the box or almirah in his presence.⁵³⁹

There are a number of ways in which these types of controls are replicated in the functionalities of digital records management systems in ICA-Req. ICA-Req compliant systems must:

19 – Ensure that only authorized users and administrators can change the content of records management metadata elements.⁵⁴⁰

⁵³⁸ *The Bengal Manual*, p.13.

⁵³⁹ *Ibid.*

⁵⁴⁰ *ICA-Req 2*, p.26.

93 – Allow the administrator to limit access to records, aggregations and records management metadata to specified users or user groups.

108 – Allow security classifications to be assigned to records.⁵⁴¹

Rule 54 and the ICA-Req requirements demonstrate how custodianship, systems and metadata are used in combination to protect records. In the former case, responsibility is assigned, ultimately to the head of the office, metadata is captured in a ‘special register’, and a system of control is instituted, comprising procedures for storing, registering and checking and infrastructure for ensuring security.

6.4.10. Records Destruction

The destruction of records is made necessary by considerations of security, privacy, efficiency, space and the resources needed to preserve records of enduring value. Unauthorised and undocumented destruction calls into question the integrity of the system and implicates custodians in negligent or corrupt behaviour. The Bengal Manual notes in rule 1.I.100 that destroyed records must be noted in a memorandum according to rule 230(c), which requires the memorandum to be signed by the Collector or Subdivisional Officer.⁵⁴² In ICA-Req, records management systems must:

71 – Prevent the deletion of an electronic aggregation or any part of its contents at all times, with the exceptions of:

- destruction in accordance with a disposal authority; and
- deletion by an administrator as part of an audited procedure.⁵⁴³

163 – Record any deletion or disposal action comprehensively in the process metadata.

164 – Automatically record and report all disposal actions to the

⁵⁴¹ Ibid., pp.35-36.

⁵⁴² *The Bengal Manual*, p.21.

⁵⁴³ *ICA-Req 2*, p.32.

administrator.⁵⁴⁴

The custodial role of the Collector or Subdivisional Officer in authorising records destruction is performed by the system administrator in ICA-Req. The Manual states that registers were to be kept permanently, according to rule 1.I.97, and ICA-Req states systems must:

183 - Have the ability to retain records management metadata for records and aggregations that have been destroyed or transferred.

184 - Allow the administrator to specify a subset of aggregation records management metadata that will be retained for aggregations which are destroyed, transferred out or moved offline.⁵⁴⁵

These provisions mean that even after records are destroyed their metadata persists and it becomes possible to see what records once existed. The metadata, read in conjunction with the rules for records destruction (a feature of the *system*) and the signed memorandum, would enable an auditor to see if records destruction adhered to established norms, or if it was suspect. In this case, we see the interplay of metadata, system and custodian in an assurance of integrity.

6.4.11. Into the Archives

Eventually, records will fall out of frequent use and be appraised for their ongoing value. Different standards and practices are used to maintain the integrity of records once they have been accepted into the archives. It is not my intention to explore examples of those in the depth that I have looked at the Bengal Manual and ICA-Req, but to provide an overview of some of the authenticity controls used in the submission of records to and their acceptance by archives for permanent preservation.

The Manual requires the transfer of records from the registries to the record-room

⁵⁴⁴ Ibid., pp.44.

⁵⁴⁵ ICA-Req 2, p.45.

at District headquarters, along with the register that relates to them. The Manual states:

1.H.78. ...Without these registers the record-keeper must decline to receive any correspondence, except under express orders of the Collector. It is also the duty of the record-keeper to make such a scrutiny of the records made over to him as will satisfy him that all "C" papers have been destroyed, that each file has a proper fly-leaf, and that the papers are arranged in due serial order.⁵⁴⁶

These checks constitute the inspection of authenticity that Duranti envisaged when writing about the acceptance of records over the archival threshold as an assurance to users. Rule 1.H.78, in combination with 1.H.81, creates an environment in which metadata, systems and custodianship are again in interplay to provide assurances of authenticity:

1.H.81. When a year's correspondence is made over to the record-keeper, if any of the files are not made over to him because they have been transferred to the next year or because they are missing, he must report the facts to the Deputy Collector in charge of the record-room, and the Deputy Collector must satisfy himself in regard to the files alleged to have been so transferred that they have been transferred and must give the record-keeper a certificate to the effect...⁵⁴⁷

In the digital environment, this process continues. Government agencies will create digital records and will, after appraisal, transfer a portion of those records to the state archives' digital repository. There are a number of standards that inform the development and management of such repositories, and metadata, systems and custodians figure strongly in them. For instance, the OAIS standard discussed in Chapter 5 requires digital repositories to have defined management roles and data management and administrative processes that document

⁵⁴⁶ *The Bengal Manual*, p.17.

⁵⁴⁷ *Ibid.*

events.⁵⁴⁸ Many governments are still working to establish digital repositories and methods for the transfer of digital records, grappling with issues of secure transfer that might threaten the integrity of the records and their metadata, sensitivity review and digital redaction, and searchability through the use of e-discovery tools and machine learning. For example, in 2016 TNA began piloting the first digital transfers from government departments.⁵⁴⁹ Nevertheless, where digital transfers to archives are being made, the equivalent of the inspections of the record-keeper described in the Bengal Manual might be found in the automated characterisation (identification of file formats) and generation and validation of checksums (which can be used to check the integrity of digital records over time). The two components of authenticity, as defined by the InterPARES work, are identity and integrity. Characterisation addresses the need to establish identity and checksums can be used to establish both identity and integrity. TNA has developed a tool that performs both functions. DROID (Digital Record Object Identification) allows users to scan their systems to identify files, capturing metadata about names, file types and sizes. The tool also allows users to generate checksums for files, which can then be validated by TNA during ingest and whenever it is necessary to ascertain that file integrity has persisted.⁵⁵⁰

6.5. Conclusion

Though decades apart, the publication of the Bengal Manual and ICA-Req demonstrates a concern for control over information. This concern is grounded in the need for trustworthy information as a basis for decision-making and accountability, which are essential for well-functioning organisations, as Yates

⁵⁴⁸ International Standardisation Organisation, *ISO 14721:2012 Space data and information transfer systems – Open archival information system (OAIS) – Reference model* (Geneva: International Standardisation Organisation, 2012).

⁵⁴⁹ National Archives, *Reviewing Digital Records Management in Government* (London: The National Archives, 2014) <http://www.nationalarchives.gov.uk/information-management/manage-information/policy-process/reviewing-digital-records-management-government/research/> [accessed 4 October 2018].

⁵⁵⁰ National Archives, *DROID: User Guide*, (London: The National Archives, 2017) <http://www.nationalarchives.gov.uk/documents/information-management/droid-user-guide.pdf> [accessed 4 October 2018].

has shown.⁵⁵¹ The Bengal Manual is an example of an administrative manual produced by a bureaucracy that has its roots in the administrative practices of the East India Company, which attempted to manage its Indian business from London. The problem of management at a distance was solved, according to Hull, by giving ‘paper some of the qualities of discourse, people, places, and time through the use of signatures, dates, stamps, and interartifactual references’.⁵⁵² Hull sees the same trend in the digital environment, with increased automation reducing the need for human mediation: ‘Eyes, minds, and hands are to be replaced by satellites, computers, and printers in an attempt to restrict the human role in referential practices to interpretation’.⁵⁵³ Certainly, many of the measures set out in the Manual are automated in ICA-Req, which, as I have shown, replicates many of the same processes as the Manual, such as registration, cross-referencing and destruction. Though the technologies of record-making and record-keeping have changed, the need for controls to protect the authenticity of the records has persisted. A close reading of the Bengal Manual and ICA-Req demonstrates that metadata, systems and custodianship have and continue to be prominent in guidance on the defence of the authenticity of records. Importantly, for this thesis, the close analysis of these documents has provided specific examples of the kinds of record-keeping controls that support informational trustworthiness. While the limited record-keeping literature that has engaged with open government data has made claims that aspects of records thinking could support data quality, this thesis attempts to show what that might mean in reality, so it is important to be explicit about what these controls look like and how they work.

Hull wrote that official discourse ‘was anchored to people, places, times, and artifacts through an elaborate use of signatures, dates, and stamps’.⁵⁵⁴ These signatures, dates and stamps are metadata that provide contextual information about the handling of records in the same way that metadata in records, such as the date of writing and the signature of the author, provide information about the

⁵⁵¹ Yates, J., *Control through Communication: The Rise of System in American Management*, (London: The John Hopkins University Press, 1989).

⁵⁵² Hull, *Government of Paper*, p.257.

⁵⁵³ Ibid.

⁵⁵⁴ Hull, *Government of Paper*, p.10.

context of the records' creation. In ICA-Req compliant systems, metadata accrues as records pass through the various processes that constitute their active use. The metadata indicates the operations of the system in which they are held and the actions of their custodians.

The materiality of paper records enables them to be inspected for authenticity, physically. The inspection of digital records requires that the systems in which they are created, used and kept are capable of facilitating inspections for authenticity digitally. This requires continuous and documented control in the kind of system articulated by ICA-Req. Systems comprise pre-defined and documented processes and procedures, which require adherence by custodians, and they comprise checks on the execution of processes and procedures, which rely on the accrual and capture of metadata. Such checks may also be made on the metadata and custodians.

According to Hull, the bureaucratic manuals of the late nineteenth century defined relations and hierarchies between divisions and agencies of government at all levels.⁵⁵⁵ Throughout the Manual, the clerks, record-keepers and other officers have their responsibilities for the care and use of information spelled out. For example: '1.D.36. When the letter is entered in Register No. 61, the clerk in charge of that register will copy this order regarding reminders across columns 8 and 9 of the register and will be responsible for duly issuing the reminders'.⁵⁵⁶ Similarly, ICA-Req compliant systems enable controlled access and certain actions to be permitted to different classes of staff. Both assign responsibilities to and require the action (a signature or the press of a button) of users, establishing a chain of custody of records.

These controls become more effective as they work in combination. This is what Hull observed when he wrote:

In recognition that stamps can be reproduced or removed from their rightful place in an office, stamp images are supplemented by

⁵⁵⁵ Ibid.

⁵⁵⁶ *The Bengal Manual*, p.103.

autographic writing, which establishes an intertextual and interartifactual relation between the stamped graphic artifact and a diary book where movements are registered.⁵⁵⁷

The provisions of the Manual and ICA-Req are echoed across many different sources of professional guidance on record-keeping, such as PD ISO/TR 15801:2017 *Document management – Electronically stored information – Recommendations for trustworthiness and reliability* and the Public Records Office of Victoria PROS11/07 Capture Standard. This record-keeping thinking has developed over centuries to support authenticity in information. The level of detail in the analysis presented in this chapter is intended to show how these controls work together to protect authenticity in records. Though the statement has been made that record-keeping controls can help to improve open government data,⁵⁵⁸ this is the first attempt to articulate for the open data community what exactly is meant by ‘record-keeping controls’ and how they work. The following case studies will look at open government data initiatives – the latest instance of government monitoring mechanisms intended to help correct information asymmetry - with a view to the presence of record-keeping safeguards of authenticity in the three broad and interrelated categories of metadata, auditable systems and documented custodianship.

⁵⁵⁷ Hull, *Government of Paper*, p.131.

⁵⁵⁸ Thurston, A., ‘Trustworthy Records and Open Data’, *The Journal of Community Informatics*, 8:2, 2012.

Chapter 7

Case Study: The Kenya Open Data Initiative

7.1. Introduction

This chapter presents a case study of the Kenya Open Data Initiative (KODI), examining the level of control in place for the establishment and maintenance of authenticity in information released through Kenya's open data portal. KODI has been chosen because Kenya, as a Commonwealth country, shares the legal and administrative traditions that provide the context for information management practices – secrecy, registry practices, legal norms of evidence - allowing a closer comparison with Australia. KODI is noteworthy as the first functioning government open data portal in sub-Saharan Africa, and as such might be expected to have attained a maturity conducive to the study of the question of authenticity in open data. The case study begins with an historical overview of the development of openness in the Kenya government, from the embedded culture of secrecy, through the attempts to introduce Freedom of Information legislation, Kenya's involvement in the Open Government Partnership and the African Data Consensus, to the creation of KODI. This draws together information from scholarship in the areas of colonial history, international development, particularly in relation to public sector reform, the openness movement and information management. The case study then describes the current state of KODI. As described in the methodology, this draws on publicly available sources as well as secondary data that I obtained while working on unrelated projects in Kenya in 2010-11 and 2016. In order to identify issues relating to the authenticity of information released via KODI, the case study looks at a particular dataset available through the portal. Firstly, the data, which relates to land use, is set in context with an overview of Kenya's land management regime. Secondly, the case study describes the extant system for the management of land information. Thirdly, it describes the process for the preparation, publication and use of the land dataset. The case study then maps the lifecycle of the data and identifies strengths and weaknesses in the controls for the protection of that data's authenticity. The questions raised about authenticity call into question the quality of the data and its ability to mitigate information asymmetry in the principal-agent relationship between Kenyans and the Kenyan government.

7.2. Historical Development of Open Data in Kenya

7.2.1. Secrecy in Kenyan Government

In order to understand the development of open data in Kenya, it is first necessary to understand the pre-existing environment. A culture of secrecy has pervaded the Kenyan public sector. Nathan Mnjama's claim that 'Kenya like many other former colonial countries inherited a government that was not used to granting access to its public records' has yet to be backed up by historical research but nevertheless suggests, if not the inheritance of a culture of secrecy, the inheritance of practices informed by the idea that access should be on a 'need to know' basis.⁵⁵⁹ Low expectations of access are likely to have been shared across the individual/state divide, with colonial government culture and practice emanating out of the closed culture and practices of the British civil service as described in Chapter 4, and limited demand for information deriving from low expectations of access, aggravated by the nature of the individual/state relationship in the colonial context, where any claims of a principal-agent relationship would be problematic at best.

The tendency towards secrecy in the Kenyan public sector was encouraged by the *Official Secrets Act* (No. 11 of 1968). The Act criminalises the use and communication of documents or information 'which is calculated to be or *might be* [my emphasis] or is intended to be directly or indirectly useful to a foreign power or disaffected person' for purposes that are prejudicial to the 'safety or interests of the Republic'.⁵⁶⁰ The Act is particular about civil servants, stating that any person who possesses or controls information that 'has been entrusted in confidence to him owing to his position as a person who holds or has held a contract made on behalf of the Government' and who 'for any purpose or in any manner prejudicial to the safety or interests of the Republic—

- (i) uses the code word, plan, article, document or information; or
- (ii) retains the plan, article or document in his possession or under his control when he has no right so to retain it or when it is contrary to his duty

⁵⁵⁹ Mnjama, N., 'Access to Records and Archives in Kenya', *African Research and Documentation*, 106 (2008) p.62.

⁵⁶⁰ Republic of Kenya, *Official Secrets Act (No. 11 of 1968)*, section 3.1(c). (National Council for Law Reporting with the Authority of the Attorney-General).

so to retain it, or fails to comply with all directions issued by lawful authority with regard to its return or disposal, shall be guilty of an offence.⁵⁶¹

The broad provisions of the Act left room for interpretation to an extent that created anxiety in the public sector about the distribution of government information.⁵⁶² Further legal controls on the release of government information can be found in the 1960 *Preservation of Public Security Act*, 1985 *Service Commissions Act*, the 1998 *National Assembly (Powers and Privileges) Act* and 2012 *National Security Intelligence Service Act*.⁵⁶³

Commentators on access to government information in Kenya also point to another cause of government secrecy: the need to conceal corruption. Leonida Mutuku and Christine Mahihu describe a public sector culture in which ‘corrupt networks in public institutions benefited greatly from [a] culture of monopolizing access to information, and used this power to advance their personal interest, usually at the expense of the citizens.’⁵⁶⁴ Mnjama cites two newspaper articles that suggest that the link between secrecy and corruption is widely recognised in Kenyan society. In her 2003 article for *The Daily Nation*, Michelle Kagari wrote that the ‘culture of Government in Kenya has long been one of secrecy, bolstered by the traditional perception that Government is Serikali or the “big secret”’.⁵⁶⁵ She argues that government secrecy is both the cover for corruption, and partly responsible for the common perception that Kenyan government officials are corrupt. Mnjama also cites a 2007 *Kenya Times* article that states that

⁵⁶¹ Ibid., section 3.3(d).

⁵⁶² Interview notes, *Aligning Records Management with ICT, e-Government and Freedom of Information in East Africa*, International Records Management Trust, 2010.

⁵⁶³ Abuya, E., ‘Towards Promoting Access to Information in Kenya’, *African Network of Constitutional Lawyers*, (April 2011)

https://spaa.newark.rutgers.edu/sites/default/files/files/Transparency_Research_Conference/Papers/Abuya_Edwin.pdf [accessed 3 January 2017].

⁵⁶⁴ Mutuku, L. & Mahihu, C., *Open Data in Developing Countries: Understanding the Impacts of Kenya Open Data Applications and Services*, (Nairobi: iHub Research, 2014) p.21.

<http://www.opendataresearch.org/sites/default/files/publications/ODDC%20Report%20iHub.pdf> [accessed 26 May 2018].

⁵⁶⁵ Kagari, M., ‘You Can’t Fight Graft in Secret’, *The Daily Nation* (17 April 2003) quoted in Mnjama, ‘Access to Records and Archives in Kenya’, p.66.

government secrecy is the ‘fodder on which thrives corruption in public offices’.⁵⁶⁶ The *Kenya Times* article recognises a symbiotic relationship between secrecy and corruption: ‘When files can not be accessed when there is need to access them, then the tendency to grease palms becomes an imperative and we strengthen corruption’.⁵⁶⁷

The link between corruption and secrecy in Kenyan government has also been acknowledged in Parliament. On 8 July 1998, the Kenya National Assembly established a Parliamentary Anti-Corruption Select Committee to investigate corruption.⁵⁶⁸ Reporting to the National Assembly on its progress two years later, its Chair, Musikari Kombo, stated:

It is important to know that corruption in Kenya has created a sophisticated culture of secrecy and protection, resulting in a situation where to attempt to fight it, you have really to look in real dark allies [*sic*].⁵⁶⁹

Thus, inherited administrative culture and practice, together with corruption, may be taken as two coexisting contributing factors of the Kenyan government culture of secrecy. It was against this legal and cultural background, which favoured secrecy, that the push for transparency emerged from Kenyan civil society and international non-government organisations.

7.2.2. Access to Information: Open Data and Freedom of Information

Despite the overwhelming culture of secrecy, the 1969 Constitution of the Republic of Kenya did give Kenyans the right to ‘freedom to receive ideas and information without interference, freedom to communicate ideas and information without interference (whether the communication be to the public generally or to

⁵⁶⁶ Kenya Times, ‘Move on Information Freedom Enlightened’, *Kenya Times*, 27 February 2007, quoted in Mnjama, ‘Access to Records and Archives in Kenya’, p.66.

⁵⁶⁷ Ibid.

⁵⁶⁸ National Assembly of the Republic of Kenya, *Hansard*, 5 July 2000, p.1440.

⁵⁶⁹ Ibid., p.1444.

any person or class of persons)...'.⁵⁷⁰ Furthermore, Mnjama notes that Kenya was a signatory to the *International Covenant on Civil and Political Rights* and the *African Charter on Human and People's Rights*, both of which contained similar provisions.⁵⁷¹ However, no legislation explicitly gave effect to these rights.

Mutuku and Mahihu see Bitange Ndemo's appointment as Permanent Secretary of the Ministry of Information and Communications as a turning point in the history of Kenya government secrecy and openness. In 2010, during fieldwork for the International Records Management Trust's *Aligning Records Management with ICT, e-Government and Freedom of Information in East Africa* research project, Justus Wamukoya and I interviewed Ndemo about the state of open government in Kenya.⁵⁷² He described an ingrained culture of secrecy that thwarted public service delivery and economic growth, and described his vision for unlocking the potential of government data through the use of technology.

Appointed by the Kibaki regime in 2005, Ndemo's first major initiative was to map the distribution of Constituency Development Fund (CDF) money. The CDF was established by the CDF Act 2003 to fund local development projects; the CDF was allocated 2.5% of the Kenya Government's ordinary income. By mapping the distribution of CDF funds, Ndemo 'exposed the unfair allocation of funds'.⁵⁷³ So controversial were his findings that Ndemo's data map was taken offline under pressure from Members of Parliament.⁵⁷⁴ As this chronology of Kenya's move towards openness will show, the political and technological environment was not yet ready to sustain an initiative of this kind.

Despite the controversy aroused by Ndemo's project, other governmental moves towards openness continued, prompted by civil society. The Kenyan Section of the International Commission of Jurists, an international non-government organisation that promotes the '[the defence of] human rights and the rule of law

⁵⁷⁰ Republic of Kenya, *The Constitution of Kenya Act 1969*, s79(1).

⁵⁷¹ Mnjama, 'Access to Records and Archives in Kenya', p.68.

⁵⁷² International Records Management Trust, *Kenya Country Report*, (August 2011) <http://www.irmt.org/portfolio/managing-records-reliable-evidence-ict-e-government-freedom-information-east-africa-2010-%e2%80%932011> [accessed 25 May 2017].

⁵⁷³ Mutuku and Mahihu, *Open Data in Developing Countries*, p.21.

⁵⁷⁴ *Ibid.*, p.22.

worldwide',⁵⁷⁵ prepared a draft Freedom of Information Bill in 2006. In 2007, the Ministry of Information and Communication published a Freedom of Information Policy that mapped out the development of the Bill, the revision of conflicting legislation and the implementation of an FOI Act. Yet this would be the last significant step towards introducing a legal basis for government openness for some years.

When Wamukoya and I interviewed Ndemo in 2010, he expressed concern about the slow progress of the Freedom of Information Bill.⁵⁷⁶ No progress had then been made on the enactment of the Bill, but the implied right to information in the 1969 Constitution was made more explicit in a new Constitution. It granted citizens the right to access public sector information. It stated:

35. (1) Every citizen has the right of access to—
 - (a) information held by the State; and
 - (b) information held by another person and required for the exercise or protection of any right or fundamental freedom.
- (2) Every person has the right to the correction or deletion of untrue or misleading information that affects the person.
- (3) The State shall publish and publicise any important information affecting the nation.⁵⁷⁷

The provisions of the 2010 Constitution therefore created a legal basis for both reactive and proactive information release from the public sector, and this was finally given effect with the 2016 enactment of FOI legislation.

As well as a legal basis, open data requires an adequate technological infrastructure. Mutuku and Mahihu's report includes a 'Kenya Open Data Timeline', a visualisation of progress towards openness. It shows the establishment of the ICT Board (now the ICT Authority) in 2007, which marked a

⁵⁷⁵ International Commission of Jurists, *About*, <https://www.icj.org/about/> [accessed 20 September 2018].

⁵⁷⁶ Interview notes, *Aligning Records Management with ICT, e-Government and Freedom of Information in East Africa*, International Records Management Trust, 2010.

⁵⁷⁷ Republic of Kenya, *The Constitution of Kenya*, p.27.

recognition of the need to engage with technology in the public sector in a coordinated way.⁵⁷⁸ Meanwhile, Kenya's technological infrastructure continued to develop. Most significantly for open data, undersea fibre optic cables brought increased bandwidth to Kenya's government and people. SEACOM (from France) and The East African Marine System (TEAMS) (from the United Arab Emirates) landed in Mombasa in 2009 and the Eastern Africa Submarine Cable System (EASSY) (running from South Africa to Sudan with nine landing points on the east coast of Africa and onward connectivity to networks in Europe, the Americas and Asia) landed in Mombasa in 2010.⁵⁷⁹ Since these connections went live, infrastructure-building projects have continued to carry fibre optic cabling further inland in Kenya.⁵⁸⁰

Concurrently, steps were being taken to create an open data portal. An Open Data Task Force had been established, and included volunteers from the World Bank and the Kenya Bureau of Statistics and data, legal and technology experts. In our interview in 2010, Ndemo identified the open data portal as a way to bypass the legislative process in which the Freedom of Information bill was still mired at that time.⁵⁸¹ Ndemo's significance to the launch of the open data portal has been identified by Mutuku and Mahihu, who point to the good working relationship Ndemo and his Minister, Mutahi Kagwe, had with President Kibaki as the key to the progress of the portal.⁵⁸² On 8 July 2011, the Kenya Open Data Initiative portal was launched by President Kibaki.⁵⁸³

Despite the promise offered by KODI, however, there were significant problems with the early data releases. I have previously reported on work led by Anne Thurston, which discussed the publication of CSV files converted from other formats in such a way that 'the data was broken for several thousand rows, and

⁵⁷⁸ Ibid., p.24.

⁵⁷⁹ Internet Society, *Africa Internet History: Highlights*, https://www.internetsociety.org/wp-content/uploads/2017/09/history_internet_africa.pdf [last accessed 15 March 2019].

⁵⁸⁰ Interview notes, *Aligning Records Management with ICT, e-Government and Freedom of Information in East Africa*, International Records Management Trust, 2010.

⁵⁸¹ Ibid.

⁵⁸² Mutuku and Mahihu, *Open Data in Developing Countries*, p.23.

⁵⁸³ Davies, T., *Open Data Policies and Practice: An International Comparison* (European Consortium for Political Research, 2014) p.14. <https://ecpr.eu/Filestore/PaperProposal/d591e267-bcee-4d5d-b699-7d0bda633e2e.pdf> [accessed 16 January 2017].

anyone seeking to use the data would have had to go through a challenging and laborious process to rectify it'.⁵⁸⁴ The first tranches of household poverty data were also several years out of date at the time of publication.⁵⁸⁵ These problems with the first data releases may be attributable to the speed with which the open data initiative proceeded. Greg Brown of the Sunlight Foundation, a US NGO that promotes government transparency, ascribes to Jay Bhalla of the Open Institute the statement that Kenya's Open Data Task Force 'ran before it could walk'.⁵⁸⁶ Brown states:

Once President Kibaki approved the idea, the [Open Data Task Force] hurriedly developed the portal, gathered the data, and did their necessary legal and policy homework. Remarkably, the portal was launched a few weeks after gaining approval from President Kibaki.⁵⁸⁷

KODI was specifically named in Kenya's first Open Government Partnership (OGP) National Action Plan. Kenya signaled its intention to join the OGP in August 2011 and adopted its first OGP National Action Plan in March 2012. The plan included nine commitments, which involved improving transparency and public participation. Only one commitment related specifically to access to information:

3. Promoting transparency, accountability in government services by providing published datasets online, and in simplified formats that relate to public expenditures and disbursements in health, education, water and other essential services on the Kenya Open Data Portal.⁵⁸⁸

⁵⁸⁴ Lowry, J., 'Opening Government: Open Data and Access to Information', in Lowry, J. & Wamukoya, J. (eds.), *Integrity in Government Through Records Management: Essays in Honour of Anne Thurston* (Ashgate: Farnham, 2014) p.163.

⁵⁸⁵ Ibid.

⁵⁸⁶ Brown, G., *Why Kenya's open data portal is failing – and why it can still succeed*, <https://sunlightfoundation.com/2013/09/23/why-kenyas-open-data-portal-is-failing-and-why-it-can-still-succeed/> [accessed 25 May 2018].

⁵⁸⁷ Ibid.

⁵⁸⁸ Republic of Kenya, *OGP National Action Plan 2012-2013*, p.6. <https://www.opengovpartnership.org/documents/kenya-action-plan-2012-2013> [accessed 20 September 2018].

The target date for the completion of this work was May 2012. A year later, Greg Brown posted a blog that stated:

Kenya's open data portal is floundering. Despite the [excitement](#) that surrounded its launch in July 2011, the portal has not been updated in eight months, has seen stagnant traffic, and is quickly losing its status as the symbolic leader of open government in Africa.⁵⁸⁹

The problems, and the potential, of KODI were reflected in the 2013 report of the United Nations' High-Level Panel of Eminent Persons on the Post-2015 Development Agenda. The report outlined ideas for new development goals, post-2015, the deadline for the Millennium Development Goals, 'the world's time-bound and quantified targets for addressing extreme poverty'.⁵⁹⁰ These ideas included work on peace, inequality, climate and demographic change. The report called for a 'data revolution' for sustainable development:

with a new international initiative to improve the quality of statistics and information available to people and governments. We should actively take advantage of new technology, crowd sourcing, and improved connectivity to empower people with information on the progress towards the targets.⁵⁹¹

The report expands on this, stating that developments in information technology, and particularly in mobile technology, have not adequately penetrated the statistics community. The report states:

The post-2015 process needs to bring them together and start now to improve development data. Data must also enable us to reach the neediest,

⁵⁸⁹ Brown, *Why Kenya's open data portal is failing – and why it can still succeed*.

⁵⁹⁰ Executive Board of the United Nations Development Programme, the United Nation Population Fund and the United Nations Office for Project Services, *Management response to the evaluation of the role of UNDP in supporting national achievement of the Millennium Development Goals DP/2015/20*, p.1. <http://undocs.org/en/DP/2015/20> [accessed 20 September 2018].

⁵⁹¹ United Nations' High-Level Panel of Eminent Persons on the Post-2015 Development Agenda, *A New Global Partnership: Eradicate Poverty and Transform Economies through Sustainable Development*, (New York: United Nations Publications, 2013) p.21 <https://sustainabledevelopment.un.org/content/documents/8932013-05%20-%20HLP%20Report%20-%20A%20New%20Global%20Partnership.pdf> [accessed 8 September 2018].

and find out whether they are receiving essential services. This means that data gathered will need to be disaggregated by gender, geography, income, disability, and other categories, to make sure that no group is being left behind.⁵⁹²

By linking data disaggregation with development needs in this way, the report positioned data curation as a key analytic tool with the potential to determine development priorities and influence international and national budgetary allocations across all areas of government service provision. The report anticipates the service delivery and financial consequences of a data revolution when it states that ‘Better data and statistics will help governments track progress and make sure their decisions are evidence-based; they can also strengthen accountability’.⁵⁹³ The report does not elaborate on the issue of accountability, but notes that ‘This is not just about governments. International agencies, CSOs and the private sector should be involved’.⁵⁹⁴ This concept, in which non-government organisations have a role to play in the use of data for accountability, is the same concept at work in KODI. The staff of the ICT Authority which manages KODI see non-government actors as part of a ‘value chain’ in which civil society organisations are intermediaries between government and individuals, auditing, analysing, augmenting and rendering government data to end users.⁵⁹⁵ This idea is explored in the following section, *The Role of Civil Society*.

In response to the High-Level Panel of Eminent Persons on the Post-2015 Development Agenda report, the United Nations Economic Commission for Africa (UNECA) coordinated, with pan-African organisations, an expert group meeting in Nairobi to brainstorm the concept of a data revolution in Africa.⁵⁹⁶ There is no publicly available information about why Nairobi was chosen as the site of the

⁵⁹² Ibid., p.23.

⁵⁹³ Ibid., p.24.

⁵⁹⁴ Ibid.

⁵⁹⁵ Interview with Sifa Mawiyoo, Open Data Specialist and GIS Technologist, Kenya Open Data Initiative, ICT Authority; and Prestone Adie, Data Analyst, Kenya Open Data Initiative, ICT Authority, Nairobi, 20 September 2016.

⁵⁹⁶ United Nations Economic Commission for Africa (UNECA), *Statisticians converge on African Data Revolution* (28 March 2014) <http://www.uneca.org/stories/statisticians-converge-african-data-revolution> [accessed 16 January 2017].

meeting, but the choice may suggest that Kenya was perceived as a leader in this area. That expert group, later referred to as the Data for African Development Working Group, presented its work on 27 March 2014, at the Seventh Joint Annual Meetings of the Economic Commission for Africa (ECA) Conference of African Ministers of Finance, Planning and Economic Development and African Union (AU) Conference of Ministers of Economy and Finance, in Abuja, Nigeria.⁵⁹⁷ A UNECA article on the Abuja meeting noted that the expert group identified four problems in the area of data collection and use in Africa: ‘that National Statistics Offices (NSOs) have limited independence and unstable budgets...; incentives are misaligned and data quality checks are limited...; ...donor priorities take precedence over national priorities; ...access to and use of data is very limited.’⁵⁹⁸ The working group outlined three priorities for correcting these problems:

First, build institutions that produce accurate, unbiased data. Secondly, improve access, accuracy, and coordination of data. Finally, NSOs should build capacity by leveraging regional and multilateral partnerships.⁵⁹⁹

Meanwhile, the problems described by the working group were becoming increasingly evident in Kenya. In early 2014, Kenya’s OGP Independent Reporting Mechanism (IRM), Geoffrey Runji Njeru, published his report on the delivery of Kenya’s first OGP national action plan. He noted:

Although the government launched the open data portal, databases related to public expenditures and disbursements in education [*sic*], water and other essential services – as committed to in the nation action plan – are absent.⁶⁰⁰

Writing in 2014, Mutuku and Mahihu attributed the decline in enthusiasm for the portal – observed by Brown and confirmed by Njeru - to the changing political

⁵⁹⁷ Ibid.

⁵⁹⁸ Ibid.

⁵⁹⁹ Ibid.

⁶⁰⁰ Runji Njeru, G., *Independent Reporting Mechanism Kenya Progress Report 2012-13*, (Open Government Partnership, 2013) p.23, http://www.opengovpartnership.org/sites/default/files/Kenya_final_2012_0.pdf [accessed 25 May 2018].

landscape following the constitutional reforms of 2010, which set in motion the devolution of powers to new county governments, and the transition of power following the 2013 elections. Also in 2014, Anne Muigai of the Open Institute ascribed the decline of the portal to the leadership vacuum created when, in 2013, Ndemo left his post to take up an associate professorship at the University of Nairobi:

Nearly three years since its launch, KODI has been noticeably inert, especially since the exit of its most prominent champion, Bitange Ndemo. Prior to his departure, Dr Ndemo was able to mount pressure on his counterparts in other ministries as well as the Kenya National Bureau of Statistics (KNBS) to obtain additional data for the initiative.⁶⁰¹

The 2013 elections took place in March. It is likely that Ndemo left public service around this time, roughly the eight months alluded to by Brown, supporting Muigai, Mutuku and Mahihu's suggestion that the change in government and the departure of Ndemo were the cause of Kenya's open data slump. Mutuku and Mahihu were hopeful that the new president, Uhuru Kenyatta, would reinvigorate the open data initiative.⁶⁰² They pointed to the Jubilee Alliance manifesto commitment to cleaning and managing databases, and although the manifesto conceives of this as a means to promote efficiency across government, rather than as a resource for open data, elsewhere the manifesto mentions aspirations to transparency.⁶⁰³ Muigai pointed to the ICT Authority's 2014 invitation to firms to deliver engagement and development work on KODI as an indicator that government activity on open data in Kenya would revive.⁶⁰⁴

In March 2015, a conference on the subject of 'Data Revolution in Africa' was held in Addis Ababa, Ethiopia, organised by the pan-African organisations at the request of the Joint Conference in response to the report and discussions in

⁶⁰¹ Muigai, A., *A Look Back at Kenya's Open Data Journey*, <http://www.openinstitute.com/a-look-back-at-kenyas-open-data-journey/> [accessed 25 May 2018].

⁶⁰² Mutuku and Mahihu, *Open Data in Developing Countries*, p.23.

⁶⁰³ Ibid.

⁶⁰⁴ Muigai, *A Look Back at Kenya's Open Data Journey*.

Abuja.⁶⁰⁵ At this conference, the Africa Data Consensus was adopted, and has been published by UNECA, the African Union, the African Development Bank and the United Nations Development Programme.⁶⁰⁶ The *Africa Data Consensus* states:

A sustained data revolution is needed to drive social, economic and structural transformation in every African country. Such a revolution will also make it easier to track our countries' progress towards meeting national and globally agreed sustainable development goals, with a view to leave no one behind.⁶⁰⁷

It provides the first formal definition of an African Data Revolution:

A profound shift in the way that data is harnessed to impact on development decision-making, with a particular emphasis on building a culture of usage. The process of embracing a wide range of data communities and diverse range of data sources, tools, and innovative technologies, to provide disaggregated data for decision-making, service delivery and citizen engagement; and information for Africa to own its narrative.⁶⁰⁸

It articulates challenges including:

- Uneven coverage, lack of disaggregated data and data stored in sector-specific silos
- Lack of accessible usable information that is open to all communities
- A lack of harmonisation of data collected by different sources in different formats
- A lack of timely, accurate, comparable and relevant data.
- Weak data governance and accountability

⁶⁰⁵ UNECA, African Union, African Development Bank & United Nations Development Programme *Africa Data Consensus* (29 March 2015) http://www.uneca.org/sites/default/files/PageAttachments/final_adc_-_english.pdf [accessed 25 May 2018].

⁶⁰⁶ Ibid.

⁶⁰⁷ Ibid., p.3.

⁶⁰⁸ Ibid., p.2.

- Concerns over privacy, data protection and intellectual property...
- Lack of common standards allowing comparison of data across sectors and countries.⁶⁰⁹

As this case study will show, these challenges are evident in the data released through KODI.

In the *New Global Partnership* Report, the United Nations' High-Level Panel of Eminent Persons on the Post-2015 Development Agenda recommended:

establishing a **Global Partnership on Development Data** that brings together diverse but interested stakeholders – government statistical offices, international organisations, CSOs, foundations and the private sector. This partnership would, as a first step, develop a global strategy to fill critical gaps, expand data accessibility, and galvanise international efforts to ensure a baseline for post-2015 targets is in place by January 2016.⁶¹⁰

This recommendation led to the establishment of the Global Partnership for Sustainable Development Data (GPSDD), which includes governments (including the Government of Kenya), civil society groups, private companies, universities and others.⁶¹¹ The GPSDD is working on a number of initiatives, including a DATA4SDGS Toolbox to help countries develop and implement 'data roadmaps for sustainable development',⁶¹² developing best practice on data architectures, and sharing information about available tools and resources.⁶¹³ As this case study will show, this kind of collaboration across government, civil society and the private sector has become characteristic of the open data ecosystem in Kenya.

7.3. Current State of Kenya's Openness Programme

⁶⁰⁹ UNECA, African Union, African Development Bank & United Nations Development Programme *Africa Data Consensus*, p.3.

⁶¹⁰ United Nations' High-Level Panel of Eminent Persons on the Post-2015 Development Agenda, *A New Global Partnership*, p.24.

⁶¹¹ A full list of its members is available at <http://www.data4sdgs.org/index.php/partner-listing> [accessed 20 September 2018].

⁶¹² Global Partnership for Sustainable Development Data, *Data4SDGs Toolbox* <http://www.data4sdgs.org/initiatives/data4sdgs-toolbox> [accessed 8 September 2018].

⁶¹³ Ibid.

This section will draw on previous scholarship and new data to provide an overview of the current state of the Kenyan openness regime.

The World Wide Web Foundation's Open Data Barometer (ODB) ranks countries on three criteria:

- Readiness: How prepared are governments for open data initiatives? What policies are in place?
- Implementation: Are governments putting their commitments into practice?
- Impact: Is open government data being used in ways that bring practical benefit?⁶¹⁴

The ODB's methodology draws on government self-assessment, peer-reviewed expert survey responses, detailed dataset assessments and secondary data.⁶¹⁵

The Open Data Barometer Regional Report for Africa (3rd Edition, 2016) examines twenty-one sub-Saharan African countries, including Kenya, which is the highest ranking sub-Saharan African country, at 42 of the 92 countries studied globally.⁶¹⁶ The key findings of the Regional Report are that:

1. *Performance across the continent is relatively poor globally, and in comparison to leading countries in the Global South.* With regard to Kenya, the report notes that Kenya 'does not publish a single, fully open dataset — health, education and legislation data are open licensed but fall short of being fully open because the data are not available in bulk'.

2. *A downward trend is all too common in the overall Barometer scores from 2013 to 2015.* The assessment shows a drop in ten points in Kenya's ODB score between 2013 and 2015, which correlates with the downturn in activity described by Muigai, Mutuku and Mahihu.

⁶¹⁴ World Wide Web Foundation, *Open Data Barometer: ODB Methodology - v1.0 28th April 2015*, p.3 <http://opendatabarometer.org/doc/3rdEdition/ODB-3rdEdition-Methodology.pdf> [accessed 25 May 2018].

⁶¹⁵ Ibid., p.3.

⁶¹⁶ World Wide Web Foundation, *Open Data Barometer, Third Edition, Regional Report, Africa*, (May 2016) p.6. <http://opendatabarometer.org/doc/3rdEdition/ODB-3rdEdition-AfricaReport.pdf> [accessed 25 May 2018].

3. *Open data initiatives lack long-term commitment and resources, resulting in short-term gains that are unsustainable.* ODB data shows a drop in scores between 2014 and 2015 in all countries but Nigeria and Cameroon. However, the report notes ‘While there is a net decline in the scores for Kenya and Mauritius between 2013 and 2015, there is a recovery in their overall scores between 2014 and 2015’.

4. *ODB implementation scores are lower than readiness scores, and the gap is not closing.* The ODB’s comparison of readiness and implementation scores show a consistent gap in the case of Kenya, so that both scores increase and decrease in parallel, with no sign of the gap closing.

5. *There is no stand-out performer in Africa.* The report states that ‘Africa is the only region without a clear open data champion... In previous editions, Ghana and Kenya looked likely to assume this role, but the data show that the performance of these countries is erratic.’ Again, this may be attributable to the impact of constitutional and political changes in Kenya during the period. The ODB Global Report acknowledges that Kenya and Ghana are in ‘a holding pattern as they try to revamp their initiatives.’⁶¹⁷

Despite the perceived problems with KODI, ODB data visualisations show that Kenya government policies and action remained consistent over the 2013-2015 period, exceeding the sub-Saharan African mean. This suggests that the Kenyatta government maintained the commitment to openness made by the Kibaki government. This is borne out by the 2015 study by François van Schalkwyk, Michelle Willmers and Tobias Schonwetter, who attempted to understand if open data practices were being embedded in African governments, using institutional theory to compare Kenya and South Africa. In relation to Kenya, they found that ‘the government executive is responding to external pressures by conforming rather than buffering itself from pressures to embed open practice’.⁶¹⁸ However, the same study notes a disconnect between political

⁶¹⁷ World Wide Web Foundation, *Open Data Barometer, Third Edition, Regional Report, Africa*, p.30.

⁶¹⁸ van Schalkwyk, F., et al, *Embedding Open Data Practice: Developing Indicators on the Institutionalisation of Open Data Practice in Two African Governments*, (Cape Town: University of Cape Town Intellectual Property Unit, 2015) p.29. <http://webfoundation.org/docs/2015/08/ODDC-2-Embedding-Open-Data-Practice-FINAL.pdf> [accessed 25 May 2018].

will at the executive level and implementation at the agency level. Studying the Kenya Institute for Public Policy Research and Analysis (KIPPRA) and the Kenya National Bureau of Statistics (KNBS), van Schalkwyk, Willmers and Schonwetter found that ‘there is little evidence of [KIPPRA] conforming to pressures – either from the middle structure or directly from external stakeholders – to embed open data practices’, while KNBS was more responsive, which the authors attribute to the key role KNBS plays in Kenya’s open data initiative.⁶¹⁹ This echoes Mutuku and Mahihu’s assertion that, since the establishment of KODI (which marks some level of acceptance of the openness agenda at the executive level after the CDF maps were taken offline), it is at the agency level that resistance to openness has been most apparent. As van Schalkwyk, Willmers and Schonwetter conclude:

Governments should be cautious when relying on open data assessments and rankings as indicators of success. In developing country contexts, assessments and rankings may capture and reflect high-level commitments to open data and fledging initiatives that result in a batch of open datasets being published, but they are unlikely to reflect the fundamental organisational dynamics and complexities that may hinder, or at least delay, long-term, sustainable – in other words, embedded – open data practice.⁶²⁰

Although there are some problems with van Schalkwyk, Willmers and Schonwetter’s data,⁶²¹ their conclusion about the significance of agency buy-in is supported by the findings of Brown, Mutuku and Mahihu.⁶²² They find that KNBS is in close alignment with the Kenyan executive, unlike its South African counterpart: ‘A possible explanation... is that KNBS and the government’s open data initiative (KODI) share data. To some extent there appears to be a pact or

⁶¹⁹ The study notes of KIPPRA’s non-compliance ‘Whether this is because of multiple and possibly conflicting institutional pressures is difficult to say based on the data available’ (p.29).

⁶²⁰ Schalkwyk et al, *Embedding Open Data Practice*, p.35.

⁶²¹ The authors attempt to draw generalisations about Africa from a comparison of only two countries, but note ‘Without the inclusion of governments from developed countries in our sample, we cannot make any claims as to whether the indicators are sensitive to national contexts or whether they are only sensitive inter-organisational differences. To this extent, the inclusion of “Africa” in the title of this report is a red herring’ (p.20). They also acknowledge the limited scope of their data in relation to the number of agencies studied in each country.

⁶²² Brown, *Why Kenya’s open data portal is failing*; Mutuku & Mahihu, *Open Data in Developing Countries*.

some form of mandate between the executive and KNBS, and therefore between the middle and under structures, in Kenya'.⁶²³ Mutuku and Mahihu's timeline flags the establishment of KNBS and the demise of its predecessor, though the authors do not explain the significance of this change. A study of the motivation for and progress of this change might reveal a political move to put in place an outward-facing bureau more aligned to the growing openness agenda.

In an interview in September 2016, Mutuku spoke of a concern about the government's data release cycle, stating that 'government is struggling to make data releases regularly', which she attributed to a poor legal framework for proactive data release.⁶²⁴ She also flagged concerns over the trustworthiness of data being issued by the Bureau of Statistics, noting that KODI is primarily concerned with format and licensing, rather than reliability and authenticity.⁶²⁵

On the KODI side, there is a good deal of enthusiasm in the ICT Authority for improving and expanding the open data programme.⁶²⁶ Members of the KODI team participate in the Data Science Africa research network and its annual conferences.⁶²⁷ The KODI team has a network of 'fellows' across the public sector who source data, and a Data Science Team that cleans data released by government agencies.⁶²⁸ Although the *African Data Consensus* has not yet been ratified by President Kenyatta,⁶²⁹ the challenges it lists are recognised by the staff of the ICT Authority and inform much of the work associated with KODI.⁶³⁰ KODI staff are aware of problems with the data they are receiving for publication through the portal. They cite many of the same problems reported by civil society

⁶²³ Schalkwyk et al, *Embedding Open Data Practice*, p.29.

⁶²⁴ Interview with Leonida Mutuku, Nairobi, Kenya, 16 September 2016.

⁶²⁵ Ibid.

⁶²⁶ Interview with Sifa Mawiyoo, Open Data Specialist and GIS Technologist, Kenya Open Data Initiative, ICT Authority, and Prestone Adie, Data Analyst, Kenya Open Data Initiative, ICT Authority, Nairobi, 20 September 2016.

⁶²⁷ *Data Science Africa*, <<http://www.datascienceafrica.org/>> [accessed 25 May 2018].

⁶²⁸ Interview with Sifa Mawiyoo, Open Data Specialist and GIS Technologist, Kenya Open Data Initiative, ICT Authority, and Prestone Adie, Data Analyst, Kenya Open Data Initiative, ICT Authority, Nairobi, 20 September 2016.

⁶²⁹ Interview with Sandra Musoga, Senior Programs Officer – Transparency, Article 19, Nairobi, Kenya, 16 September 2016.

⁶³⁰ Interview with Sifa Mawiyoo, Open Data Specialist and GIS Technologist, Kenya Open Data Initiative, ICT Authority, and Prestone Adie, Data Analyst, Kenya Open Data Initiative, ICT Authority, Nairobi, 20 September 2016.

and the ODB, including the questionable accuracy and completeness of the data, the risk of the introduction of errors during the cleaning and curation of data, problems of timeliness and infrequent updates, and ongoing resistance to data release within the public sector, noting that ‘the quantity of data [they] receive does not seem to reflect the enactment of the FOI law’ in August 2016.⁶³¹ At the time of our interview in September 2016, KODI team members had not considered authenticity and the significance of contextual information, but they noted that they ‘hardly ever get data with metadata’.⁶³²

However, they have put in place a number of quality control measures. KODI staff take a snapshot of the portal with every new data upload. They also keep copies of original datasets as received from ministries, departments and agencies, so that sources can be checked in the event that curated data is queried. Regarding the issue of frequency of data releases, the KODI team has created a ‘data release calendar’, which they use to schedule and monitor updates. They have also created templates for datasets, which encourages more complete and consistent data.⁶³³

Mutuku and Mahihu noted that government agencies continued to withhold data after the promulgation of the new constitution.⁶³⁴ This echoes Darch and Underwood’s statement that the ‘framing of the access to information right in the adversarial context of most freedom of information laws recognises to some extent that “the state” is not a monolithic actor in the gathering and archiving of information, but is itself a site of struggles between different and real bureaucratic and political interests’.⁶³⁵ Rather than seeing Freedom of Information legislation as a separate and different means of opening government, Mutuku and Mahihu suggested that the enactment of the long-awaited bill would give effect to Article 35 of the 2010 national Constitution and create the legal basis that would compel

⁶³¹ Ibid.

⁶³² Ibid.

⁶³³ Interview with Sifa Mawiyoo, Open Data Specialist and GIS Technologist, Kenya Open Data Initiative, ICT Authority, and Prestone Adie, Data Analyst, Kenya Open Data Initiative, ICT Authority, Nairobi, 20 September 2016.

⁶³⁴ Mutuku and Mahihu, *Open Data in Developing Countries*, p.26.

⁶³⁵ Darch, C. & Underwood, P., *Freedom of Information and the Developing World: The Citizen, the State and Models of Openness* (Oxford: Chandos, 2010) p.249.

agencies to release information. Similarly, Brown sees a dependency between freedom of information and open data in Kenya:

A more open legal framework is also critical to the portal's success. The [lack of a freedom of information law](#) has made it difficult to get government institutions to release information and has weakened the standing of pro-openness organizations and individuals.⁶³⁶

On 31 August 2016, President Kenyatta assented to the Access to Information bill.⁶³⁷ The Access to Information Act 2016 sets out the process by which citizens can request information from public sector bodies and includes sanctions against public servants who fail to comply with the Act.⁶³⁸ At the time of my work in Kenya in September 2016, the implementation of the Act had not started. A national open data policy was being drafted, which, if adopted, would be an additional imperative for data releases.⁶³⁹

Also in August 2016, the GPSDD supported the East Africa Open Data Conference in Nairobi, in which members of the KODI team participated.⁶⁴⁰ The range of organisations that supported the conference, the membership of the GPSDD, and my interviews with KODI and Kenyan civil society all demonstrate the close collaboration between governments, civil society, companies and others that exists in the international and Kenyan open data ecosystems.

Further development of Kenya's open data programme is likely to be driven by the commitments in Kenya's second OGP national action plan (2016-2018), which do not include developments in the portal or the processes and structures

⁶³⁶ Brown, *Why Kenya's open data portal is failing – and why it can still succeed*.

⁶³⁷ Republic of Kenya, *Access to Information Act 2016*

<http://kenyalaw.org/lex/actview.xql?actid=No.%2031%20of%202016> [accessed 20 September 2016].

⁶³⁸ Ibid.

⁶³⁹ Interview with Sifa Mawiyoo, Open Data Specialist and GIS Technologist, Kenya Open Data Initiative, ICT Authority, and Prestone Adie, Data Analyst, Kenya Open Data Initiative, ICT Authority, Nairobi, 20 September 2016.

⁶⁴⁰ Carranza, J., *East Africa's Open Data Revolution* <http://www.data4sdgs.org/master-blog/2016/9/7/east-africas-open-data-revolution> [accessed 25 May 2018].

that support it, but do call for the release of specific types of data, for instance in the areas of climate change and extractive industries.⁶⁴¹

7.4. The Role of Civil Society

Given that this research is concerned with information asymmetry in the principal-agent problem of the social contract, and in view of the traditionally adversarial relationship between government and civil society on the question of access to information, it is useful to consider, separately, the role of the principal – individuals and civil society - in attempting to enable monitoring by opening Kenyan government information.

The openness movement in Kenya has benefitted from the support of champions within government, such as Ndemo, the current staff in the Deputy President's office,⁶⁴² and technologists within the ICT Authority, but much of the push for government openness has come from local and international civil society actors. The relationship between the Kenyan government and civil society organisations has historically been adversarial where access to information is concerned, particularly as it relates to the accountability and anti-corruption agendas. This antagonism is seen most starkly in Michaela Wrong's best-selling account of John Githongo's detection of the Anglo Leasing and Finance Company fraud, following which he received death threats and was forced into exile in Britain.⁶⁴³

Civil society's push for more openness has been instrumental in improving government acceptance of the open data agenda, which in turn has created an environment in which civil society engagement with government, and government data, has flourished. Kenya's IRM report notes:

⁶⁴¹ Republic of Kenya, *OGP National Action Plan 2*, http://www.opengovpartnership.org/sites/default/files/Kenya_AP2_2016_0.pdf [accessed 25 May 2018].

⁶⁴² Interview with Sandra Musoga, Senior Programs Officer – Transparency, Article 19, Nairobi, Kenya, 16 September 2016.

⁶⁴³ Wrong, M., *It's Our Turn to Eat*, (London: HarperCollins, 2009).

The launch of the open data portal spurred efforts from outside the government, mainly from civil society, to strengthen the open government data space. Two examples are Data Bootcamp (www.databootcamp.org) and Code for Kenya (www.code4kenya.org). Data Bootcamp is an effort to provide training to journalists and civil society organizations interested in using open government data. Code for Kenya places fellows, who are computer or data experts, in media and civil society organizations to improve the host's use and understanding of open government in Kenya.⁶⁴⁴

In the wake of the advances in Freedom of Information and open data, those civil society actors who drove the openness movement are turning from lobbying to implementation support. For example, the NGO Article 19 has long been an advocate for Freedom of Information in Kenya.⁶⁴⁵ It is now focusing on public and public sector sensitisation, educating citizens in their rights to access to information and public servants in their obligations under the Access to Information Act, encouraging them to think about the systemic and cultural changes that will be required.

The IRM Report recognises the crucial role of civil society in what the KODI staff call the value chain: 'These civil society efforts are encouraging, as buy-in and participation by civil society is necessary for open government data to be meaningful and relevant'.⁶⁴⁶ This may be seen in the critical use civil society is making of open government data. In its work on open contracting, Article 19 found that KODI data references other government sites, but that these sites were often offline. Article 19 has also frequently encountered file formats that could not be opened, 'numbers that don't mean anything' and 'data without assurances'.⁶⁴⁷

⁶⁴⁴ Njeru, *Independent Reporting Mechanism Kenya Progress Report 2012-13*, pp.22-23.

⁶⁴⁵ Article 19, *Eastern Africa*, <<https://www.article19.org/regional-office/eastern-africa/>> [accessed 20 September 2018].

⁶⁴⁶ Njeru, *Independent Reporting Mechanism Kenya Progress Report 2012-13*, p.23.

⁶⁴⁷ Interview with Sandra Musoga, Senior Programs Officer – Transparency, Article 19, Nairobi, Kenya, 16 September 2016.

The now symbiotic relationship between government and civil society in the open data arena is clear from the government support for work that scrutinises government performance. As the IRM for Kenya has reported:

The Kenya ICT Board is in the process of awarding grants to developers to create “high impact” applications using data from the portal. Already, the team behind the Ushahidi-powered platform, Huduma (Kiswahili for ‘services’), has used the data to map and explore access to health, infrastructure and education. Virtual Kenya (www.virtualkenya.org) has built an application for mapping counties where MPs have refused to pay taxes.⁶⁴⁸

Aside from its role in implementation and scrutiny, civil society is also tackling public issues through the use of data outside of the public sector. Open Duka (‘duka’ is Swahili for ‘shop’⁶⁴⁹), an initiative of the Open Institute, is scraping data from various sources, including legal cases and company registration databases, to provide information on Kenyan ‘entities’, such as organisations (contractors, litigants, etc.), people (public appointees, litigants, etc.), land parcels (owners), etc. Open Duka aims to ‘provide citizens, journalists, and civic activists with a practical and easy-to-use tool to understand the ownership structure of the world they live in’.⁶⁵⁰ Open Duka has experienced a number of common problems with opening data. Andrew Young and Stefaan Verhulst found that its ‘... organizers underestimated the difficulty of sourcing data, of government resistance, and of convincing a somewhat disenchanted population to use it’.⁶⁵¹

At a more grassroots level, civil society in Kenya has had a number of successes with crowd sourcing data. In 2009, the Map Kibera Trust initiated a citizen-led mapping project to map Kibera, the largest urban slum in Africa. Using GIS devices, Kibera community members mapped Kibera, documenting health clinics,

⁶⁴⁸ Njeru, *Independent Reporting Mechanism Kenya Progress Report 2012-13*, pp.23.

⁶⁴⁹ Young, A. & Verhulst, S., *Kenya’s Open Duka: Open Data for Transactional Transparency*, (OMIDYAR Network, January 2016) p.7, <http://odimpact.org/files/case-study-kenya.pdf> [accessed 25 May 2018].

⁶⁵⁰ Open Duka, *About Us*, <http://www.openduka.org/index.php/about> [accessed 25 May 2018].

⁶⁵¹ Young & Verhulst, *Kenya’s Open Duka*, p.15.

water points, and other services via OpenStreetMap.org.⁶⁵² This information is shared with the community and is used as a basis for local action plans and engagement with government. In another example of crowdsourced data, Chief Francis Kariuki, administrative chief of Lanet Umoja, Nakuru North District, known as ‘The Tweeting Chief’ for his use of Twitter to mobilise his community against local crises, has secured services for his district.⁶⁵³ Although information about Kariuki’s data work is limited, an Open Institute video of a presentation by Kariuki shows him explaining a community data collection initiative that produced baseline data on access to clean water, which resulted in the donation of water filters to approximately 3000 households.⁶⁵⁴

Civil society was instrumental in driving forward open government in Kenya. It has gained a foothold and is now recognised by government as part of the open data value chain.⁶⁵⁵ Currently, civil society organisations in Kenya are engaging communities, developing civic technologies, scrutinising government data, and sourcing data for civic purposes from the people of Kenya.⁶⁵⁶ However, the ODB data visualisations for Kenya show that the social impact of open data has been consistently low.⁶⁵⁷ This points to unavailability of data that meets user needs, low user uptake, and a lack of intermediary technologies.⁶⁵⁸ Notwithstanding a natural lag between local developments and ODB assessments, the open data ecosystem in Kenya, while dynamic, is still ‘emerging’ and open data’s utility as a monitoring mechanism remains limited.

7.5. Examination of Land Data

⁶⁵² Map Kibera, *Methods*, <http://mapkibera.org/work/methods/> [accessed 25 May 2018].

⁶⁵³ Open Institute, *Chief Kariuki presenting Lanet Umoja's data experiences* (29 September 2016) <https://www.youtube.com/watch?v=7EXkkYzUyLU> [accessed 25 May 2018].

⁶⁵⁴ Open Institute, *Chief Kariuki presenting Lanet Umoja's data experiences*.

⁶⁵⁵ Interview with Sifa Mawiyoo, Open Data Specialist and GIS Technologist, Kenya Open Data Initiative, ICT Authority; and Prestone Adie, Data Analyst, Kenya Open Data Initiative, ICT Authority, Nairobi, 20 September 2016; Interview with Sandra Musoga, Senior Programs Officer – Transparency, Article 19, Nairobi, Kenya, 16 September 2016.

⁶⁵⁶ Ibid.

⁶⁵⁷ Open Data Barometer, *Kenya*, https://opendatabarometer.org/data-explorer/?_year=2015&indicator=ODB&lang=en&open=KEN [accessed 20 September 2018].

⁶⁵⁸ Mutuku & Mahihu, *Open Data in Developing Countries*, p.5.

This case study will now examine a sample dataset from KODI with a view to determining whether controls for authenticity are in place. In view of the significance of land in Kenyan political and economic life (explained below), the study will look at an available dataset relating to land use. It will begin with an introduction to Kenya's land management regime and an overview of Kenyan government land information management, before considering the dataset itself.

7.5.1. Land information management

'...Kenya was among the first [independent African] countries to experience comprehensive land reform and has, in many ways, gone the farthest in promoting individual freehold...'⁶⁵⁹ nonetheless, access to information about land and related resources has been a recurring problem in Kenya. On 30 October 2015, the Thomas Reuters Foundation reported on the Kenya Groundwater Mapping Programme, a project to map Kenya's water resources for the benefit of farmers and others. Reuters reported:

One key problem is lack of data... According to the Kenya Water Industry Association, not one of the country's several water regulation agencies, including the Water Resources Management Authority, has reliable data that captures the distribution, quantity and quality of available groundwater.⁶⁶⁰

Following Yates, Vismann and Pasquinelli, this data problem becomes a governance problem. Since the colonial period, Kenya's land management system has generally functioned through the creation, transmission and exchange of paper records. Under the *Registration of Titles Act* (1982), a Central Registry was established in Nairobi and a Coastal Registry established in Mombasa for the management of paper records. Under the *Registered Land Act* (1989), a registry was established in every county (56 registries in total). The 2012 *Registration Act*, which aims to rationalise and devolve the land registration

⁶⁵⁹ Herbst, J., *States and Power in Africa: Comparative Lessons in Authority and Control*, (Princeton: Princeton University Press, 2000) p.185.

⁶⁶⁰ Waruru, M., 'To arm against drought, Kenya maps its water resources', *Thomas Reuters Foundation*, <http://news.trust.org/item/20180610090015-agcuf/> [accessed 25 May 2018].

process, has yet to take effect: regulations are in development. Darch and Underwood observed that ‘...as far as freedom of information is concerned, good record-keeping and archival practices – an essential pre-condition for compliance – are often lacking [on the African continent]...’⁶⁶¹ and, historically, government record-keeping in Kenya has experienced numerous challenges. These have included a lack of cohesive policies, lack of compliance with procedures, ad hoc systems, and lack of staffing and other resources.⁶⁶²

In 2010 and 2011, during our research into the readiness of Kenyan government record-keeping for e-government and Freedom of Information, Wamukoya and I found that what was then the Ministry of Lands (now the Ministry of Lands and Physical Planning) had experienced a period of record-keeping reform.⁶⁶³ At the time of our study, the registries were well-functioning and monitored, and sanctions were imposed for infringements:

Records management is not audited, but when breaches of records management procedure are identified, they are investigated. At the time of the interview a member of the registry staff was on suspension for removing a file that he was not permitted to access...⁶⁶⁴

At the time of that study, a digitisation project was underway, motivated by a sense that computerisation and digitisation would, amongst other benefits, reduce delays in work processes, which accounted for the majority of complaints received by the Ministry (approximately 90% of written complaints, according to one member of staff).⁶⁶⁵ The digitisation project was being led by the Land Management Systems Technical

⁶⁶¹ Darch & Underwood, *Freedom of Information and the Developing World*, p.206.

⁶⁶² Administrative histories of record-keeping in sub Saharan Africa are limited. For an overview of the problems that are common across many of those countries, see Wamukoya, J., ‘Records Management and Governance in Africa in the Digital Age’ and Mnjama, N., ‘Anne Thurston and Record-keeping Reform in Commonwealth Africa’, in Lowry, J. & Wamukoya, J. (eds.) *Integrity in Government through Records Management*. The roots of these problems are to be found in the colonial period. Again, this history is under researched, though the colonial origins of more recent record-keeping problems is noted by Musembi, M., ‘Development of Archive Services in East Africa’, in *Historical Development of Archival Services in Eastern and Southern Africa: Proceedings of the 9th Biennial General Conference* (Mbabane;Roma: ESCARBICA, 1986) p.116.

⁶⁶³ International Records Management Trust, *Kenya Country Report*, p.13.

⁶⁶⁴ Ibid.

⁶⁶⁵ Ibid.

Working Group in the Lands Reform Unit, which was leading on the Ministry's target under the government-wide Vision 2030 strategy: an improved land title acquisition process.⁶⁶⁶

Wamukoya and I found that 'Few members of staff use computers at present, and e-mail is not yet used for work purposes', indicating a working environment not yet ready for digital processes or digital information. We reported that records staff were confident that the organisation of paper records was adequate for land title process improvements – the digital systems and technologies were seen as the key to meeting the Vision 2030 target – but the members of the Working Group (none of whom were records staff) identified significant gaps in the paper records that were being digitised. Digital surrogates were comprehensively and regularly backed up, but there was no plan in place to develop digital preservation measures, and no consideration of the need to eventually transfer some records to the Kenya National Archives and Documentation Service (KNADS). There also seemed to be no planning for moving from the creation and scanning of paper records to the management of born-digital records.⁶⁶⁷

In September 2016, I visited the Ministry of Lands and Physical Planning again. The digitisation of paper records was still underway, including the repair and digitisation of maps. The process was conceived as an information technology issue, so there was no involvement of records officers in the planning of the project or its technical parameters. Nevertheless, the project has had some success: it is now possible for staff to conduct online searches for Nairobi, and the Ministry is setting up one central database of land titles as part of the Kenya National Spatial Data Infrastructure, which should expand the Ministry's capacity to conduct online searches.⁶⁶⁸

⁶⁶⁶ Ibid.

⁶⁶⁷ Ibid., p.5.

⁶⁶⁸ Interview with Edward Kosgei, Head of Lands Administration, and Emily Ndungi, Principal Records Management Officer, Lands Department, Ministry of Lands and Physical Planning, Nairobi, Kenya, 19 September 2016.

Record-keeping practices in the Ministry continue to be guided by public service guidance on records – there is no separate manual for the Ministry.⁶⁶⁹ Although less complex than the *Bengal Records Manual*, the *Records Management Procedures Manual for the Public Service* (the *Manual*) issued by the Office of the Prime Minister, Ministry of State for Public Service, in May 2010, sets out a system that provides for the registration and management of records. It makes provisions for mail management, filing, indexing, cross-referencing, classification, file tracking, ‘bring-up’, storage, survey, appraisal and disposal (transfer to archives or destruction). The *Manual* also discusses disaster management and capacity building (training and performance monitoring). Its final chapter explains the institutional framework for record-keeping, which includes ministerial responsibility for compliance.

The ‘Security of Records’ chapter discusses access to records. It states that ‘Access to classified records [Top Secret, Secret, Confidential, Restricted] is restricted to the “need to know” principle’, reflecting the long-standing civil service bias towards secrecy.⁶⁷⁰ However, it also notes that ‘Access to public records shall be provided within the existing legislative and regulatory framework’,⁶⁷¹ which now includes Freedom of Information legislation, and the *Manual* warns staff to ‘guard against a natural tendency to over classify documents’.⁶⁷²

The *Manual* includes a chapter on digital records management that sets out responsibilities, provides guidance on naming conventions, media handling and storage, and empowers KNADS to set procedures for digital records management and authorise the destruction of digital records. The chapter’s description of the components and characteristics of digital records is grounded in archival theory as informed by diplomatics, speaking of content, context and structure, and trustworthiness. The chapter states that all public institutions are required to install the Integrated Records Management System (IRMS). At the time of the *Aligning* project, the Personnel Office in the Ministry of Public Service

⁶⁶⁹ Ibid.

⁶⁷⁰ Republic of Kenya, *Records Management Procedures Manual for the Public Service*, (Nairobi, Kenya: Office of the Prime Minister, Ministry of State for Public Service, May 2010), p.50.

⁶⁷¹ Ibid., p.50.

⁶⁷² Ibid., p.51.

had just developed and introduced the system. Although it has not been audited against ICA-Req or similar standards, it has the core records system functionalities of registration, file tracking, 'bring up' and reporting.⁶⁷³ Although no provisions have been made for the preservation of born-digital records or digital surrogates, Kenya's current OGP National Action Plan includes a commitment to 'Enhance right to information by strengthening records management and access to information', which is being led by KNADS and requires KNADS to 'Establish a central digital repository to provide lasting access to government records and data and all information of public interest' by June 2018.⁶⁷⁴

7.5.2. Examination of land dataset

In order for KODI information to help to correct information asymmetry, as already argued, it is essential that its data can be shown to be authentic. Mutuku and Mahihu found that 'data quality' was a key issue in the performance of civic technologies, which can inherit faults from the data issued through the data portal. Mutuku and Mahihu's assessment of quality deals with the relevance of data to citizens, the regularity of data updates or dataset releases, and the utility of data, which concerns incomplete data and data that is poorly structured or formatted.⁶⁷⁵ While their conception of utility as an aspect of quality approaches the question of trustworthiness, inasmuch as it evokes the notion that datasets often must be manipulated and/or supplemented to be useful, it does not closely consider authenticity. The InterPARES work on trustworthiness in digital records shows that metadata, documented custodianship and auditable systems are prerequisites for establishing the authenticity of records. In order to examine the extent to which these three prerequisites are present in the production and publication of data released through KODI, it is necessary to examine a sample dataset. In October 2016, I selected the dataset 'Proportion of Parcels Using Fertiliser 2006' because it related to the contentious issue of land use and was accompanied by contextual information that suggested it might be possible to determine if the prerequisites for authenticity were evident in its production and

⁶⁷³ International Records Management Trust, *Kenya Country Report*, p.9.

⁶⁷⁴ Republic of Kenya, *OGP National Action Plan 2*.

⁶⁷⁵ Mutuku & Mahihu, *Open Data in Developing Countries*, p.47.

publication: the dataset was published in 2015, uploaded by 'Knoema', and its metadata identified the Kenya National Bureau of Statistics as the original data source.

Accessing KODI at opendata.go.ke on 11 January 2017, the site rendered a log-in interface (see *Figure 4*).

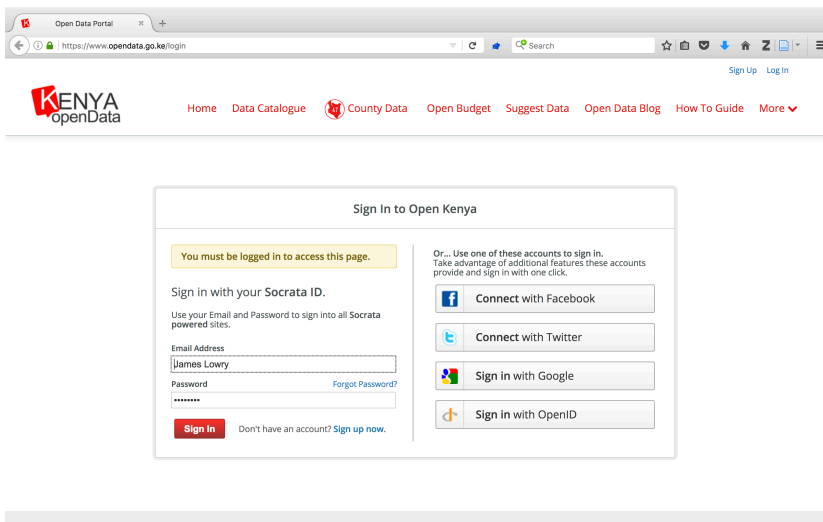


Figure 4. KODI interface on 11 January 2017.

Creating an account and signing in brought me to an error message (*Figure 5*) which indicated that permission was required to access the data (*Figure 6*).

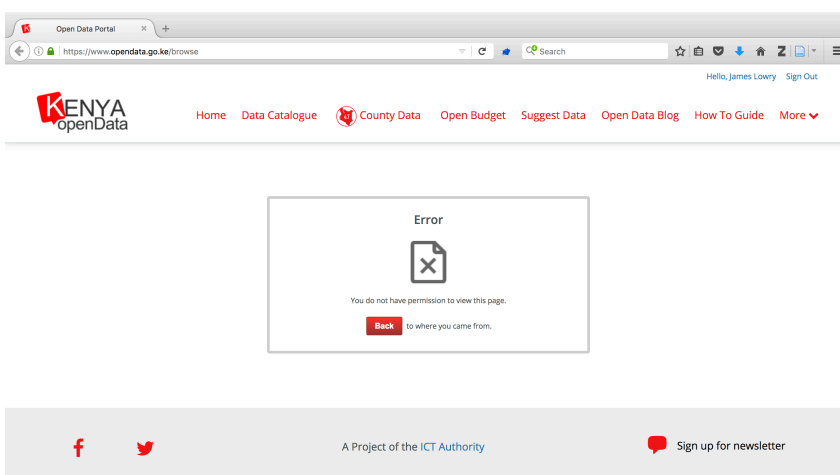


Figure 5. KODI sign in error message.

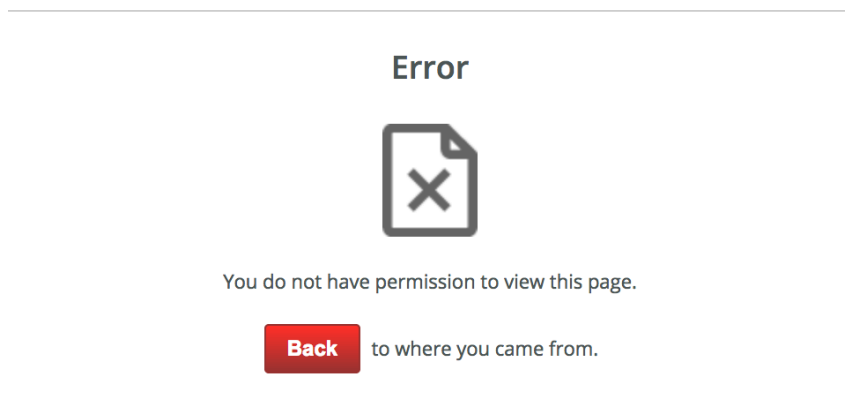


Figure 6. KODI sign in error message (close up).

On querying the staff of the ICT Authority, I was directed to use a mirror site while the portal was migrated between service providers. A search of the interim platform for the title of the dataset on 16 January 2017 produced no results. A search for 'parcels' produced a short list of results that included a dataset called 'Proportion of Parcels Using Fertiliser County Estimates 2005/6'. The metadata for the dataset indicated that the dataset was shared by 'kodipublisher' rather than 'Knoema', with a publication date of 20 December 2016, likely the date of the migration to the interim platform. There is nothing to indicate that the dataset was first published in 2015. Comparable datasets are now available for 2005 and 2007-2009.

The dataset, viewed as a CSV file, consists of seven columns: A – object ID (a sequence of ascending numbers from 1 to 47); B – 'county_nam' (a list of county names); C – 'proportion_of_parcels_using_f' (the same list of county names); D – 'proportion_of_parcels_using_1'; E - 'proportion_of_parcels_using_2'; F - 'proportion_of_parcels_using_3'; G - 'proportion_of_parcels_using_4'. Columns D to F provide figures ranging from 0 to 0.94. Column G provides figures ranging from 0 to 428581.8. There is no data or metadata within the CSV file to guide the interpretation of these figures, such as the meaning of 'using_1', 'using_2', etc., and no formulae that provide a key to the relationship between the figures in columns D through F and the figures in column G. Columns D through F for Wajir County each have a value of 0 and column G has a value of 0. Columns D through F for Mombasa County each have a value of 0, but column G has a value

of 2425.2. The content, context and structure of the dataset, in themselves, are insufficient for the interpretation of the dataset.

With regard to this particular dataset, KODI does the work of interpretation for its users. The portal includes three tabs: 'Overview', 'Data' and 'Visualization'.⁶⁷⁶

'Overview' provides basic metadata about the dataset; its title, publisher, last modified date (only), licensing information, and it attributes the dataset to the Kenya National Bureau of Statistics. Viewing the 'Data' tab produces a generic error message ('There was an error'). The 'Visualization' tab allows users to generate a map of Kenya's counties coloured to show the proportion of parcels using fertiliser. To the general user, an understanding of the dataset requires this visualisation feature.

The 'Proportion of Parcels Using Fertiliser County Estimates 2005/6' dataset demonstrates the absence of assurances of authenticity. In terms of metadata, the content of the CSV file itself lacks sufficient metadata to facilitate the interpretation of the data. This proves only that this particular dataset does not constitute a record (its structure does not tie together its data and metadata in a way that produces meaning); there have been no claims that the datasets available through KODI are records. Furthermore, that human interpretation of the CSV is not possible does not necessarily undermine an assumption of authenticity. However, since it is only when rendered as a visualisation by KODI that the data becomes interpretable, a question arises about the mechanisms of visualisation: what formulae and algorithms are being used to render the dataset? If machine-readable information can constitute a record when, amongst other things, it is rendered human-readable, the software that does that rendering work constitutes a significant component of the system for managing the information. The rendering function constitutes part of the system that the InterPARES work has shown must be auditable to support a presumption of authenticity. There is no technical information available through KODI that would allow users to

⁶⁷⁶ Kenya Open Data Initiative (KODI) < <http://www.opendata.go.ke> > [Accessed 21 September 2018].

understand the process that produces the map, which effectively constitutes a gap in the chain of custody.

Of more significance, for the purposes of ascertaining authenticity, is the lack of metadata available outside the file itself, through the platform (the system). Treating 'Proportion of Parcels Using Fertiliser County Estimates 2005/6' in isolation, it appears to have 'identity', which, as defined by the InterPARES ontology, 'refers to the distinguishing character of a record, that is, the attributes of a record that uniquely characterize it and distinguish it from other records'.⁶⁷⁷ This 'identity' is supplied by its metadata, in particular its unique title, author and dates. However treating the dataset in comparison with the 'Proportion of Parcels Using Fertiliser 2006' dataset available in October 2016, the discrepancy in publishers (Knoema and kodipublisher) and dates (uploaded 2015 and last modified 20 December 2016, without reference to an upload date) raises a question about the identity of the dataset. Is it the same dataset? If identity is called into question through the lack of version control, one of the two fundamental elements of authenticity (identity, with integrity) is absent.

In addition, the provenance of the dataset is obscured by the lack of metadata documenting the custodianship of the data from the point of collection. Working backwards, there is metadata about the publisher of the data (though some ambiguity about the identity of the publisher when the two versions of the dataset are considered together) and metadata about the source of the dataset (the Kenya Bureau of Statistics), but beyond this, it is not possible to see the sources used by the Bureau of Statistics. In this dataset, there are two types of data: 1) the definition of land parcels, which depends on information generated by the land registration process conducted by the Ministry of Lands and Physical Planning, and 2) the estimates of fertiliser use, which are likely to come from the Ministry of Agriculture, Livestock and Fisheries or one of its agencies. In order to have an assurance of the integrity of the dataset, it must be possible to know the sources of the data. In this case, such as assurance would require interaction

⁶⁷⁷ InterPARES 2, *Terminology Database*, http://www.interpares.org/ip2/ip2_terminology_db.cfm [last accessed 15 March 2019].

with the Bureau of Statistics. My research has shown that the record-keeping system of the Ministry of Lands and Physical Planning is well-functioning and conforms with the system set out in the *Manual*,⁶⁷⁸ which, being government-wide, is likely to also guide the record-keeping of the Ministry of Agriculture, Livestock and Fisheries. This record-keeping system documents custodianship throughout the lifecycle of the Ministries' records, though this feature of the system, and the particular custodians of particular information, is likely to be unknown by KODI users. The break in the documentation of the chain of custody occurs between the creation or capture of the data by the Ministries and the aggregation of the data by the Bureau of Statistics. What processes and controls are used by the Bureau to manage data in such a way that responsibility for the integrity of the data is assigned and documented?

Regarding the InterPARES requirement for an auditable system, we see in the publication of a dataset that there are a series of systems through which the data passes. At the point of creation and throughout use within the Ministry, land parcel information is subject to an auditable system. The data may then pass to the Ministry of Agriculture, Livestock and Fisheries, where it serves as a parameter for the ministry's estimates of fertiliser use. No maps of this data flow are available to KODI users. At the point of aggregation, the data passes into the systems of the Bureau of Statistics, which are not transparent to KODI users. In preparation for publication, the data passes into the custody and systems of KODI. This transition of custody, unlike the previous transitions, is documented in the publicly available metadata. As outlined above, KODI staff have implemented certain measures for ensuring the integrity of data, and, while they are not brought together in a formal system that allows each action and custodian to be audited, they do facilitate the auditing of the data itself. There is no information about these controls on the KODI portal. Rather than assurances of integrity, KODI users have ambiguous information about the sources and treatment of the data.

⁶⁷⁸ International Records Management Trust, *Kenya Country Report*, pp.12-13.

Though ‘Proportion of Parcels Using Fertiliser County Estimates 2005/6’ may be authentic in the sense that it ‘is what it purports to be’, it is not possible to ascertain its authenticity. Doubts about its identity and integrity are introduced by its partial metadata, opaque provenance, undocumented custody, particularly during aggregation, and the lack of available information about the systems for its management. This assessment of the dataset has also shown that KODI itself should be considered part of the system insofar as the interpretation of the dataset requires the visualisation performed by KODI: as part of the system, its methods for rendering the data human-readable must be open to audit if the identity and integrity of the data as rendered are to be established.

7.6. Conclusion

Between 2006 and 2016, Kenyan government experienced a rapid transition from an ingrained culture of secrecy that had evolved from the colonial legacy of the ‘need to know’ stance on access to information and the imperative to conceal corruption; a culture codified in restrictive legislation such as the *Official Secrets Act* (No. 11 of 1968). That transition from secrecy began with Ndemo’s attempt to expose the unfair distribution of CDF funds, which was thwarted by Members of Parliament, and the efforts of local and international civil society actors, most notably the efforts of the Kenyan Section of the International Commission of Jurists and Article 19, to push Freedom of Information. In 2010, with no progress on the Bill, but a victory in the form of clause 35 of the new national Constitution, openness advocates increasingly looked to open data as a means of opening government. Ndemo’s success in securing presidential support for an open data portal prompted the rapid development of KODI, which published problematic data. The enthusiasm for open data was reflected in Kenya’s first OGP action plan, which called for the publication of a range of datasets. The problems with KODI started to become apparent to observers in 2013, and similar problems were identified by the Data for African Development Working Group as a trend across the African continent. This study has shown that, in the Kenyan case, these early problems resulted from the rapid development of KODI, the political upheaval of the devolvement of powers to counties and the transition between governments. The Open Data Barometer Regional Report for Africa (3rd Edition,

2016) describes Kenyan open data as being in a holding pattern, but signs of ongoing commitment to the open data project are evident in the literature and my data. In 2016, a new OGP action plan reiterated the call for the release of more datasets and the Access to Information Act received presidential assent. It is too early to gauge the effects of the Access to Information Act on the participation of public bodies in the open data programme, but my research has shown that although open data was initially identified as a means of bypassing the legislative route to openness, the common view amongst Kenyan civil society and public sector stakeholders is that the lack of a legislative basis for data release has hampered the open data initiative.

The transition towards openness in Kenya has also been characterised by a change in the relationship between government and civil society. My research with Wamukoya in 2010 revealed an adversarial relationship between civil society organisations that were pushing for openness and a resistant public sector. In 2016, both sides were engaged in the openness project, both acknowledged the challenges still to be addressed, and a symbiotic relationship was evident in the role being played by civil society in aiding the government's implementation of the Access to Information Act and enriching the open data ecosystem. The caveat here is that my data reflects the position of a limited number of government officials, most engaged in the technical work of opening data. The true extent of the shift within the public sector – the kind of work undertaken by van Schalkwyk, Willmers and Schonwetter - is a study that remains to be done. Outside of this relationship, a vibrant open data culture exists within civil society, evident in the numerous community-led data projects that are effecting social and infrastructural changes across Kenya. These have been made possible by education and technological investment, such as the spread of fibre-optic cabling and the growth in the use of mobile technology. In hindsight, it is clear that the CDF distribution map was premature, and that political and technological developments were necessary to create an environment in which an open data initiative might survive.

There are problems with the data that is being released through KODI, so that the potential for agent monitoring remains limited. Many of these problems are known

to KODI and recognised by civil society. The ODB data visualisations make it clear that open data is still having a low impact – KODI continue to solicit different datasets from government, while Article 19 is educating citizens on their data rights, which is likely to increase demand, and therefore impact. On a more technical level, there are problems of work flow that include frequency of release, which is tied to the embeddedness of openness culture. KODI is addressing this through its network of fellows.

My study has charted the lifecycle of a dataset and found that an assurance of authenticity may be expected from the record-keeping practices of public sector bodies, though this would be enhanced by the publication of record-keeping audit reports. This concerns the beginning of the data lifecycle. Later in that lifecycle, when data is published on KODI, authenticity is called into question by partial metadata about data's provenance and management. Leaving aside the problems that may be unique to an examination of datasets during KODI's transition between platforms, there are unanswered questions about the management of the data throughout its lifecycle.

KODI is taking steps to build controls into its processes. Through its snapshots of the portal with every new upload, its retention of original datasets as received from government and its normalisation of forms through the introduction of templates for datasets, KODI is taking steps to build an audit trail for the accuracy and completeness of data and metadata. If these could be brought together in a systematic way, documented and published, KODI will have moved towards an assurance of authenticity. KODI's visualisation of the dataset is necessary for the dataset's interpretation by most users, which makes KODI a component of the system for the management of the data. Systems must be transparent – custodianship must be done *in public*; at present, KODI's mechanisms for visualisation are not published. The stage of the data curation lifecycle that most jeopardises a presumption of authenticity is during aggregation. The data aggregation practices of the Bureau of Statistics are not publicly documented. It is in this space between creation and publication that provenance is obscured, and custody is not auditable. The chain of custody appears to be broken in several places.

Chapter 8

Case Study: data.gov.au

8.1. Introduction

This chapter reports the findings of an assessment of controls for authenticity in open data in Australia, at the federal government level. Like the other case study in this thesis, this study focuses on a Commonwealth country because of the possibility of comparisons across inherited cultures of secrecy, record-keeping practices and legal traditions. The case study begins with an historical overview of the development of openness in Australian government, focusing on the federal level. The case study then describes the current state of Australia's open data initiative and the country's involvement with the Open Government Partnership. The case study closely examines a sample dataset prepared by the Australian Bureau of Statistics and published on the federal government's data portal, data.gov.au. The dataset relates to land management practices, as in the Kenyan case study. The study of the dataset is set in the context of the Australian land management regime. It notes that there are differences between Kenya and Australia with regard to land registration practices, related to the level of government at which registration occurs, the method of land registration and the record-keeping infrastructure in place for land registration documentation. The case study provides a point of comparison with the Kenyan case study, enabling observations about the place of authenticity controls in the redress of information asymmetry in the principal-agent problem under the social contract. That comparison is made in Chapter 9.

8.2. Historical Development of Open Data in Australia

8.2.1. *Secrecy in Australian Government*

Judith Bannister et al have asserted that 'Australian public administration inherited a culture of secrecy from England, which was reinforced by official secrets legislation'.⁶⁷⁹ Australia's first secrecy legislation was introduced in the colony of Victoria in 1867, and the first national secrecy legislation was passed by

⁶⁷⁹ Bannister, J., et al., *Government Accountability: Australian Administrative Law* (Cambridge: Cambridge University Press, 2015) p.205.

the first session of the parliament of the newly federated nation in 1901.⁶⁸⁰ It is therefore possible to say that state secrecy is as old as the state in Australia. Subsequent commentators on Australian government have observed that this secrecy is implicit in the Westminster system of government. For example, a 2008 Independent Review Panel report on FOI legislation in the state of Queensland stated:

Secrecy had been an essential ingredient of the system—secrecy to protect the deliberations of the cabinet, secrecy to protect the advice proffered by public servants to their ministers, secrecy to hide what happened within the public service. The democratic element that allowed this closed system to function was provided by the concept of ministerial responsibility—ministers were responsible, collectively and individually, directly to parliament and indirectly to the electorate, for what the government did, and for what their departments did.⁶⁸¹

In 1914, the national parliament passed the *Crimes Act*, which, as the Honourable Justice Susan Kenny shows, demonstrates the cultural and legal inheritance Australia received from the United Kingdom:

Section 70 of the *Crimes Act* originated with s 86 of the Queensland Criminal Code of 1889. Section 79(3) derived from s 2 of the UK's *Official Secrets Act 1911*. Both provisions prohibit the disclosure of Commonwealth information acquired by a person in the course of his or her duties. They remain two of the most significant secrecy provisions in Australian law.⁶⁸²

The Australian Law Reform Commission (ALRC) report on *Secrecy Laws and Open Government in Australia*, issued in 2009 in response to the Attorney-

⁶⁸⁰ Australian Government: Australian Law Reform Commission (ALRC), *Secrecy Laws and Open Government in Australia: Report 112* (December 2009) p.43.

<https://www.alrc.gov.au/sites/default/files/pdfs/publications/ALRC112.pdf> [accessed 18 July 2018]

⁶⁸¹ Freedom of Information Independent Review Panel (2008) quoted in Australian Law Reform Commission, *Secrecy Laws and Open Government in Australia* (ALRC Report 112)

https://www.alrc.gov.au/publications/secracy-open-government#_ftn3 [accessed 24 June 2017].

⁶⁸² Kenny, J. C., *Secrecy Provisions: Policy and Practice*, (2011)

<http://www.fedcourt.gov.au/digital-law-library/judges-speeches/justice-kenny/kenny-j-20110324> [accessed 21 September 2018].

General's 2008 request for an inquiry into the safeguarding of government information, provides a brief history of secrecy in Australian government.⁶⁸³ The report states that further secrecy measures were introduced by or into national legislation following the Second World War and during the Cold War:

Periods of international conflict have precipitated an awareness of the need for, and experience of, secrecy provisions. For example, World War II and the Cold War 'provided a setting where secrecy was linked to military strength'. In 1960, amendments were made to s 70 of the Crimes Act 1914 (Cth), inspired in part by the anti-communist climate of the Cold War. The amendment, which extended the reach of s 70 to former Commonwealth officers, was 'just one of many secrecy provisions inserted or strengthened in legislation after the war'.⁶⁸⁴

This link between secrecy and the protection of the state is also seen in the UK's introduction of secrecy legislation following the German gunboat incident.⁶⁸⁵

8.2.2. Access to Information: Open Data and Freedom of Information

The ALRC report notes that interest in and calls for FOI legislation became evident in Australian political discourse at the national and state levels through the 1960s and 1970s:

In 1970, the then Leader of the Opposition, the Hon Gough Whitlam MP, noted with concern that 'excessive secrecy has become commonplace in governmental decision making'. Introduction of FOI legislation became an issue in the lead up to the 1972 federal election, at which time the Australian Labor Party claimed that the government's monopoly of knowledge had 'led to bad decisions and bad government'.⁶⁸⁶

⁶⁸³ Australian Law Reform Commission, *Report 112*, pp.41-6.

⁶⁸⁴ *Ibid.*, p.43.

⁶⁸⁵ Hooper, D., *Official Secrets: The Use and Abuse of the Act* (London: Secker and Warburg, 1987) p.28.

⁶⁸⁶ Australian Law Reform Commission, *Report 112*, pp.44-5.

Through the 1970s, openness measures were introduced in various Acts of Parliament, including the *Ombudsman Act 1976*, the *Administrative Appeals Tribunal Act 1975* and the *Administrative Decisions (Judicial Review) Act 1977*, before the *Freedom of Information Act* was passed in 1982 – the first FOI legislation in a Westminster-style government.⁶⁸⁷ Kenny asserts that the Act had its origins in the Whitlam government's 1972 policy commitment; as in the UK, FOI was the result of Labour policy.⁶⁸⁸

On 22 December 2009, the Australian government's Government 2.0 Taskforce submitted its report, *Engage: Getting on with Government 2.0*.⁶⁸⁹ Government 2.0 is conceived of as the use of tools and approaches from Web 2.0 in government, with a view to increasing collaboration, participation and openness.⁶⁹⁰ The political project of Government 2.0 is elucidated in the report: 'Though it involves new technology, Government 2.0 is really about a new approach to organising and governing. It will draw people into a closer and more collaborative relationship with their government.'⁶⁹¹ The report acknowledges that Government 2.0 is not only about the use of new technologies, but requires a culture change in the public sector, one that moves from secrecy to openness.⁶⁹² The treatment of information is central in this culture change:

Information collected by or for the public sector is a national resource which should be managed for public purposes. That means that we should reverse the current presumption that it is secret unless there are good reasons for release and presume instead that it should be freely available for anyone to use and transform unless there are compelling privacy, confidentiality or security considerations which require otherwise.⁶⁹³

⁶⁸⁷ Ibid., p.45; Kenny, *Secrecy Provisions: Policy and Practice*.

⁶⁸⁸ Kenny, *Secrecy Provisions: Policy and Practice*.

⁶⁸⁹ Tanner, L., 'Declaration of Open Government', *Australian Government Department of Finance* (2010) <http://www.finance.gov.au/blog/2010/07/16/declaration-open-government> [accessed 19 July 2018].

⁶⁹⁰ Government 2.0 Taskforce, *Engage Getting on with Government 2.0* (2009) p.x. <https://www.finance.gov.au/sites/default/files/Government20TaskforceReport.pdf?v=1> [accessed 19 July 2018].

⁶⁹¹ Ibid.

⁶⁹² Ibid.

⁶⁹³ Ibid.

Indeed, open access to public sector information was identified as one of three pillars of the Government 2.0 agenda, with culture change and the application of Web 2.0 technologies to government business.⁶⁹⁴ The *Engage* report called on the government to make a Declaration of Open Government that would articulate the connection between Web 2.0 approaches and tools, access to public sector information as a national resource, and positioning Government 2.0 as a component part of the government's wider reform agenda.⁶⁹⁵ The *Engage* report was highly specific about the character of 'open' information, stating:

This means that unless there are good reasons to the contrary, information should be:

- Free (10)
- easily discoverable
- based on open standards and therefore machine-readable (11)
- properly documented and therefore understandable (12)
- licensed to permit free reuse and transformation by others.⁶⁹⁶

Footnote 11 states 'The Semantic web involves a vision of a machine-readable web, where intelligent agents would be capable of understanding data presented online by interpreting the accompanying metadata' and footnote 12 states 'Supported by metadata that will aid in the understanding the quality and interpretability of the information.'⁶⁹⁷ These footnotes demonstrate an awareness of the technical requirements for open information, and in particular the significance of metadata for the contextualisation and interpretation of data.

In March 2010, further calls for increased open government came when the Advisory Group on Reform of Australian Government Administration issued its report *Ahead of the Game: Blueprint for the Reform of Australian Government Administration*.⁶⁹⁸ The report stated:

⁶⁹⁴ Ibid., p.xi.

⁶⁹⁵ Government 2.0 Taskforce, *Engage Getting on with Government 2.0*, p.xiii.

⁶⁹⁶ Ibid., pp.xiv-xv.

⁶⁹⁷ Ibid.

⁶⁹⁸ Advisory Group on Reform of Australian Government Administration, *Ahead of the Game: Blueprint for the Reform of Australian Government Administration* (2010) <http://apo.org.au/system/files/20863/apo-nid20863-24401.pdf> [accessed 19 July 2018].

An important component of open government is enabling citizens to collaborate on policy and service design... Advances in information technology are making a stronger relationship between citizens and government possible. Today it is often more convenient for citizens to use online mechanisms to communicate their views to government. The Blueprint recommends that the Australian Government become more open and that public sector data be more widely available, consistent with privacy and secrecy laws.⁶⁹⁹

Although the link between policy and service design, online engagement and the openness of public sector data is not explicated in this statement, the connection is evidently recognised by the Advisory Group.

On 16 July 2010, the Hon. Lindsay Tanner MP, Minister for Finance and Deregulation in Julia Gillard's Labor Party government, issued a press release announcing the Declaration on Open Government.⁷⁰⁰ The document begins with a clear statement of the role of data in promoting civic participation, and technology as a key component of information reuse:

The Australian Government now declares that, in order to promote greater participation in Australia's democracy, it is committed to open government based on a culture of engagement, built on better access to and use of government held information, and sustained by the innovative use of technology.⁷⁰¹

In the press release, Tanner is quoted as saying: 'The Declaration is about making government information available to the public online and encouraging reuse of that information in new, valuable and potentially unexpected ways. This

⁶⁹⁹ Ibid., p.ix.

⁷⁰⁰ The Office of the Hon. Lindsay Tanner MP, *Media Release 41/2010: Declaration of Open Government* (2010)

http://parlinfo.aph.gov.au/parlInfo/download/media/pressrel/AKCX6/upload_binary/akcx60.pdf;fileType=application%2Fpdf#search=%22media/pressrel/AKCX6%22 [accessed 19 July 2018]. The OAIC website states that the Declaration on Open Government was made in July 2012:

<https://www.oaic.gov.au/information-policy/about-information-policy> [accessed 19 July 2018].

⁷⁰¹ Tanner, 'Declaration of Open Government'.

is very much in line with our Government 2.0 agenda.⁷⁰² It is at this point, then, that we can see the overt connection between information, technology and government in the Australian government's conceptualisation of open government.

The Declaration outlines some 'supporting initiatives', including reforms to the *Freedom of Information Act*, the establishment of the Office of the Australian Information Commissioner and the requirement for a government-wide Information Publication Scheme.⁷⁰³ Tanner is quoted as saying: 'Through the establishment of the Office of the Australian Information Commissioner and the Government's broader freedom of information reforms we are restoring trust and integrity in government',⁷⁰⁴ echoing the idea of the link between trust in public institutions and the openness of information familiar from the FOI and OGP rhetoric.

The Freedom of Information reforms cited by Tanner were given effect in the Freedom of Information (Removal of Conclusive Certificates and Other Measures) Act 2009 *and the* Freedom of Information (Reform) Act 2010. The 2009 reform abolished the power of government ministers to issue 'conclusive certificates', a mechanism enshrined in the Freedom of Information Act 1982 and the Archives Act 1983 that were used to exempt records indefinitely.⁷⁰⁵ The 2010 reform required government agencies to institute 'contractual measures' to ensure that they received copies of documents from contractors and subcontractors, effectively bringing outsourced public services into the FOI regime.⁷⁰⁶ The 2010 reform also saw the introduction of the Freedom of Information (Fees and Charges) Amendment Regulations 2010 (No. 1), which removed fees for applications and internal reviews of decisions, made requests for the data subject's personal data free of charge, stipulated that the first five hours of decision-making in relation to other FOI requests would be free of

⁷⁰² Ibid.

⁷⁰³ Ibid.

⁷⁰⁴ The Office of the Hon. Lindsay Tanner MP, *Media Release*.

⁷⁰⁵ Office of the Australian Information Commissioner, *FOI Guide*, <<https://www.oaic.gov.au/freedom-of-information/foi-resources/foi-guide/part-a-freedom-of-information-past-and-present>> [last accessed 15 March 2019].

⁷⁰⁶ Ibid.

charge, and waived fees in cases where agencies failed to meet stipulated timeframes for decisions.⁷⁰⁷

Immediately following the passage of the FOI legislation, the *Australian Information Commissioner Act 2010 (No. 52 of 2010)* was passed. As a result of the Act, the Office of the Australian Information Commissioner (OAIC) was established as an independent statutory agency within the Attorney General's portfolio.⁷⁰⁸ The OAIC took over the duties of the Privacy Commissioner, which was established in 1989 by the *Privacy Act 1988*.⁷⁰⁹ It has responsibilities under freedom of information and privacy legislation, and for advising the government on information policy: 'This requires the Information Commissioner to report to the Attorney-General on how public sector information is collected, used, disclosed, administered, stored and accessed.'⁷¹⁰ These responsibilities clearly concern the specific management arrangements for public sector information, bringing the responsibilities of the Information Commissioner into close alignment with the record-keeping function. The OAIC has developed a number of tools and policy documents to support the opening of public sector information, most notably the De-identification Decision-Making Framework (DDF), developed with the Commonwealth Scientific and Industrial Research Organisation (CSIRO) from the UK Anonymisation Networks' *Anonymisation Decision-Making Framework*.⁷¹¹ The DDF comprises ten components:

1. Describe your data situation
2. Understand your legal responsibilities
3. Know your data
4. Understand the use case

⁷⁰⁷ Australian Government, *Freedom of Information (Fees and Charges) Amendment Regulations (2010) No. 1*. <https://www.legislation.gov.au/Details/F2010L02781> [accessed 21 September 2018].

⁷⁰⁸ Australian Government, Office of the Australian Information Commissioner, *About Us* <https://www.oaic.gov.au/about-us/> [accessed 19 July 2018].

⁷⁰⁹ Australian Government, Office of the Australian Information Commissioner, *History of the Privacy Act* <https://www.oaic.gov.au/about-us/who-we-are/history-of-the-privacy-act> [accessed 19 July 2018].

⁷¹⁰ Australian Government, Office of the Australian Information Commissioner, *Australian Information Commissioner Act* <https://www.oaic.gov.au/about-us/who-we-are/australian-information-commissioner-act> [accessed 19 July 2018].

⁷¹¹ Australian Government, Office of the Australian Information Commissioner, *De-identification Decision-Making Framework* <https://www.oaic.gov.au/agencies-and-organisations/guides/de-identification-decision-making-framework> [accessed 24 July 2018].

5. Meet your ethical obligations
6. Identify the processes you will need to assess disclosure risk
7. Identify the disclosure control processes that are relevant to your data situation
8. Identify who your stakeholders are and plan how you will communicate
9. Plan what happens next once you have shared or released the data
10. Plan what you will do if things go wrong

Guidance under components 5 and 10 demonstrate a recognition of the value of documenting data management. Users are advised, under component 5, to 'Establish a comprehensive record-keeping system across all your operational activities, related to your privacy policies and procedures, to ensure there is a clear audit trail'.⁷¹² Under component 10, users are advised to ensure they have robust audit trails, because:

Being able to provide a clear audit trail which takes into account all relevant de-identification activities and processes will be crucial for the purposes of:

1. demonstrating that you have followed all correct procedures, and
2. identifying where, if at all, in your processing activities you might need to make changes to prevent a similar occurrence.

In practice this means keeping clear and up-to-date records of all your processing activities, detailing who did what, when and how. Some of this information can itself increase disclosure risk and thus these records may by default be internally facing. Not being transparent about the de-identification process may, however, impact on utility and for this reason you may wish to provide a top level public narrative about your de-identification processes.⁷¹³

This guidance shows a close consideration of the need to document adherence to procedures around openness for the sake of compliance as well as the value

⁷¹² O'Keefe, C. M., et al. *The De-Identification Decision-Making Framework: CSIRO Reports EP173122 and EP175702* (2017) <https://publications.csiro.au/rpr/download?pid=csiro:EP173122&dsid=DS2> [accessed 24 July 2018].

⁷¹³ O'Keefe et al., *De-Identification*, p.59.

for doing so in relation to demonstrating transparency about data management practices. The guidance notes the tension between transparency about these practices and the related risk of improper information release: striking an appropriate balance between openness and privacy is therefore not only a consideration in the review of information but, at the metatextual level, in the framing and transparency of data handling practices.

The Information Publication Scheme (IPS) mentioned in Tanner's press release was established by the FOI legislation of 2010. The IPS requires government agencies to proactively publish certain information, including '...the agencies' structure, functions, appointments, annual reports, consultation arrangements, and details of the agency's freedom of information (FOI) officer. Information routinely released through FOI requests and routinely provided to parliament must also be published online.'⁷¹⁴ The IPS effectively helps to move the Australian public sector towards a default position of openness, while reducing the need for FOI requests for certain classes of data. This shift away from reactive disclosure towards proactive disclosure is also seen in the developments in open data in Australia.

The Australian federal government's open data portal (data.gov.au) was established following the government's Declaration of Open Government and in line with the recommendation of the *Engage* report, which called for 'a central portal (data.gov.au) that will enable access to and discovery of the data and skills necessary in preparing government information to be released as open PSI [public sector information]'.⁷¹⁵ [Data.gov.au](http://data.gov.au) is managed by the Public Data Branch of the Department of the Prime Minister and Cabinet, the same branch that manages the whole of government data policy.⁷¹⁶ This central positioning within the structure of government suggests not only a desire to manage the portal as an extension of the policy in a harmonised way, but the significance of the open data initiative as a core governance issue. This positioning responds to a

⁷¹⁴ Australian Government, Office of the Australian Information Commissioner, *Our Information Publication Scheme* <https://www.oaic.gov.au/about-us/access-our-information/our-information-publication-scheme/> [accessed 24 July 2018].

⁷¹⁵ Government 2.0 Taskforce, *Engage Getting on with Government 2.0*, p.xv.

⁷¹⁶ Government of Australia, <https://data.gov.au/page/about> [last accessed 12 July 2017].

recommendation in the *Engage* report that ‘a lead agency be appointed from within one of the central portfolios — either within Finance and Deregulation or Prime Minister and Cabinet — to take responsibility for Government 2.0 policy and provide leadership, guidance and support to agencies and public servants.’⁷¹⁷

In May 2011, the OAIC published its *Principles on open public sector information*, which it stated would be applied in its monitoring of FOI compliance, but not establishing any legal obligations on agencies.⁷¹⁸ The *Principles* reiterate the presumption of openness, the need for clear reuse rights, appropriate charges for access and transparent complaints procedures. Community engagement is prioritised throughout the document. The *Principles* state that, in complying with government policy on the need to engage the community in policy formulation, agencies should consult the community on what information should be published, and on ‘agency publication practices’, which could be interpreted to mean the formats and frequency of releases, amongst other things.⁷¹⁹

The *Principles* demonstrate a high level of awareness within the Australian government of the need for openness to be underpinned by technical measures that facilitate information retrieval and access, as well as interpretation. Principle 3 positions information governance as a significant component of an openness regime, calling for the appointment of a senior executive ‘information champion’ with responsibilities including ‘managing agency information to ensure its integrity, security and accessibility’.⁷²⁰ Principle 4, ‘Robust information asset management’, calls on agencies to ‘identify the custodian of each information holding and the responsibilities of that officer’ and ‘establish clear procedures and lines of authority for decisions on information publication and release’; this guidance echoes the traditional record-keeping controls that require defined custodianship.⁷²¹ Principle 4 also advises agencies to ‘document known

⁷¹⁷ Government 2.0 Taskforce, *Engage Getting on with Government 2.0*, p.xiii.

⁷¹⁸ Australian Government, Office of the Australian Information Commissioner, *Principles on Open Public Sector Information* <https://www.oaic.gov.au/resources/information-policy/information-policy-resources/principles-on-open-public-sector-information.pdf> [accessed 24 July 2018].

⁷¹⁹ *Ibid.*, p.1.

⁷²⁰ *Ibid.*, p.2.

⁷²¹ *Ibid.*, p.2; Principle 4 also advises agencies to ‘decide if information should be prepared for publication at the time it is created and the form of publication’, which is in keeping with current practice in digital records management, where appraisal is done before creation.

limitations on data quality’, which underpins a presumption of trustworthiness.⁷²² Finally, it advises agencies to ‘preserve information for an appropriate period of time based on sound archival practices’.⁷²³ Principle 5, ‘Discoverable and useable information’, states ‘attach high quality metadata to information so that it can be easily located and linked to similar information using web search applications’.⁷²⁴ Taken together, these aspects of the *Principles* demonstrate the government’s recognition of information management’s value to openness.

In 2013, Australia first signalled its intention to join the Open Government partnership in a letter from the Attorney General.⁷²⁵ In 2015, another letter, this time from Prime Minister Malcolm Turnbull, informed the OGP of Australia’s intention to finalise its membership of the OGP through the adoption of its first National Action Plan (NAP), covering the period 2016 to 2018.⁷²⁶ Australia’s first NAP included fifteen commitments in five broad categories: transparency and accountability in government (whistleblowing protection in the tax and corporate sectors, beneficial ownership transparency, extractive industries transparency, combating corporate crime); open data and digital transformation (releasing high-value datasets, building public trust around data sharing, digitally transform government service delivery); access to government information (update information management and access laws, understand the use of FOI, improve the discoverability and accessibility of government data); integrity in the public sector (improve confidence in the electoral system, establish a National Integrity Framework, open contracting); public participation and engagement (delivery of the NAP, enhance public participation in government decision making). Information management and access figure prominently in these commitments. For example, Commitment 2.1 (Release high-value datasets and enable data driven innovation) notes that non-sensitive Australian government data is open by default, but aims to ‘develop an open dialogue between government and non-government sectors to identify characteristics of high-value data and to stimulate

⁷²² Ibid.

⁷²³ Ibid.

⁷²⁴ Ibid.

⁷²⁵ Open Government Partnership, *Australian Letter of Intent to Join OGP*, <https://www.opengovpartnership.org/documents/australia-letter-of-intent-join-ogp> [accessed 29 September 2018].

⁷²⁶ Open Government Partnership, *Australian Letter of Intent to Join OGP*.

greater use and re use of public data in innovative ways.⁷²⁷ Commitment 3.1 (Information management and access laws for the 21st century) concerns a review and harmonisation of the legal framework for information access, including an examination of the effectiveness of the FOI law and the *Archives Act 1983* and the roles of the OAIC and the National Archives of Australia (NAA).⁷²⁸ Commitment 3.3 (Improve the discoverability and accessibility of government data and information) calls for the further development of data.gov.au and the increased digitisation of paper records in the NAA.⁷²⁹ Although Australia has not yet been assessed by an Independent Reporting Mechanism (IRM), its *Midterm Self-Assessment Report* of 2017 reported significant progress against its commitments.⁷³⁰

8.3. Current State of Australia's Openness Programme

Data.gov.au is operated within a cohesive regulatory environment. As the portal states, it is 'bound by the provisions of the Privacy Act 1988 (Cth), including the Australian Privacy Principles' and the 'Guidelines for Federal and ACT government websites issued by the Office of the Australian Information Commissioner.'⁷³¹ Information is also often released free of charge and under open licence. As the *Engage* report states:

In 2001, the Australian Government's Spatial Data Access and Pricing Policy was one of the first substantial programs in the world in which government

⁷²⁷ Open Government Partnership Australia, *Australia's first Open Government National Action Plan: Commitment 2.1 - Release high value datasets and enable data driven innovation* <https://ogpau.pmc.gov.au/australias-first-open-government-national-action-plan-2016-18/commitments/ii-open-data-and-digital-0> [accessed 24 July 2018].

⁷²⁸ Open Government Partnership Australia, *Australia's first Open Government National Action Plan: Commitment 3.1 – Information management and access laws for the 21st century* <https://ogpau.pmc.gov.au/australias-first-open-government-national-action-plan-2016-18/commitments/iii-access-government-0> [accessed 24 July 2018].

⁷²⁹ Open Government Partnership Australia, *Australia's first Open Government National Action Plan: Commitment 3.3 – Improve the discoverability and accessibility of government data and information* <https://ogpau.pmc.gov.au/australias-first-open-government-national-action-plan-2016-18/commitments/iii-access-government-2> [accessed 24 July 2018].

⁷³⁰ Government of Australia, Department of the Prime Minister and Cabinet, *Midterm Self-Assessment Report for Australia's First Open Government National Action Plan 2016-2018* https://www.opengovpartnership.org/sites/default/files/Australia_Midterm-Self-Assessment_2016-2018.pdf [accessed 24 July 2018].

⁷³¹ Australian Government, *data.gov.au* <https://data.gov.au/about> [accessed 24 July 2018].

data which had previously been sold was made available without charge.

Today both the Australian Bureau of Statistics and Geoscience Australia are licensing much of their output using Creative Commons licences which permit others to freely use and remix it.⁷³²

The Australian government has stated that data.gov.au is a work in progress, acknowledging that ‘Improving the quantity and quality of the government data and the data.gov.au stack will be an ongoing process.’⁷³³ Nevertheless, there are signs that cultures of secrecy persist in the Australian federal government. The literature suggests that a culture of secrecy has persisted in particular parts of the government, perhaps most prominently in relation to the operation of migrant detention centres. Amy Nethery and Rosa Holman’s study of human rights abuses in offshore migrant detention centres cited the report of a Senate Committee investigation into conditions in a Nauru detention centre which found a ‘pervasive culture of secrecy’ in the Department of Immigration, exacerbated by the opaque reporting lines between the Department and the private sector ‘service providers’ who run the centres.⁷³⁴ Nethery and Holman argue that the culture has been enshrined in law in the form of the *Australian Border Force Act* of 2015: ‘Under the ABF Act, it is an offence to disclose “documents and information about the provision of services to persons who are not Australian citizens”. Staff working in the centres risk imprisonment for up to two years if they reveal information to the media concerning the detention centres.’⁷³⁵ In discussing the ethical issues facing healthcare professionals working in the centres, David Isaacs has observed limitations in the information shared with refugees on their own immigration cases, as well as the problems of restrictive contracts and gagging laws.⁷³⁶ Luke Heemsbergen and Angela Daly link the problem of government secrecy in this area with the surveillance of individuals,

⁷³² Government 2.0 Taskforce, *Engage Getting on with Government 2.0*, p.xii.

⁷³³ Australian Government, *data.gov.au* <https://data.gov.au/about> [accessed 24 July 2018].

⁷³⁴ Nethery, A. & Holman, R., ‘Secrecy and human rights abuse in Australia’s offshore immigration detention centres’, *The International Journal of Human Rights*, 20.7 (2016) pp.1025-1026.

⁷³⁵ *Ibid.*, p.1026.

⁷³⁶ Isaacs, D., ‘Are healthcare professionals working in Australia’s immigration detention centres condoning torture?’, *Journal of Medical Ethics*, 42 (2016).

and particularly refugees, citing ‘an acceleration of prohibitive privacy for government and prohibitive transparency for individuals’.⁷³⁷

In the same year as the *Australian Border Force Act* was introduced, the federal government passed the *Telecommunications (Interception and Access) Act*, which required telecommunications companies to retain metadata for two years. Benedetta Brevini has stated that this law would allow the government to trace journalists’ sources in the event of information leaks.⁷³⁸ Brevini situates the law in the context of, firstly, increased national security legislation following the September 11 attacks - ‘Over fifty laws were passed to create new criminal offences, new detention, extended investigative powers for security and police officers, new tools to control people’s movements and activities without criminal convictions’ – and secondly, legal reforms intended, she argues, to tighten information controls following the Snowden leaks.⁷³⁹

In view of this increased interest, and legislation, around information about particular government functions and activities (such as immigration) and telecommunications metadata relating to the communications of individuals, it is clear that information asymmetry continues to shift as government attempts to understand and control the informational landscape. The gains made through FOI reforms, publication schemes and open government data projects have been offset, to some extent, by further information asymmetries in particular areas of government operations relating to the surveillance of individuals and particularly those outside the demarcation of ‘citizen’. The current state of legislative provision for information control and access, in the context of its evolution since 2011, demonstrates the slippery nature of the principal-agent problem when the agent is large and functionally diverse. Government information continues to be a site of contestation.

8.4. The Role of Civil Society

⁷³⁷ Heemsbergen, L. & Daly, A., ‘Leaking Boats and Borders: The Virtue of Surveilling Australia’s Refugee Population’, *Surveillance & Society*, 15.3-4 (2017) p.389.

⁷³⁸ Brevini, B., ‘Metadata Laws, Journalism and Resistance in Australia’, *Media and Communication*, 5.1 (2017).

⁷³⁹ Brevini, ‘Metadata Laws’, p.77.

The history of this change is, at least in part, about changing relationships between citizens and government; re-evaluation of Westminster-style responsible government; a greater appreciation of participatory democracy; and the adoption of the language of human rights in the civil and political spheres.⁷⁴⁰

The literature is almost silent on the civil society role in the introduction of Freedom of Information legislation in Australia. Worthy points out that the ‘first attempt to map the process of passing FOI was Snell’s (2000) exploration of two early developers, Australia and New Zealand...’⁷⁴¹ and that Snell found that the push for these laws came from outside of government. Rhys Stubbs alludes to a campaign for FOI in the late 1960s but swiftly moves on to the first government action towards FOI in the 1970s.⁷⁴² Daniel Stewart’s descriptions of Australia’s FOI regime also starts here⁷⁴³ and Greg Terrill only notes that Australian historians made representations to the 1979 Senate review.⁷⁴⁴ Information on civil society engagement in the OGP process in Australia is similarly limited. The Australian Open Government Civil Society Network was established in December 2015 in order to facilitate the OGP NAP formulation process.⁷⁴⁵ Although a complete list of members is not available, there were eighty members (organisations and individuals) at March 2017.⁷⁴⁶ The membership of the Network’s Steering Group and founders shows that many of the groups involved in the UK’s civil society network have Australian counterparts working in the same way (Open Knowledge Australia, Transparency International Australia and

⁷⁴⁰ Kenny, *Secrecy Provisions: Policy and Practice*.

⁷⁴¹ Worthy, B., *The Politics of Freedom of Information: How and Why Governments Pass Laws that Threaten their Power* (Manchester: Manchester University Press, 2017) p.14.

⁷⁴² Stubbs, R., ‘Freedom of Information and Democracy in Australia and Beyond’, *Australian Journal of Political Science*, 43.4 (2008) p.671.

⁷⁴³ Stewart, D., ‘Assessing Access to Information in Australia: The impact of freedom of information laws on the scrutiny and operation of the Commonwealth Government’, in Wanna, J., Lindquist, E. A. & Marshall, P. (eds), *New Accountabilities, New Challenges* (Australia: ANU Press, 2015) p.79.

⁷⁴⁴ Terrill, G., ‘“A bit fast money, a bit white shoe brigade”? Freedom of Information and Australian History’, *Twentieth Century British History*, 12.2, (2001) p.233.

⁷⁴⁵ Australian Open Government Civil Society Network, *About*, <https://opengovernment.org.au/about/> [last accessed 15 March 2019].

⁷⁴⁶ Ibid.

Publish What You Pay Australia, Open Data Institute Australia).⁷⁴⁷ In a letter of 6 September 2016, Christopher Snow, an individual member of the Network, alleged that the Network demonstrated ‘oligarchical behavior [sic]’, an allegation that Australia’s OGP Support Unit dismissed after an investigation.⁷⁴⁸

Nevertheless, the lack of information about community-based, grassroots and individual members of the network suggest a network dominated by prominent non-state organisations. If there is or was limited principal engagement in the open government movement in Australia, this may suggest that there is motivation towards openness from the agent’s side, such as from within the bureaucratic supersubject, but it is not possible to confirm that on the basis of the available information.

8.5. Examination of Land Data

As a point of comparison with the KODI dataset, this case study will now examine the ‘Land Management Practices (2011 – 2012)’ dataset published on data.gov.au.⁷⁴⁹ As in the Kenyan case study, the purpose of this examination is to see whether or not controls for authenticity are in place. The assessment will begin with an overview of Australia’s land information management, before turning to an analysis of the dataset.

8.5.1. Land information management

In 1770, Lt James Cook of the British Royal Navy encountered the Gwiyagal people at what is now called Botany Bay, New South Wales. He pronounced the land to be *terra nullius* - nobody’s land – and from 1788 the British began to invade and settle *Terra Australis*. Over the next two hundred years, the white settlers, their heirs and subsequent immigrants to Australia would ignore indigenous sovereignty. In the Australian colonies, land registration was

⁷⁴⁷ Ibid.

⁷⁴⁸ Open Government Partnership, *OGP Response Policy Case Review and Resolution of Christopher Snow’s Letter of Concern, Dated September 6, 2016 Against the Government of Australia* https://www.opengovpartnership.org/sites/default/files/Australia_RP-Report_August2017.pdf [accessed 24 July 2018].

⁷⁴⁹ Australian Bureau of Statistics, *Land Management Practices (2011-12)*, <https://data.gov.au/data/dataset/land-management-practices> [last accessed 15 March 2019].

performed through the deeds registration system, which conferred priority on registered deeds over all other (i.e. non-registered) instruments.⁷⁵⁰ This system became unwieldy during the land speculation boom of the early 1880s.⁷⁵¹ In 1858, the Torrens title system of registration was introduced in South Australia, and would later spread to the other colonies.⁷⁵² Today, it is the method of land registration across all Australian states and territories, and it continues to be adopted internationally.⁷⁵³ The Torrens title system involves registering the names of land owners by a central registry, so that it is always possible to establish who is the owner of a piece of land through reference to the register.⁷⁵⁴ The registration system ties land parcels to owners through the use of unique land parcel numbers (volume-folio numbers) that correspond to files containing a description and plan of the parcel's boundaries and documentation of the owner's other rights and obligations such as mortgages and easements.⁷⁵⁵

The Aboriginal rights movement of the 1960s and 1970s provided the impetus for state and federal legislation that would begin to recognise indigenous land rights through the late 1970s and early 1980s.⁷⁵⁶ These laws established or recognised Aboriginal Land Councils and began the transfer of ownership over Aboriginal reserve lands and provided the Councils with opportunities to claim unowned and unleased land.⁷⁵⁷ In 1992, the Mabo case⁷⁵⁸ was heard by the High Court, which rejected the idea that Australia had been *terra nullius*, and recognised native title as a principle in Australian common law. Native title is a concept that 'reflects the entitlement of the indigenous inhabitants, in accordance with their laws or

⁷⁵⁰ Fox, P. M., 'The story behind the Torrens System,' *Australian Law Journal* 489 (1950).

⁷⁵¹ Ibid.

⁷⁵² Ibid.

⁷⁵³ Department of Planning, Transport and Infrastructure, Government of South Australia, *What is the Torrens Title System?* https://www.landservices.sa.gov.au/1public/Education_-_Careers/Documentary_Movies/4What_is_the_Torrens_Title_System.asp [last accessed 15 March 2019].

⁷⁵⁴ Fox, 'The story behind the Torrens System.'

⁷⁵⁵ Ibid.

⁷⁵⁶ Central Land Council, *The History of the Central Land Council*, <https://www.clc.org.au/articles/cat/our-history/> [last accessed 15 March 2019].

⁷⁵⁷ Central Land Council, *Functions and Responsibilities*, <https://www.clc.org.au/index.php?/articles/info/what-land-councils-must-do/> [last accessed 15 March 2019].

⁷⁵⁸ *Mabo v Queensland (No 2) ("Mabo case")* [1992] HCA 23 <http://www8.austlii.edu.au/cgi-bin/viewdoc/au/cases/cth/HCA/1992/23.html> [accessed 29 September 2018].

customs, to their traditional lands'.⁷⁵⁹ Although native title has been established in case law, official record-keeping has not responded to this shift, and Aboriginal knowledge and understanding of the land is not part of or represented in government land management systems or documentation.

The registration of land ownership is done at the state or territory level, increasingly through digital systems. For instance, the South Australian Integrated Land Information System (SAILIS) provides members (typically conveyancers, surveyors, valuers, banks and real estate agents) with integrated textual and spatial information, Certificates of Title, survey plans, Registered Dealings, Property Interest Reports, valuation and titling information and historical searches.⁷⁶⁰ Conveyancing across state borders is increasingly handled digitally. Currently, five states are using Property Exchange Australia (PEXA), an e-conveyancing platform that operates under a regulatory framework developed by the Australian Registrars' National Electronic Conveyancing Council (ARNECC).⁷⁶¹ PEXA issues digital certificates and it is possible for users to trade in paper certificates for digital certificates of title.⁷⁶²

PEXA is an interstate initiative, rather than an initiative of the federal government. At the federal government level, at least three departments deal with matters relating to land; the Department of Agriculture and Water Resources, the Department of Infrastructure and Regional Development and the Department of Environment and Energy. Record-keeping across these departments is likely to be strong given the strong regulatory environment and culture of record-keeping in Australian government. Within the record-keeping profession internationally, Australia has a reputation for excellence and innovation in records management and archival administration. The international standard for records management, ISO15489, is based on the Australian standard AS 4390-1996, the development of ICA-Req was led by the National Archives of Australia, an early leader in the

⁷⁵⁹ Ibid.

⁷⁶⁰ South Australian Integrated Land Information System, *Access to Land Property Information* http://www.sa.gov.au/_data/assets/pdf_file/0003/156567/Access-Land-and-Property-Information.PDF [accessed 24 July 2018].

⁷⁶¹ Property Exchange Australia, *About e-Conveyancing* <https://www.pexa.com.au/lawyers-conveyancers> [accessed 24 July 2018].

⁷⁶² Ibid.

EDRMS market was TRIM, a system developed by the Australian firm Tower Software, the Australian Series System was an innovation in archival arrangement, and the widely used continuum model of record-keeping was developed by archival theorists at Monash University in Melbourne.⁷⁶³ At federal and state levels of government, record-keeping is often supported by a cohesive and comprehensive regulatory framework, clear mandates established in law and policy, professional education at post-graduate levels and opportunities and resources for training and ongoing education provided by employers, universities and the private sector.

At the federal level, the National Archives of Australia (NAA) has statutory authority under the *Archives Act 1983* to set standards for government record-keeping.⁷⁶⁴ NAA has developed the Check-up PLUS platform, which allows government agencies to assess their information management practices against the requirements of the Digital Continuity 2020 Policy, a government policy designed to integrate information management practices into government work.⁷⁶⁵ NAA is leading on the policy implementation, and reports to parliament on the progress of government agencies.⁷⁶⁶ The policy includes ten recommended actions:

- Information governance reporting
- Agencies have established an Information Governance Committee
- Agencies have an information governance framework

⁷⁶³ The development of ISO15489 is discussed in Healy, S., 'ISO 15489 Records Management: its development and significance', *Records Management Journal*, Vol. 11 No. 2, (2001) pp.133-42; for information on ICA-Req see the International Council on Archives' ICA-Req pages at <https://www.ica.org/en/ica-reg-principles-and-functional-requirements-records-electronic-office-environments-guidelines-a-1> [last accessed 15 March 2019]; Tower Software's success with TRIM is evident in this Information and Data Manager blog post from 2006 <https://idm.net.au/blog/002207tower's-trim-does-it-again>. Two years later Tower Software was acquired by Hewlett Packard. For information on the Australian Series System, see Hurley, C. 'The Australian 'Series' System: An Exposition' in McKemmish, S. & Piggott, M. (eds.) *The Records Continuum: Ian Maclean and Australian Archives First Fifty Years* (Clayton: Ancora Press, 1994). The records continuum model is described in McKemmish, S., Upward, F. & Reed, B., 'Records Continuum Model', *Encyclopedia of Library and Information Sciences* (3rd ed.) (CRC Press, 2010) pp. 4447–8.

⁷⁶⁴ National Archives of Australia, *Our Organisation*, <http://www.naa.gov.au/about-us/organisation/> [last accessed 15 March 2019].

⁷⁶⁵ National Archives of Australia, Check-up PLUS, <http://www.naa.gov.au/information-management/check-up/> [last accessed 15 March 2019].

⁷⁶⁶ Ibid.

- Agencies manage their information assets for as long as they are required
- Agencies meet targets for skilled staff
- Agencies work digitally, with business interactions, decisions and authorisations recorded digitally
- Information in analogue formats is migrated to digital format, where there is value for business
- Information is managed based on format and metadata standards for information governance and interoperability
- All business systems meet functional requirements for information management
- Cross-agency and whole of government processes incorporate information governance requirements and specifications

These recommendations promote a move to more comprehensive records creation, standardised file types and metadata schemas and records management functionality across business systems. Each step in this direction will enable higher levels of control across metadata, documented custodianship and auditable systems.

8.5.2. Examination of land dataset

The ‘Land Management Practices (2011-2012)’ dataset published on data.gov.au is a rich dataset covering a range of land management topics, including fertiliser use. The landing page for the dataset includes a summary of the topics covered, as well as a metadata element set that includes the status of the data (active), the update frequency (infrequently), the date of publication (2016-02-18) and update (2016-10-21) – and though there is nothing to indicate that the second date is the date of the most recent update, the timing of the update suggests there has only been one update since it was published – and the name of the publisher (Australian Bureau of Statistics), with contact information. The metadata also includes information on the dataset’s geospatial coverage (<http://www.ga.gov.au/place-names/PlaceDetails.jsp?submit1=GA1>). This URL directs users to the Gazetteer of Australia Place Name Search page, which loads

a map of the entire country.⁷⁶⁷ However, to the left of the element set is an embedded map titled ‘Dataset extent’, which pinpoints Eidsvold, Queensland but does not provide any boundary information for the dataset’s coverage. This seeming discrepancy is easily cleared up: the dataset’s nation-wide coverage becomes clear from a cursory look at the dataset’s elements, which include the names of all of the states and territories. Temporal coverage is also indicated in the element set (2011-06-30 - 2012-06-30). This date range raises a question about what changes were made during the October 2016 update, given that the dataset covers a one year period that ended over four years earlier. Unlike the geographical coverage, discrepancies about temporal coverage cannot easily be reconciled as the changes made in October 2016 are not documented in a way that the user can see.

Organisations and individuals may sign up as users of data.gov.au. A user’s profile will include their name and a profile image (usually a logo in the case of government bodies), a user biography (optional), the date their membership commenced, the number of followers they have, datasets (published), and the number of edits they have made. Three tabs (Datasets, Activity Stream, and Data Requests) provide information on the users’ interaction with datasets, the portal and the user community. Activity Stream provides a high level audit trail associated with the user, and Data Requests documents the user’s requests for data access, establishing a level of transparency around users as well as providing data about response times and data release decisions that could conceivably be used to gauge government performance around openness. It is possible to search previous data requests. A review of the open requests shows that users also use the data request function to report problems with published datasets, such as broken links. Comments on data requests are permitted, so that users may answer each other’s requests, pointing their peers towards relevant datasets, but the responsibility to respond to and/or close the request lies with a moderator (user: DGA_Custodian). The ability to sign up and

⁷⁶⁷ The site states ‘The Gazetteer provides information on the location and spelling of more than 370,000 geographical names across Australia and its external Territories as at January 2012. The supply of data is coordinated by the Intergovernmental Committee on Surveying and Mapping and is derived from State, Territory and Commonwealth Government agencies.’ <http://www.ga.gov.au/placename?submit1=GA1> [accessed 24 July 2018].

participate in a data.gov.au user community is noteworthy because it creates a forum for participation in the curation and use of the data, which will be discussed in the following chapter, and because it provides publicly available metadata about problems with datasets and the responses of data providers. This is an expanded form of custodianship that allows a surer presumption of authenticity than if these conversations between users and providers were happening in private. Custodianship is also documented at the portal level: 'Content on the data.gov.au main site, search index, administrator account information and completed form information is hosted on Amazon servers located within Australia.'⁷⁶⁸

The Land Management Practices dataset is also tagged with metadata terms to allow users to discover other relevant datasets. These tags range from the general ('trees', 'land', 'fruit') to the specific ('sugarcane', 'water buffalo', 'gypsum'). A 'More information' tab takes users to a much smaller metadata element set ('Additional Information') that repeats some of the elements of the larger metadata element set:

Field	Value
Last updated	Unknown
Created	Unknown
Format	HTML
License	Creative Commons Attribution 3.0 Australia
Created	over 1 year ago
format	HTML
has views	1
id	8321b6dc-d7d3-479a-b276-4dd0cc73ed93
package id	67e1f959-4272-4608-83a9-edb7e59f4406
revision id	4fb9ed11-1643-4656-9b14-2a48f704bfd5
state	Active

⁷⁶⁸ Australian Government, data.gov.au <https://data.gov.au/about> [accessed 24 July 2018].

Figure 7. Additional Information metadata element set for *Land Management Practices* dataset.

Some metadata values do not agree across the sets: creation and update dates are given as 'unknown' in the Additional Information set, presumably because the metadata has not been carried over. Discrepancies in metadata across matching metadata element sets theoretically introduce uncertainty around identity and integrity, although it seems likely in this case that one instantiation of the metadata is simply more complete than the other. The Additional Information set includes checksums for 'id', 'package id' and 'revision id', and while checksums should provide surety around identity and integrity, searching data.gov.au using the checksums returns thousands of results (id returns 27,832, package id returns 28,077 and revision id returns 27,848 results) and while 'Land Management Practices' does not appear in the top 50 results for any of these searches, all of the results appear to be datasets about land or natural resource management, suggesting that the datasets reference, draw on or link to 'Land Management Practices'. A cursory exploration of the documentation around a sample of these datasets does not reveal how they relate to 'Land Management Practices'; the checksums for the Land Management Practices dataset do not appear in the metadata for the datasets returned by a search on the checksums. That the relationships between these datasets and their checksums is not explained or easily identifiable by users reinforces uncertainty around identity and integrity.

Returning to the landing page, tabs at the top of the page are 'Dataset' (the landing page described above), 'Groups', 'Activity Stream' and 'Use Cases'. There are no groups associated with this dataset. Groups are not user groups such as government agencies or communities of practice, but groups of datasets associated with particular functions. For example, the *Australian Dryland Salinity Assessment Spatial Data (1:2,500,000) – NLWRA 2001* dataset is associated with two groups: 'Business Support and Regulation' and 'Science'. The Business Support and Regulation group is described as:

Formulating policy to regulate and support the private sector, including small business and non-profit organisations. Developing strategies to assist business growth and management. Implementing advocacy programs, providing funding and administering regulatory bodies.

This description links to the National Archives of Australia's *Australian Governments' Interactive Functions Thesaurus* (AGIFT), which is a 'three-level hierarchical thesaurus that describes the business functions carried out across Commonwealth, state and local governments in Australia.'⁷⁶⁹ This is significant because it, firstly, demonstrates some synergy across areas of information management practice, as discussed in more depth in Chapter 9, and secondly because it relates the data to the context of its creation, and in particular the government functions that it relates to. This significantly enriches the contextual information around the dataset.

All groups are associated with datasets (Business Support and Regulation is associated with 81 datasets), have an 'Activity Stream' in the same way that datasets do and may be followed by users. The 'Activity Stream' tab for the Land Management Practices dataset provides a chronological list of events in the dataset's life:

- [The ABS.Stat Team](#) updated the resource [Land Management Practices](#) in the dataset [Land Management Practices](#) over 1 year ago
- [The ABS.Stat Team](#) updated the resource [Land Management Practices](#) in the dataset [Land Management Practices](#) over 1 year ago
- [The ABS.Stat Team](#) changed the extra "contact_point" of the dataset [Land Management Practices](#) over 1 year ago
- [The ABS.Stat Team](#) updated the dataset [Land Management Practices](#) over 1 year ago
- [The ABS.Stat Team](#) updated the dataset [Land Management Practices](#) over 1 year ago

⁷⁶⁹ National Archives of Australia, *Australian Governments' Interactive Functions Thesaurus* <https://data.naa.gov.au/def/agift/AGIFT.html> [accessed 24 July 2018].

- [The ABS.Stat Team](#) created the dataset [Land Management Practices](#) over 1 year ago

Although the date and time of the events and the exact nature of changes to the dataset are not documented in the Activity Stream, users are alerted to the active and ongoing curation of the dataset in sufficient detail that it should be possible to enquire into particular events.

Selecting the ‘Use Cases’ tab shows that there are no use cases associated with the Land Management Practices dataset. The portal defines use cases as ‘any apps, articles, visualisations or ideas using datasets’.⁷⁷⁰ For example, Mappify.io – a free geocoding, reverse geocoding and coordinate classification service - is listed as a use case for the *PSMA Geocoded National Address File* dataset. Linking datasets to use cases creates a connection that improves data traceability: app users’ trust in data used by tools like Mappify.io may be increased by the formal connection to the government’s data portal.

Selecting ‘Go to resource’ (the Land Management Practices dataset itself) takes the user to the dataset on the Australian Bureau of Statistics (ABS) *ABS.Stat Beta* website. The site is a free tool that provides access to ABS datasets through a web browser interface for human interaction with the data and a web services interface for machine-to-machine data use.⁷⁷¹ The site describes its functionality in terms of searching, querying, viewing, downloading, customising and exporting data, and notes that it allows users to ‘view valuable metadata alongside the data’.⁷⁷² The site is a beta release and maintains a public list of current issues with the site, and users are encouraged to provide feedback on the site’s functionality. The site notes ‘data in this ABS.Stat beta release may not necessarily be the most up to date. You are advised to check the ABS website

⁷⁷⁰ Australian Government, ‘Use Cases’, *data.gov.au* <https://data.gov.au/showcase> [accessed 24 July 2018].

⁷⁷¹ Australian Bureau of Statistics, *ABS.Stat beta* <http://stat.data.abs.gov.au//Index.aspx> [accessed 24 July 2018].

⁷⁷² Ibid.

(abs.gov.au) for the latest data'.⁷⁷³ The Land Management Practices dataset does not appear to be available on the ABS website.

The web services interface uses the Statistical Data and Metadata Standard (SDMX) to enable machine processing. SDMX considers the prime conceptualisation of metadata as it is primarily used in statistics, as the 'named dimensions of a multi-dimensional co-ordinate system'.⁷⁷⁴ The standard then differentiates 'structural' and 'reference' metadata:

A distinction can be made between "structural" metadata – those concepts used in the description and identification of statistical data and metadata – and "reference" metadata – the larger set of concepts that describe and qualify statistical data sets and processing more generally, and which are often associated not with specific observations or series of data, but with entire collections of data or even the institutions which provide that data.

The SDMX Information Model enables this reference metadata to be attached to data and/or its structural metadata, so that layers of connections and descriptions may be built up around data. The standard recognises the value of metadata in the documentation of data curation or processing, and the Information Model includes metadata about 'data provisioning', which contributes to data quality and 'allows for automation of some validation and control functions, as well as supporting management of data reporting'. The Land Management Practices dataset can be exported from ABS.Stat Beta in SDMX format (expressed in eXtensible Markup Language). A close analysis of the SDMX file is outside of the scope of this research – methodologically it is a separate activity that could form a study of the work practices of statisticians in relation to data quality – but such an analysis could illuminate the documentation of the data provisioning practices of the Australian Bureau of Statistics.

⁷⁷³ Ibid.

⁷⁷⁴ Statistical Data and Metadata Standard (SDMX), *SDMX Standards: Section 1; Framework for SDMX Technical Standards, Version 2.1* https://sdmx.org/wp-content/uploads/SDMX_2-1_SECTION_1_Framework.pdf [accessed 24 July 2018].

Returning to the Land Management Practices dataset as displayed on the ABS site, a panel on the left side of the interface allows users to find 'data by theme', with a free text search field, a drop down menu of themes such as 'Industry', 'People' and 'Snapshots of Australia'. Below the menu is a field that shows the user where the dataset they are viewing is situated within the classification scheme. For instance, the Land Management Practices dataset is found under the 'Industry' theme > 'Agriculture' > 'Land Use'. It is through this part of the interface that a comparable dataset for 2013/14 is discoverable, though it does not appear in the data.gov.au catalogue.

In the centre of the interface is a panel that allows users to interact with the dataset. Using drop down menus, users can focus on particular industries (Floriculture Production, Grape Growing, etc.), regions (national, state or territory, or Australian Agricultural Environments (AAE) regional levels such as tropics, semi-arid, etc.). The data itself is presented in a table that lists management practices on the left (Pasture management, etc.) alongside a column for 'Estimate' and one for 'Number of businesses'. The interface allows users to customise the data, making selections of data elements, changing the layout of the dataset and altering the structure of the table. The interface also allows users to export the data as a CSV file, PC-Axis file (PC-Axis is a statistical program) or SDMX document (XML) and includes a developer API query builder.

On the right side of the interface, in a panel called 'Information' additional metadata is provided, comprising the abstract that is available at data.gov.au, contact information, a note about the collection/source ('This dataset relates to the publication Agricultural Resource Management Practices, Australia (cat. No. 4630.0), available from the ABS website'), date last input received (30-Jun-12), direct source (Survey instrument: The Land Management Practices Survey), source periodicity (Biennial), and data sources used, for which the following note is included:

This survey has been compiled using data collected from the 2011-12 Land Management Practices Survey (LaMPS). The LaMPS aims to collect information on agricultural land management practices used on Australian

farms. Estimates from the survey will be used by departments of agriculture, other government agencies, industry groups and the wider community in developing and implementing policies and programs on agricultural land management practices.

This situates the dataset within the context of the process that produced it, deepening the context provided by the dataset's association with groups.

Also on the right side of the interface is a panel providing 'Data Characteristics', which includes the date and time of last update, and 'Link to Release calendar', which has the note:

The survey has also been conducted for the Financial year ending 30/06/2014 reference period. Data was initially released in July 2015 with additional data being released in October 2015 and can be found in the Land Management and Farming in Australia 2013-14 publication.

This is followed by rich contextual metadata on sampling, units of measure, geographic coverage ('dataset provides estimates at national, state/territory and Australian Agricultural Environments (AAE) regional levels'), item, population and sector coverage, statistical population, and information about data aggregation and consolidation, classification (referencing the Australian Statistical Geography Standard 2011 and Australia and New Zealand Standard Industrial Classification (ANZSIC 2006)), estimation, imputation, dissemination formats, notes on any seasonal adjustment, validation (referencing ABS procedures) and weights. All of this information enables the interrogation and interpretation of the dataset, as well as its reuse.

Finally, the right side panel links to a Quality Declaration. The link to the Quality Declaration takes users to a statement specific to this dataset, which includes notes on the 'Institutional Environment' (the regulatory environment of the ABS, including legal obligations, and financial and governance arrangements and scrutiny of ABS operations, with a link to a statement about that environment), 'Relevance', 'Timeliness' (noting that the dataset was published within twelve

months after the reference period), 'Accuracy' ('Data from the 2011-12 LaMPS are based on a response rate of 85.3% and has been extensively quality assured. The resultant value estimates have been checked against relevant sources, including contact with data providers, to ensure the accuracy of the estimates produced. '), 'Coherence' (pointing to Australian Taxation Office and census data), 'Interpretability' (providing a short statement on the subject of the dataset and noting that explanatory notes are available) and 'Accessibility' (directing users looking for other data to a contact phone number for the ABS's National Information and Referral Service). All of this additional information alerts users to the potential for errors or omissions in the data, and helps to frame the interpretation and use of the dataset.

At a number of places in this part of the interface, a link is provided to 'Explanatory Notes', which are also specific to this dataset, and expand on various aspects and parameters of the dataset. Importantly, the Notes include a statement on 'Reliability of Data' ('The estimates in this publication are subject to sampling and non-sampling errors') and descriptions of the likely sampling and non-sampling errors. Selected statistics within the dataset are listed with relative standard error (RSE) estimates:

RELATIVE STANDARD ERRORS OF SELECTED ESTIMATES, by State and Territory - 2011-12

	NSW	Vic.	Qld	SA	WA	Tas.	NT	ACT	Aust.
	%	%	%	%	%	%	%	%	%
Total area of intercropping (ha)	7.8	16.5	6.4	12.0	8.8	12.0	1.4	-	4.3
Land mainly used for crops - Area (ha)	1.8	3.7	2.5	3.0	1.8	3.7	4.9	21.8	1.0
Total area cultivated - Area (ha)	2.9	5.8	2.5	5.1	2.8	4.4	10.9	14.8	1.6
Crop stubble/trash managed - Cereals								32.1	
- Total area (ha)	2.6	5.1	4.6	5.0	2.7	2.6	3.1		1.7
Fertigation - Area (ha)	5.1	17.4	8.2	8.7	36.3	36.0	5.0	-	5.4

Water logging avoidance practices - All practices - Area (ha)	9.1	15.7	4.5	34.2	16.2	4.9	2.1	-	5.5
Livestock supplied with feed additives - Total - Head (n)	5.9	8.8	5.3	8.0	8.6	7.6	2.8	-	3.3
Lime and/or dolomite - Area applied to (ha)	7.8	8.4	4.0	15.7	4.1	9.6	5.0	11.9	3.3

Figure 8. Relative Standard Errors of Selected Estimates

The note on Non-Sampling Errors states:

Errors other than those due to sampling may occur because of deficiencies in the list of units from which the sample was selected, non-response, and errors in reporting by providers. Inaccuracies of this kind are referred to as non-sampling error, which may occur in any collection, whether it be a census or a sample. Every effort has been made to reduce non-sampling error to a minimum by careful design and testing of questionnaires, operating procedures and systems used to compile the statistics.⁷⁷⁵

This overt statement regarding possible errors echoes similar statements elsewhere in the interface, and rather than undermining trust, they are likely to reinforce it.

The Land Management Practices dataset is rich in content and in context. There are some weaknesses in the metadata that raise questions about the integrity of the dataset, but as in other places in the dataset's metadata and supporting documentation, the interface flags issues of data quality, so that users are not only alerted to possible errors or omissions, but advised on where and how these errors and omissions might manifest in the dataset.

8.6. Conclusion

⁷⁷⁵ Australian Bureau of Statistics, *ABS.Stat beta* <http://stat.data.abs.gov.au/Index.aspx> [accessed 8 September 2018].

The Australian federal government was the first Westminster-style government to introduce Freedom of Information legislation. The seeming ease with which FOI was taken up by government and passed into law, and the coordination across the access and archival legislation, seems to have set a precedent for innovation and harmonisation in government approaches to information management and access provision. This pattern is repeated in the open data movement in Australia, where linkages are present between data.gov.au, the Australian Bureau of Statistics and the National Archives of Australia. The portal directs users of the Land Management Practices dataset to *ABS.Stat Beta* and makes use of NAA's AGIFT. This final connection is particularly innovative because it links datasets to the government functions to which they relate, situating the data within government work in a way that provides context.

Although there are some problems with the data released through data.gov.au, the portal flags the likelihood of errors and omissions at multiple places in the user interface. Unlike in the Kenyan case study, this study of the Land Management Practices dataset released through data.gov.au has not had to guess at the lifecycle of the dataset; a series of questions about the chain of custody of the dataset are forestalled by the provision of metadata about provenance, data curation practices and post-publication changes to the dataset. This may reflect the embeddedness of record-keeping in Australian government culture but it is more likely to be attributable to transparency around statistical practices. However, the use of checksums for the Land Management Practices dataset should allow it to be located and identified as a unique digital object, but the checksums associated with it return numerous results. It is also not discernible which algorithm has been used to generate the checksums. There are still minor issues that raise questions about data quality. Data.gov.au points users to *ABS.Stat Beta*, which is still in development, and minor issues such as the demarcation of geographical coverage through embedded maps and the transfer of metadata across duplicate elements in different parts of the interface may be dealt with over time. Also, errors can be reported by a user community whose communication with official data providers is documented publicly. Over all, in terms of the controls supporting trustworthiness in digital records as distilled from the InterPARES work, the Land Management Practices dataset is rich in

metadata, has documented custodianship to a high level, though it is not complete, and auditable systems seem to be in place.

Chapter 9

Applying Record-Keeping Controls to Government Data Curation

...a power relationship can only be articulated on the basis of two elements which are each indispensable if it is really to be a power relationship: that “the other” (the one over whom power is exercised) be thoroughly recognized and maintained to the very end as a person who acts; and that, faced with a relationship of power, a whole field of responses, reactions, results, and possible inventions may open up.⁷⁷⁶

Michel Foucault, *The Subject and Power*

It is this trust – or more specifically, the conventions for establishing and sustaining this trust – that is now under question.⁷⁷⁷

Heather MacNeil and Bonnie Mak, *Constructions of Authenticity*

⁷⁷⁶ Foucault, M., ‘The Subject and Power’, *Critical Inquiry*, 8.4 (1982) p.789.

⁷⁷⁷ MacNeil, H. & Mak, B., ‘Constructions of Authenticity’, *Library Trends*, 56.1 (2007) p.27.

9.1. Introduction

Efforts to open government data are framed as transparency and accountability initiatives, but the trustworthiness of data is rarely a consideration in open data initiatives. Increasingly, NGOs and others are taking a critical approach to data quality, but much of this centres on access to quantities of data, the usability of the data, and its completeness, rather than its trustworthiness. In 2017, Transparency International, an international civil society network working against corruption, published a study of five G20 countries.⁷⁷⁸ The study found that none of the countries released all the key anti-corruption datasets, and that when released 'data is not often useful and useable'.⁷⁷⁹ The report does not flag issues of trustworthiness, and in this respect it is typical of analyses of open data programmes.

In this chapter, I will compare the findings of the Kenyan and Australian cases across the three concepts identified in Chapter 5 - metadata, documented custodianship and auditable systems - to demonstrate the significance of these controls for the trustworthiness of open government data. I will then deal with two further considerations in the application of these controls. Firstly, I will consider the differences between data and records and the ways in which the controls define data as records and records as data. Secondly, I will argue that the controls must be exercised in public to be truly effective.

9.2. Comparison between Kenya and Australia

The Kenyan and Australian case studies illustrate that the colonial legacy of official secrecy, in particular a presumption of secrecy rather than openness, characterised government information culture in the two countries in the twentieth century. Although Australia may have come further than Kenya in moving towards openness, secrecy persists in parts of its bureaucracy where citizens

⁷⁷⁸ Transparency International, *Open Data: Promise, but not enough progress from G20 countries* (2017)

https://www.transparency.org/news/feature/open_data_promise_but_not_enough_progress_from_g20_countries [accessed 13 September 2018].

⁷⁷⁹ Ibid.

and other actors are eager to see more transparency. With Freedom of Information legislation in both countries, some form of legislative foundation for openness has been established, though neither should be read as fixed in place, as the political and bureaucratic enactment of open cultures and practices fluctuate with circumstance and policy direction: the secrecy around migrant detention centres in Australia is a powerful example that FOI legislation is only one part of a more complex set of arrangements that are necessary for effective agent monitoring in the social contract. Both countries are engaged with open data initiatives, and though they are in different stages of maturity, similar technical issues are seen in the curation of data with assurances of authenticity. I will compare the findings of the Kenyan and Australian case studies in relation to the three key concepts I derived from archival and diplomatic thinking about authentic records, as described in Chapter 5, *Trustworthy Information*: metadata, documented chains of custody and auditable systems.

9.2.1. Metadata

The 'Proportion of Parcels Using Fertiliser County Estimates 2005/6' dataset published by KODI lacks detail and consistency in its metadata, detracting from a sense of authenticity, and impeding the interpretation of the data. Patchy metadata is a problem with the dataset both internally and externally: without sufficient metadata to make the CSV file human-readable, the user must depend on KODI's visualisation of the dataset, which requires trust in the visualisation algorithm, which is not publicly documented and therefore cannot be interrogated. The absence of internal metadata directly relates, then, to a gap in the chain of custody (the undocumented algorithm and its processing). As argued in the case study, though, the more problematic issue of metadata is the absence of metadata around the dataset. Problems with this metadata call into question the identity of the dataset (it appeared on KODI with different names at different times). Furthermore, provenance and chain of custody are undocumented in places, because of a lack of adequate metadata, calling into question the integrity of the dataset. In relation to the InterPARES ontology of trustworthiness, in which

the two constituent parts of authenticity are identity and integrity, this dataset is suspect in both respects as a result of poor metadata.

The 'Land Management Practices (2011-2012)' dataset released through data.gov.au is much richer in metadata, and the portal provides links to sources of information about the curation and publication of the dataset. However, there is still room to question the identity and integrity of this dataset because the checksums associated with the dataset return numerous results when used as a search term within the data catalogue, and the dataset is not discernible amongst those results. The checksum is a piece of metadata that should serve as a unique identifier for an unchanged file, offering an assurance of both identity and integrity. That the checksums for the dataset return different results may be attributable to the variety of datasets that make use of 'Land Management Practices (2011-2012)', but there is no explanation of this, and the problem remains that the dataset itself is not returned in the results. This is a problem of metadata but also one of system, and in particular the information retrieval capabilities of the data catalogue.

A government file is 'an unusual sort of artifact because', according to Hull, 'signs of its history are continuously and deliberately inscribed upon the artifact itself, a peculiarity that gives it an event-like quality'.⁷⁸⁰ J. Peter Sigmund's description of the *kantbeschikking* illustrates the same principle, and extends it to suggest that the accumulation of metadata is a prolongation of the record-making process:

When... in the seventeenth century a request was made by a certain person to an administration, that individual had to address himself via a special form. Someone in the administration who received the request probably would give his opinion on the same piece of paper, while another person would decide and record his decision, again on the same document. Then it would be returned to the applicant. This form of document is called a *kantbeschikking* (decision-in-the-margin). Administratively, the decision in the margin is the most important part of the document, for it is the end of the

⁷⁸⁰ Hull, M. S., *Government of Paper: The Materiality of Bureaucracy in Urban Pakistan*, (Berkeley; Los Angeles; London: University of California Press, 2012) p.116.

procedure. What we see here is, in fact, the complete administrative procedure in a nutshell, since all decisions on similar subjects were made up in the same way and can still be recognized on the document.⁷⁸¹

This ever-developing or emergent quality of the record or file – being enriched in contextual metadata over time – is a model for improved data curation. As Geoffrey Yeo has observed, ‘Record-keeping is, or should be, concerned not only with propositions, but also with the contextualized acts of assertion that generated them. Some information managers and data administrators are now beginning to take an interest in context..., but concerns for contextualization are generally more characteristics of archives and records management’.⁷⁸² The documentation of context through metadata is essential to establishing authenticity and therefore trustworthiness in information, and the application of this record-keeping principle to government datasets would therefore help to correct information asymmetry. Duranti wrote that for:

accountability to exist it is necessary that people be able to exercise their right of scrutiny. But the establishment of metadata systems, the creation of large information locator systems, and the provision of navigational means are not sufficient to empower the people to exercise their right. The people should be able to rely on the authenticity over time not only of the documents they see, but also and foremost of their context.⁷⁸³

The significance of the documentation of context to allow meaningful interpretation of data suggests the need to closely control that contextualisation so that it is itself reliable and authentic. As with its physical counterpart, as described by Hull, every step in the progress of a digital file or a dataset is another opportunity for the addition of erroneous metadata or the omission of some crucial metadata value. The foliation, the columns of the various registers, and the metadata elements formalised in the stamps of the Bengal Manual, all open up spaces for recording inaccurate metadata or not recording at all: it is

⁷⁸¹ Sigmund, J. P., ‘Form, Function and Archival Value’, *Archivaria*, 33.2 (1991) p.142.

⁷⁸² Yeo, G., *Records, Information, and Data: Exploring the Role of Record Keeping in an Information Culture* (London: Facet Publishing, 2018) p.168.

⁷⁸³ Duranti, L., ‘Archives as Place’, *Archives and Manuscripts*, 24.2 (1996) p.253.

through the systemisation of metadata inscription practices and formulae that errors or problems become most visible. In 'Proportion of Parcels Using Fertiliser County Estimates 2005/6', discrepancies in metadata only become apparent when using the system (KODI) to retrieve the dataset on different occasions. In 'Land Management Practices (2011-2012)', the checksum is unproblematic metadata until it is put into play with the system (the data catalogue in this case).

9.2.2. Documented Custodianship

This metadata about events and interventions must continue to accrue around data as long as the data persists if users are to be able to interrogate the data and ascertain its trustworthiness. This metadata must also assign responsibility for each event and intervention, whether to a physical or juridical person or to a device, algorithm or network. It is the assignation of responsibility that creates an impetus for ethical custodianship, as the examination of the Bengal Manual in Chapter 6, *Practices for Establishing and Maintain Authenticity in Information*, demonstrated.

A double chain is formed when custodianship is documented: one *in mundo* and one 'in the medium of literality'.⁷⁸⁴ Ribes and Jackson reflect on variations in data collection practices, including changes in staff, instruments, equipment, changes of site and new 'subprojects in data collection',⁷⁸⁵ 'all the space between the archive and the field site unaccounted for'.⁷⁸⁶ Understanding this work of custodianship is crucial to judgments of authenticity in data. The capture of metadata about custodianship forms a chain of documentation that parallels the chain of custody. It is the data double of a process and, in places, it is the process itself. Writing about the data capture practices of ecoscientists in Baltimore, Ribes and Jackson said of using identification numbers to link scientific samples with their documentation, 'If this simple numbering ritual fails,

⁷⁸⁴ Vismann, C., *Files: Law and Media Technology*, trans. by Winthrop-Young, G., (Stanford: Stanford University Press, 2008) p.56.

⁷⁸⁵ Ribes, D. & Jackson, S. J., 'Data, Bite, Man, The Work of Sustaining a Long-Term Study' in Gitelman, L. (ed.), *"Raw Data" is an Oxymoron*, (Cambridge: MIT Press, 2013) p.156.

⁷⁸⁶ *Ibid.*, p.161.

so too does the chain that connects the field to the lab⁷⁸⁷ and that this ‘labeling ritual, the notations on samples, the checks at each point of transport, are the cascades of rituals that tie together field sites to samples to databases’.⁷⁸⁸ Similarly, the series of processes that constitute the custodianship of other forms of data must correlate to metadata values linked with the data if it is to be trusted. Ribes and Jackson state that the ‘metaphor of a chain is revealing: it helps us understand the heterogeneous work of custodianship stretching from field site to lab and from lab to databases’.⁷⁸⁹ The notion of the heterogeneous work of custodianship helps us to relate metadata values to events and interventions by human and non-human actors: to remember that each piece of metadata is added, chosen or phrased according to an intentional act.

This custodial work is also curatorial work. Gitelman observed that ‘...the imagination of data is in some measure always an act of classification, of lumping and splitting, nesting and ranking, though the underlying principles at work can be hard to recover’.⁷⁹⁰ The interventions of custodians / curators shape the dataset just as they shape records. In the record-keeping literature, co-creation and the ‘rights in records’ of records’ subjects is gaining ground as a framework for thinking about access rights and systems design, amongst other issues.⁷⁹¹ As with that work, the critical work around data curation could think through the rights and contributions of those involved across the chain of events that constitute curation.

Surfacing the curatorial role of the custodian has important social consequences, at the level of the datum or metadatum, the level of the record or dataset, the file, series, fonds, collection, archive, data catalogue, etc. etc. Each of these tells a story about events, people and places and these stories can be about rights, understanding, injustice, and so on, and importantly, each story is being told by

⁷⁸⁷ Ibid., p.162.

⁷⁸⁸ Ibid., p.163.

⁷⁸⁹ Ibid.

⁷⁹⁰ Gitelman, L. & Jackson, V., ‘Introduction’, in Gitelman, L. (ed.), *Raw Data*, p.8.

⁷⁹¹ Gilliland, A. J. & McKemmish, S., ‘Rights in Records as a Platform for Participative Archiving’ in Cox, R. J., Langmead, A. & Mattern, E. (eds.) *Archival Education and Research: Selected Papers from the 2014 AERI Conference* (Sacramento: Litwin Press, 2015); Gilliland, A. J. & McKemmish, S., ‘The Role of Participatory Archives in Furthering Human Rights, Reconciliation and Recovery’ in *Atlanti* 24.1 (2014).

someone with agendas, beliefs, alliances, etc. Or, as Jeannette Bastian claimed ‘...the early role of the custodian was often one of power. Control of the records meant control of the subjects of the records’.⁷⁹² The person reading the data should know who these curators are in order to make judgements about the story they are being told. This concept is well-established in the field of record-keeping: ‘If the process of preservation is made visible, users are better equipped to make an informed decision about whether the materials meet their specific requirements for authenticity’.⁷⁹³ It is not yet well-established in open government data curation.

Furthermore, each event of metadata capture represents an assertion that performs a non-repudiation function. Each metadata element is theoretically attributable to a physical or juridical person - or an instrument or algorithm that is attributable to a physical or juridical person – and in the choice of metadata, the author is making a claim. If that person can be identified, the trustworthiness of the metadata can be construed to reflect the trustworthiness of the author in the same way that signatures have historically done. This complicates the field of metadata analysis by making possible any number of configurations of authorship, enacted as metadata capture, about custodianship. The way that processes and systems are designed, how and where they automate actions, allow documentation by custodians or about them by third parties, whether or not they make room for dissent; these considerations establish system design as a space of dissensus. Each record-keeping system, data management workflow and databank is in fact a forum where assertions are made. MacNeil wrote that a ‘digital signature is distinguished by its “non-repudiation function,” and, like the traditional signature, its purpose is to foreclose any further speech concerning the truth-value of the record to which the signature is attached’,⁷⁹⁴ but in processes of co-creation and chains of custody, reading assignments of responsibility in metadata opens up rather than forecloses speech about the truth-value of the data. These bureaucratic systems are places of contestation, even as they appear monolithic in their authority. Information asymmetry is implicit whenever a

⁷⁹² Bastian, J. A., ‘Taking Custody, Giving Access: A Postcustodial Role for a New Century’, *Archivaria*, 53.1 (2002) p.82.

⁷⁹³ MacNeil & Mak, ‘Constructions of Authenticity’, p.17.

⁷⁹⁴ MacNeil, H., ‘Trusting Records in a Postmodern World’, *Archivaria*, 51.1 (2002) p.44.

user interacts with such systems without knowing how to, or that it is possible to, refute them. As Yeo observes, 'data are seldom as uncomplicated as they seem. When we look at a database, no-one seems to be making statements; no-one is affirming that they can vouch for the data; the apparent absence of signs of authorship gives the impression that the data are uncontroversial and objective'.⁷⁹⁵ This echoes Gitelman:

At first glance data are apparently before the fact: they are the starting point for what we know, who we are, and how we communicate. This shared sense of starting with data often leads to an unnoticed assumption that data are transparent, that information is self-evident, the fundamental stuff of truth itself.⁷⁹⁶

In the case of KODI and data.gov.au, without sign-posting, users may read the data in an uncritical way.

I would argue that the publication of a dataset online performs a non-repudiation function – that each dataset is an utterance from government, with all the power that Vismann attributes to files. Data-as-command. MacNeil and Mak claimed that there 'is no final act of closure in the digital environment that corresponds to the traditional notion of "publication", indicating the final, finished piece. Rather, the "publication" of a digital work might entail the opposite – that it is now open and ready for the copying'.⁷⁹⁷ This is debatable, since at the time they were writing, EDRMSs enabled users to 'Declare as Records' any digital document considered final, and to close aggregations and archive them in processes that mirrored registry practices. Nevertheless, MacNeil and Mak anticipated the open data ethos of distribution and reuse: publication does entail opening datasets to copying. However raw, partial or erroneous the data is, in publishing datasets, governments are asserting something: where the data is known to be problematic, it is in the interests of governments to flag that for users through metadata. Furthermore, the attribution of data curation activities to the various

⁷⁹⁵ Yeo, *Records, Information, and Data*, p.142.

⁷⁹⁶ Gitelman & Jackson, 'Introduction', p.2.

⁷⁹⁷ MacNeil & Mak, 'Constructions of Authenticity', p.16.

stakeholders engaged in the production of the dataset distribute that responsibility and non-reputability in a way that echoes more recent archival thinking about multiple or parallel provenance.⁷⁹⁸ Bureaucracies are formed of individuals organised into units, agencies, ministries and departments that will curate and share data across organisational boundaries, as well as the state / civil society boundary, and data may have single or multiple and complex provenances. Documenting these improves trustworthiness. The notion that publications of datasets are non-repudiation functions suggests that web archiving offers an important proof of the pronouncements of governments: the webcrawler's snapshot becomes a proof of signature.

Where systems for managing data do not require this 'signature' (linking assertions to authors), the system, or its data, but chiefly the institution that operates the system, is always speaking *ex cathedra*. This is, in fact, the root of information asymmetry in the public sphere. Through the organisation of processes of information capture, management, use and access in such a way that these processes *exclude* by refusing interrogation, or the linking of events in the life of data to custodians, institutionalised information appears monolithic and unassailable. This is a manifestation of the phenomenon of lumping and individualising discussed by Deleuze and Virilio, where the same apparatus that collects information datum by datum, that dissects every datafied object, person and event and gathers that information into itself simultaneously presents as a unity that is encyclopedic, omniscient and authoritative. There is a parallel here with the theory of the firm where the body corporate is composed of individual actors that form its fuzzy boundaries and conduits of valorised information. The formation of the institution is mirrored in the formation of its corpus of information. Access to one might be expected to reflect access to the other, and in the public sphere, this analogy suggests a direct correlation between democratic participation and access to information.

9.2.3. Auditable Systems

⁷⁹⁸ Hurley, C., 'Parallel Provenance: What, if Anything, is Archival Description?' *Archives and Manuscripts*, 33.1 (2005).

The documentation of the custodianship of information using metadata requires a mechanism for joining the information to the custodian. This mechanism is a system for inscribing, capturing, and binding metadata to that information across time and space. As MacNeil notes in relation to electronic records systems, the integrity of the system 'is ensured by verification controls and security procedures, which, like bureaucratic controls, constitute a kind of surveillance over the system and the individuals using it'.⁷⁹⁹ Metadata is used to bring control to information, but also the process and people that created the information. This is evident in the Bengal Manual, which states that in 'each fair copy of a letter the date of the draft should be noted under the place for the Collector's signature; this will enable the Collector at a glance to check delay in the Copying Department'.⁸⁰⁰ Its parallel in ICA-Req is:

124 – Ensure that metadata is available for inspection on request, so that a specific event can be identified and all related data made accessible, and that this can be achieved by authorized external personnel who have little or no familiarity with the system.

In Chapter 6, I showed how systems such as registry systems have worked to create metadata around records so that custodianship is documented. This analysis showed the way such systems create multiple and distributed traces of the record in a way that improves the trustworthiness of the record by attesting to its authenticity. An analysis of the systems at work in the open data initiatives of the Kenyan and Australian governments illustrates the significance of systems for eliciting trust in information.

In Kenya, numerous government systems are involved in the curation of the various data sources used to create the 'Proportion of Parcels Using Fertiliser 2006' dataset. The case study showed that the traditional paper registry system used in the Ministry of Lands and Physical Planning is well-functioning. The Integrated Records Management System that was being introduced across

⁷⁹⁹ MacNeil, H., *Trusting Records: Legal, Historical and Diplomatic Perspectives* (Dordrecht; London: Kluwer Academic, 2000) p.56.

⁸⁰⁰ Board of Revenue, Government of Bengal, *The Bengal Records Manual 1943*, (Alipore: Bengal Government Press, 1943) p.9.

government did not appear to be in use in the management of land records; if it had been in use, an assessment of the system for compliance with a standard such as ICA-Req, discussed in Chapter 6, would be required to make a judgment about the ability of the system to effectively document custodianship. KODI's documentation of the dataset does not reveal the journey of the data from its source to the portal, so it is not possible to see which other systems the data passes through, but it is clear that it spends some time in the custody of the Kenya National Bureau of Statistics before coming to KODI. KODI itself is a system with a part to play in the curation of the data, since it has custody of the dataset at the point at which it becomes public. The case study shows how problematic KODI is as a data curation system: on different dates, searches of the system revealed differing metadata for titles, dates and publishers. These metadata elements are essential for establishing trustworthiness – the title of the dataset is analogous with the identity of the dataset, while dates also help to establish identity, and the 'publisher' element attributes authorship, which is important in judgments of reliability. The discrepancies in metadata in KODI show that the system works against attributions of trustworthiness. Most importantly, KODI provides visualisations that are essential to the interpretation of the dataset. The case study showed that the metadata within the dataset was insufficient for its interpretation by a layperson. Clause 124 of ICA-Req, cited above, specifically states that inspection should be able to 'be achieved by authorized external personnel who have little or no familiarity with the system'.⁸⁰¹ It was only through KODI's visualisation of the dataset that the data could be read by the untrained eye. This makes KODI a pivotal system in the curation of the dataset, since it is the mechanism for rendering the dataset. There is no data on KODI that tells the user how this is done, so the algorithms cannot easily be audited.

In contrast, the systems at work in the curation of the 'Land Management Practices (2011-2012)' dataset published on data.gov.au are more visible at the point of publication. While it was not possible to assess the record-keeping

⁸⁰¹ International Council on Archives, *Principles and Functional Requirements for Records in Electronic Environments - Module 2: Guidelines and Functional Requirements for Electronic Records Management Systems* (International Council on Archives, 2008) p.39.

practices of the government agencies whose data contributed to the dataset, the case study in Chapter 8 showed that there is a strong regulatory framework and culture of record-keeping in the Australian public service. Treating data.gov.au as a data curation system, the public can assess its performance in documenting the custodianship of datasets by looking to the metadata provided: names of datasets, dates of publication and updates, and names of publishers with contact information. In contrast to KODI, there are no discrepancies in the identity information of the dataset examined in the case study. This is not to say that there are not errors in the data or its metadata, but questions are not raised in the data.gov.au case because the most basic metadata is consistent: trustworthiness is not immediately called into question as in the KODI case. In addition to consistency across metadata, the data.gov.au case shows that volume of metadata has a part to play in assuring trustworthiness. The metadata about datasets is further enriched by the documentation of conversations around data and its management in user profiles, data requests and 'activity streams'. In a sense, data.gov.au is also a forum for the interrogation of data. Importantly, it demonstrates transparency in the criticism and correction of data, which improves trust in the data and the system that provides it. There are some weaknesses, however: it is not possible to see what changes were made to the data in updates, some metadata values do not agree across instantiations of the metadata element sets that appear in different parts of the portal, and there is ambiguity around the way that checksums are used.

In both cases, the points between systems are weak points in the chain of custody and its documentation. In the passing over of data from one curation system to another, documentation can be lost, changed or never created. In particular, provenance information should, but rarely does, in these case studies, travel with data across systems. In data.gov.au, users can see that the National Bureau of Statistics published the data, but users cannot see its sources. For that information, users would need to engage with National Bureau of Statistics, and to trust what they found there, users would need to be able to interrogate its systems. In data.gov.au, the system refers to and draws on other systems or tools, such as the *ABS.Stat Beta* website, the *Gazetteer of Australia Place Name Search* and the National Archives of Australia's *Australian Governments'*

Interactive Functions Thesaurus (AGIFT). The same questions and assurances of trustworthiness that apply to the datasets are relevant to trust in the system, or interconnected web of systems, that are used to create data. As Gilliland has said 'metadata, like archives, are endless'⁸⁰²; this can be extended to systems – like bureaucracies, their boundaries are not always clear, and they often depend on and feed into other systems. Surfacing these connections is important in establishing trust in the data they hold and process.

As MacNeil has stated:

The most pervasive characteristic of bureaucracy is the existence of a system of control based on rational rules, that is, rules meant to design and regulate the entire organization on the basis of technical knowledge and with the aim of achieving maximum efficiency. Reliable and authentic records are critical to that control because they are the primary means by which bureaucrats account for their actions.⁸⁰³

KODI and data.gov.au are systems of control in two ways: they control data for part of its life, indeed that crucial part when it becomes public, and, like the record-keeping systems that are connected to them via data and metadata, they are systems within a larger system of control 'by which bureaucrats account for their actions'.⁸⁰⁴ Roberts observes two changes in the production of what he considers unstructured data: increased diversity and quantity – both resulting from technological changes in office work.⁸⁰⁵ It is the system that introduces control over this diversity and quantity. Weld wrote that the 'documentary flow within any bureaucracy reflects that bureaucracy's structures, relationships between substructures, and internal hierarchies', but it is the system that makes

⁸⁰² Gilliland, A. J., 'The Wink that's Worth a Thousand Words' in, Smit, F., Glaudemans, A. & Jonker, R. (eds.) *Archives in Liquid Times* (Stiching Archieffpublicaties, 2017) pp.212-227.

⁸⁰³ MacNeil, 'Trusting Records in a Postmodern World', p.41.

⁸⁰⁴ Ibid.

⁸⁰⁵ Roberts, A., *Blacked Out: Government Secrecy in the Information Age* (New York: Cambridge University Press, 2008) p.212.

these structures and relationships visible, through the attribution of accumulative metadata.⁸⁰⁶

9.3. The Application of Record-Keeping Controls to Datasets

Agency theory holds that principal-agent relationships become problematic because the goals of the agents do not necessarily align with the goals of principals, and that it is through the contract that these goals can be aligned. According to agency theory, '[r]esolution of these problems is difficult... because of information asymmetry favoring agents...'.⁸⁰⁷ Theorists of information asymmetry have proposed mechanisms to address information asymmetry and of these mechanisms, monitoring by principals is key to addressing information asymmetry in the social contract. Open data is a monitoring mechanism, but monitoring requires both accessibility and trustworthiness. The case studies have shown that the open data initiatives of governments succeed or fail as monitoring mechanisms with respect to trustworthiness on the basis of the controls in place for metadata, documented custodianship and auditable systems. These controls are derived from hundreds of years of record-keeping practice, as discussed in Chapters 5 and 6, but how does the notion of a record resonate with the nature of data?

When government data is published it is delivered from official control. The repurposing of government data by the private and third sectors is a key driver of the open data movement, and much of its rhetoric concerns reuse as a commodity for commercial innovation, as shown in Chapter 4, *Opening Government: A History of Monitoring*. This repurposing is possible because data is fissiparous: it is prone to being broken up.

The controls that make trust in information viable, such as those analysed in Chapters 5 and 6, are derived from record-keeping, a field of work concerned with the management of records over time to ensure that their administrative,

⁸⁰⁶ Weld, K., *Paper Cadavers: The Archives of Dictatorship in Guatemala* (Durham: Duke University Press, 2014) p.71.

⁸⁰⁷ Sharma, A., 'Professional as Agent: Knowledge Asymmetry in Agency Exchange,' *The Academy of Management Review*, 22.3 (1997) p.760.

legal, historical and other uses remain possible for as long as required by the community or legal system served by those records. Records do not share fissiparity with data. Instead, records are bound and structured objects. Traditionally, they have been defined as information objects with content (data), context (metadata) and structure (form or format). The extraction of data is possible, and indeed ICA-Req states that electronic records management systems must ‘249 – Allow the re-use or re-purposing of content’,⁸⁰⁸ but the evidential qualities of records are conferred by their context and structure, as discussed in Chapter 5, *Trustworthy Information*. The principles of systematic control and the documentation of custodianship through metadata concern data as an informational object that is wrapped in metadata that accrues as the object moves and is used.

In record-keeping, the definition of a record has been debated for many years. This debate has concerned qualities such as fixedness, originality, as well as trustworthiness. Amelia Acker has argued that ‘the project of defining a record should be abandoned in the age of networked records. Instead we must expand our focus to the ever-broadening layers of infrastructure and context through which digital records move’.⁸⁰⁹ She states that ‘...in the digital age the ontological purity and drive for the true nature of the record is over. It cannot be exclusively or definitively located because it is not in a single place, it never ends, and is “always in a process of becoming” in the continuum’.⁸¹⁰ Instead, we should be ‘Analysing how a record moves through networks, and asking when a record is becoming, and how and where it persists’ because this ‘is a more productive method for understanding the nature and effects of infrastructure and ultimately how electronic evidence is socially constituted with technologies’.⁸¹¹ This view of records as evolving information objects in motion brings ‘records’ much closer to

⁸⁰⁸ International Standardisation Organisation, *ISO 16175:2012 - Information and documentation - Principles and functional requirements for records in electronic office environments*, (Geneva: International Standardisation Organisation, 2012) p.58.

⁸⁰⁹ Acker, A., ‘When is a Record: A Research Framework for Locating Electronic Records in Infrastructure’, in Gilliland, A. J., McKemmish, S. & Lau, A. J. (eds.) *Research in the Archival Multiverse* (Clayton: Monash University Publishing, 2017) p.316.

⁸¹⁰ Ibid.

⁸¹¹ Ibid.

the notion of data as opened and circulating across diverse networks. It echoes with what MacNeil and Mak have said about digital images:

It is part of the character of these resources to be copied and reinterpreted in different contexts. Thus, if one of the qualities of digital materials is to be allographic, that is, to enable copying and manipulation and to be used in different ways and for different purposes, this aspect may need to be accommodated in the process of preservation.⁸¹²

In addition to the allographic nature of digital resources, which involves multiple and simultaneous intentionalities, also brought to the fore is their polysemic character. That is, in addition to their ability to support multiple and simultaneous intentionalities, they also encourage multiple, simultaneous, and possibly dissonant meanings.⁸¹³

Can data and records, then, be conceived of as existing on the same scales of quality, richness or evidential weight? Can we observe records ‘becoming’ from data and devolving to data through their movement into different contexts and systems? If so, the crucial question in relation to information asymmetry and open government data’s role in correcting it, is ‘How do we make open government data more like a record?’ The case studies have shown that there is a scale: some data are more trustworthy than others because of the extent of the implementation of controls developed in record-keeping. Data will be repurposed once it is in the public domain, but throughout its curation up to and including its publication on portals such as KODI and data.gov.au, auditable controls must be in place if the data is to be trustworthy. Rosalind Krauss asked ‘Is the “support” of film the celluloid strip, the screen, the splices of the edited footage, the projector’s beam of light, the circular reels?’⁸¹⁴ and a similar question can be asked of datasets. For data to ‘happen’ a whole set of supports are called into concert. The user needs to be able to see these, and their impacts on the data. The acceptance of the reliability and authenticity of the information will be judged

⁸¹² MacNeil & Mak, ‘Constructions of Authenticity’, p.16.

⁸¹³ Ibid.

⁸¹⁴ Krauss, R., ‘Two Moments from the Post-Medium Condition’, *October*, 116 (2006) p.56.

differently by different users: as MacNeil has argued, this ‘cannot be defined in any monolithic sense’.⁸¹⁵ There is necessarily an element of subjectivity in assessments of the trustworthiness of information, but as the case studies have shown, some data are more trustworthy than others.

Where metadata, custodianship and systems have a correcting effect on information asymmetry, they share in common the characteristic of *publicity*. In order for metadata to assure users of the source of data, users must be able to read it. In order for users to know that custodianship has been consistent or that curation has been appropriate, users must be able to see it, either in operation, or to see its traces in metadata. To trust the system that unites these elements and presents them to users, users must be able to watch the system work. It is not enough to have transparency of information: users must have transparency in the controls over the information. Data curation must be done *in public*. In contrast to the Victorian concept of the public service ethos in which ‘publicity has no part...’, publicity is necessary for the avoidance of moral hazard in public service, and this can be extended to moral hazard in data management.⁸¹⁶ As MacNeil argued, ‘...the very invisibility of technological control mechanisms is precisely what makes them even more insidious than traditional mechanisms of bureaucratic control’.⁸¹⁷ Visibility is essential. To have authority, the public seal must be *bene cognitum et famosum*.⁸¹⁸ In the same way, to be ‘recognised by all subjects’, each attribution of metadata, act of the custodian or step in the system’s processes should be ‘well known and famous’.⁸¹⁹ In this sense, each curatorial intervention is an attribution of authorship and authority. As with livery of seisin, the validity of the act is judged by its witnesses: all curatorial acts, whether by living custodians or by automated systems, are open to question. The authority of the data can only be judged by the evidence of its proper treatment, and this treatment must be witnessable. Information symmetry guards against moral

⁸¹⁵ MacNeil & Mak, ‘Constructions of Authenticity’, p.17.

⁸¹⁶ Robertson, K. G., *Public Secrets: A Study in the Development of Government Secrecy* (London: MacMillan Press, 1982) p.53.

⁸¹⁷ MacNeil, ‘Trusting Records in a Postmodern World’, p.44.

⁸¹⁸ Maftai, M. & McAndrew, I., *Lineage of Elements Included in the Template for Analysis (pre-InterPARES): From Traditional Diplomats to Contemporary Archival Diplomats* (University of British Columbia, 2000) p.4.

⁸¹⁹ Ibid.

hazard, both in public life and in systems. In the digital environment, responsible data management relates directly to the *fides publica*.

Roberts describes the legal requirement in Sweden that all public authorities should maintain a public register of records, including those that are inaccessible due to legal restrictions. Under this requirement, information may be secret, but the information about the existence of the information is public. I have described the Norwegian equivalent, a digital register that unifies the metadata of documents created and received across public authorities.⁸²⁰ Applying the principle that there must be transparency in curation, if these Swedish and Norwegian registers are to be effective, it must be possible for users to understand what will and will not be included in them. Roberts states that many working documents are excluded from the Swedish register by a clause in the Freedom of the Press Act. It would take a certain level of critical information literacy for users to discover this. Attempting to find out if the Norwegian register details its operations and the information types that it excludes, I found the portal offline.⁸²¹

In a datafied world, Ranciere's impertinent dialectic will necessarily be acted out in examinations and critiques of data, perhaps even within the systems that manage and publish the data, such as the user engagement that is possible at data.gov.au. System design should allow us to answer back. However, recalling Feldman's observation that 'Filing delimits both the terrain of possibility and modes of objection',⁸²² there is a risk here that participation is seen as endorsement, and that such close engagement with systems would enact the *dispositif* of bureaucracy 'through every citizen'.⁸²³ The co-option of the citizen-bureaucrat as a quality controller for government information, enriching and improving the databanks of the state, is one possible outcome of participatory

⁸²⁰ Lowry, J., 'Opening Government: Open Data and Access to Information' in Lowry, J. & Wamukoya, J. (eds.), *Integrity in Government through Records Management: Essays in Honour of Anne Thurston* (Farnham: Ashgate, 2014).

⁸²¹ *Offentlig Elektronisk Postjournal (OEP)* <https://www.oep.no/?lang=en> [accessed 14 September 2018].

⁸²² Feldman, I., *Governing Gaza: Bureaucracy, Authority, and the Work of Rule, 1917–1967*, (Durham: Duke University Press, 2008) p.32.

⁸²³ Vismann, *Files*, p.110.

systems. Is the critique of data in fact Alquati's valorizing information: is it 'feeding the circuits of the whole factory'?⁸²⁴ Participatory systems can only become dissensual systems through a critical engagement with them. Not only do users need to be critical of the data they are seeing, but critical of the systems themselves. Virilio provides a warning that is relevant here:

Today, *control of the environment* is very largely supplanting the *social control* of the constitutional state and, to this end, it has to establish a new type of transparency: *the transparency of appearances instantaneously transmitted over a distance...* This is the meaning of the commerce of the visible, the latest form of "publicity".⁸²⁵

Are systems enabling engagement to be a form of control? For instance, through data matching, are organisations using user engagement to profile users, creating better formed data doubles? Sarangi and Slemrouck argue that the management of people is done through the management of information: the more fully formed the data double, the more knowable and manageable the person.⁸²⁶ One could swap 'body' for 'data double' in Butler's statement: 'Constituted as a social phenomenon in the public sphere, my body is and is not mine'.⁸²⁷ Furthermore, opening up information systems as fora for debate turns them into possible sites of dissensus that, like the streets, may be policed, curfewed and closed down. The state's monopoly on force lies behind the information system as it does behind the political system.

9.4. Conclusion

Analyses of data published on KODI and data.gov.au demonstrate the value of applying record-keeping controls to data curation processes. These controls can be broadly conceived of as the attribution of metadata, documented

⁸²⁴ Pasquinelli, M., 'To Anticipate and Accelerate: Italian Operaismo and Reading Marx's Notion of the Organic Composition of Capital,' *Rethinking Marxism*, 26.2 (2014) p.183.

⁸²⁵ Virilio, P., *The Information Bomb*, 2nd ed. (London; New York: Verso, 2005) p.61.

⁸²⁶ Sarangi, S. & Slemrouck, S., *Language, Bureaucracy, and Social Control* (New York: Routledge, 2014).

⁸²⁷ Butler, J., *Precarious Life: The Powers of Mourning and Violence*, (London; New York: Verso, 2004) p.82.

custodianship and auditable systems, but as shown in Chapter 6, the record-keeping profession has developed highly specific technical controls in order to protect the trustworthiness of information. The preceding discussion of the findings of the case studies shows how the rigour of these controls assures or calls into question the trustworthiness of open government data: the richness and consistency of the metadata, the documentation of the custodial / curatorial process and the openness of the systems directly affected data trustworthiness. It is not possible to define a universally applicable statement of what is 'good enough' as this will vary depending on users, uses and contexts, as well as with systems, processes and documentary requirements for contextualisation, but a scale exists, along which data becomes more or less like a record as the qualities that make it trustworthy are more or less present. Open government data can be made more like a record, and this helps to create assurances of trustworthiness.

It is clear that, with access, trustworthiness is an important feature of information that can redress information asymmetry.

It is not such an abrupt shift at this point to return to social contract, and to remind ourselves that the social contract and its apparatus of constraint was designed by the theorist just to cleanse the life of individuals of fraud, deception, and other such ills arising from men's hedonistic pursuit of their boundless desires. Society will - after the covenant - vouch for the match between intention and action. But how will it do this? One answer is as follows: through the archetypal speech act of contract, society establishes speech acts as a way of rendering states of mind transparent. That, in a sense, is the central work of community.⁸²⁸

In the next chapter, I will return to the social contract to consider the broader implications of archival and data management practices in open government data curation.

⁸²⁸ Turner, R., 'Speech and the social contract,' *Inquiry*, 28.1-4 (1985) p.50.

Chapter 10

Conclusion

The biopolitical project of knowing the population to manage it better was deeply compromised by the methods used to collect such data.⁸²⁹

Akhil Gupta, *Red Tape: Bureaucracy, Structural Violence and Poverty in India*

⁸²⁹ Gupta, A., *Red Tape: Bureaucracy, Structural Violence, and Poverty in India*, (Durham: Duke University Press, 2012) p.43.

10.1. Introduction

This thesis began with the premise that open government data is becoming an important mode of government openness, but one that has not yet fully addressed the need for data quality with particular respect to the trustworthiness of opened data. It has applied concepts from record-keeping, particularly archival-diplomatics, to open government data curation to understand if and how open data can be improved in relation to its trustworthiness. The research shows that the trustworthiness of government data has repercussions for the dynamic of the principal-agent relationship in the social contract. In this final chapter, I will summarise the argument of this thesis, draw conclusions about the intersection of record-keeping controls and open government data and suggest further research.

10.2. Summary of Thesis

In the chapter *Information and the Social Contract*, I argued that information is crucial to the operations of power. I argued for a conceptualisation of the ‘record-as-command’, a form of bureaucratic writing that enables power to operate at a distance, giving effect to the governance of people, resources and territories. Responses to records-as-commands form a dialogue between the governing and governed that is essential for the sustenance of that power relationship, which I illustrated through an analogy to the Pasolini film *Salo*. This recursive dialogue that happens through the exchange of written records sustains bureaucracies and helps to manufacture authority. I argued that bureaucracies are *dispositifs* that are sustained by a dialogic engagement with the governed that is enacted through records. This animating quality of information flows has been characterised by Alquati as control information – information flowing down from the system to the individual - and valorising information – ‘the flow running upstream and feeding the circuits of the whole factory’.⁸³⁰ The difficulty of opting out of this exchange of information lies in the role of information in characterising the individual in the bureaucratic mind: the ongoing record-making of the state constructs a data double of each of its subjects. This consumption of records by

⁸³⁰ Pasquinelli, M., 'To Anticipate and Accelerate: Italian Operaismo and Reading Marx's Notion of the Organic Composition of Capital,' *Rethinking Marxism*, 26.2 (2014) p.183.

the machineries of the state privileges a certain kind of record-making: the law 'operates not *in mundo* but in the medium of literality; it believes only what is written – more precisely, what it has itself written down'.⁸³¹ This suggests that the quality of authenticity is significant: there must be ways of assessing the validity of writing as a command or response.

The thesis then considers the nature of bureaucracies as apparatuses, or *dispositifs*, as conceived of by Foucault as heterogeneous ensembles of diverse elements. Bureaucracies are super-subjects composed of subjects. Bureaucracies, following Pasquinelli, are literally constituted in the bodies of workers, who are also its subjects. This boundary between the individual and the state is blurry. As I argue later in the thesis, this characteristic of the state is what allows the citizen to be subsumed when the citizen engages uncritically in the enrichment of open government data. Considering the implications of refusing to participate in this dialogical record-making, I argued that current trends and events in government information management, such as the increased use of e-government portals to centralise citizen data and the revelation of large scale state surveillance, are suggestive of the move from a society of discipline to a society of control, articulated by Deleuze. This is marked by a 'crisis of the institutions' and, later in the present chapter, I return to this idea to consider the archive in this crisis. In Chapter 2, I suggest that Ranciere's 'impertinent dialectic' is a viable response to this move into a society of control. Ranciere's notion is one in which a community of equals forms to assert its rights through direct means, creating a dynamic form of democracy that is always in a state of 'self-correction'. I asserted that: 'The impertinent dialectic is not possible without a change in the power dynamic that operates through bureaucracy and the record.' These arguments form a theoretical and political context for the description of information asymmetry in the third chapter.

In Chapter 3, I argued that information asymmetry is a defining feature of the social contract. I first suggested that the social contract is one of agency, and I used literature from the social sciences to describe agency theory's

⁸³¹ Vismann, C., *Files: Law and Media Technology*, trans. by Winthrop-Young, G. (Stanford: Stanford University Press, 2008) p.56.

conceptualisation of 'principal-agent problems'. I use Mitnick's typology of agent behaviours to show the significance of what Mitnick called 'policing' in reducing dissonance between an agent's 'self-goals' and 'other-goals' (goals of the principal). Before describing the means of aligning these goals that have been developed in the field of economics, I consider the characteristics of the principal-agent problem and in particular the six features of the canonical form of the problem as defined by Miller. I find that three of the core assumptions persist in the application of the model to the citizen-state relationship: the core assumptions (1) agent impact, (2) information asymmetry, and (3) asymmetry in preferences are applicable to a study of the social contract. I suggest that information asymmetry – an imbalance of information between agents and principals (conceived of in terms of quantity and quality) is the most significant feature of the model in its political application because the state operates through the transmission of records-as-commands, as established in Chapter 2.

Given the significance of information asymmetry for the social contract as a principal-agent problem, I briefly trace the development of the idea in information economics in the context of western economic thought before outlining the solutions to information asymmetry that have been proposed by information economists. These solutions fall into two categories: those that solve selection problems, or the problem of adverse selection, and those that solve incentive problems, or the problem of moral hazard. Selection problems concern information asymmetry before a contract is made. Incentive problems concern information asymmetry after a contract is made. Information economists have proposed monitoring and metering as solutions to these post-contract problems. Monitoring is any form of oversight of the agents by principals, while metering is the comparison of contract terms with observable deliverables. Looking at this typology of information problems and solutions, I finally argue that:

It is monitoring that facilitates accountability, in the sense of rendering an account. Post-contract, moral hazard is the primary problem arising from the information asymmetry that exists between individuals and the state, the principals and the agent, and monitoring becomes indispensable in redressing that information asymmetry.

This prompts the question: how is monitoring done in the principal-agent problem of the social contract? Chapter 4, *Opening Government: A History of Monitoring*, posits that open government data is the latest mechanism in a history of monitoring. After pointing to some more ancient roots of government secrecy, the chapter reviews the literature to provide a history of the development of a culture of secrecy within the British government from the Victorian period, when the present structures and cultures of public service began to take shape. UK government secrecy – ‘the British disease’ – is described in detail since, as the later Kenyan and Australian case studies describe, that disease was spread during the colonial period. The chapter then turns to archival access, considering the connection between access to archives and the character of the nation state, using record-keeping and political science literature to show a general association between openness and democracy, and secrecy and more autocratic forms of government. The chapter then provides an overview of the legal instruments devised to normalise archival transfers and control access, which is context for legal developments around whistleblowing and Freedom of Information. Regarding whistleblowing, I argue that it augments access to documented information by encouraging the reporting of events that occur outside of bureaucratic record-making processes, and that it reinforces other monitoring mechanisms by encouraging the reporting of breaches to documentary custodianship, record-keeping systems and record-making norms such as adequate and accurate metadata capture. Regarding Freedom of Information, again, the chapter provides a précis of the historical development of FOI in the UK, noting the barrier posed by the government culture of secrecy. I show that the UK’s FOI regime continues to be a site of contestation, both in terms of the adversarial nature of the FOI request process, and the boundaries of the regime itself, or what falls within the discoverable, e.g. non-government participants in the hollow state, or information that is undocumented because of the chilling effect.

The chapter then looks at open government data, noting that unlike FOI, there is not yet a comprehensive history of its development internationally or in the UK. The chapter points to some early landmarks in the emergence of the open data

trend from 2006, and notes the initial question about open data's viability as a cheaper alternative to FOI. I describe how the establishment of the Open Government Partnership in 2011 gave new impetus to open government data, and the work since 2012 of civil society actors interested in strengthening open data by aligning it with record-keeping frameworks and techniques. Finally, I describe the use of open data in civic technologies to realise the aspirations of the openness movement in relation to citizen participation through the use of opened public sector information. Overall, I argue that there has been an accelerating trend towards government openness, and I suggest that this raises a question about the trustworthiness of the information that individuals are being given access to.

In the fifth chapter, *Trustworthy Information*, I provide an historical overview of thinking about the trustworthiness of information, from Roman law through to common law legal systems in the feudal period, noting the importance of the role of witnesses to transactions such as the livery of seisin. I use the work of MacNeil to show that the introduction of written charters to replace the oral testimony of witnesses still recognised the necessarily public nature of significant transactions by recording the names of witnesses to the act and facilitating non-repudiation. Publicity, witnessing and non-repudiation are concepts that become significant in the later discussion of metadata for open data curation and publication.

The chapter describes the development of diplomatics as an approach to authenticating written records. I argue that this diplomatic and legal thinking forms the epistemological foundation for archival science, before reviewing how archival science has worked to understand trustworthiness in digital records. Trustworthiness is constituted by reliability, accuracy and authenticity. I note that accuracy and reliability relate to record-making, but in the use and preservation of records, the third constituent part of trustworthiness – authenticity – becomes key. According to the ontology, authenticity requires integrity (the record is complete and uncorrupted) and identity (the record is unique and distinguishable from other records). I argue that new constellations of citizen-actors, technology and government information platforms open up a connection between metering and reliability, such as in the case of crowdsourced data, but monitoring of

government through open data or other inspections of recorded information require judgments of authenticity. I review the InterPARES work on assurances of authenticity and find that three fundamental components emerge as controls over information that are necessary for authenticity to be assessed. These three components are the presence of metadata that documents the creation and management of the record, the operation of systems that control the record and capture metadata about its management over time, and custodianship, or the assignation of responsibility for every act of creation or intervention, which is documented by the system in the metadata. When these three elements are operationalised together, it becomes possible, in an open environment, to inspect and interrogate records, people and processes to judge authenticity.

This hypothesis is tested in Chapter 6, *Practices for Establishing and Maintaining Authenticity in Information*, which provides a detailed analysis of the Bengal Records Manual, a 1943 manual of registry procedures used by the colonial Government of Bengal. The system described by the Manual is a highly developed instantiation of a registry system, one that tightly controls the registration and movement of records and documents custodianship through the attribution of metadata across numerous locations. The provisions of the Manual are compared throughout with ICA-Req, an international standard – though one developed by experts predominantly trained in the Commonwealth record-keeping tradition – articulating functional requirements for digital record-keeping systems. The comparison shows that approaches to protecting authenticity in digital record-keeping systems have the same fundamental features as older paper systems – metadata capture, systematised controls, and documented custodianship. Aside from demonstrating that these three elements are essential for assurances of authenticity in recorded information, supporting the findings of Chapter 5, the comparison of the Manual and ICA-Req illustrates the practical methods of applying these controls to information. This prepares the way for the following case studies of Kenyan and Australian open government data initiatives, which are reviewed for evidence that the three basic controls are present in the curation and publication of the data.

Chapters 7 and 8 are case studies of open government data in Kenya and Australia respectively. The countries were chosen because they are both Commonwealth countries and share some of the legacies of British colonialism: importantly, for this study, common law legal systems, cultures of government secrecy and registry-based record-keeping practices. Together, the two case studies, it was hoped, would reflect differences in the trustworthiness of open data related to the presence or absence of record-keeping controls in the data curation and publication process.

Both case studies begin with an account of the development of open data in that country, from the culture of secrecy inherited from British administration, to moves to access to information through FOI and open data. Both case studies describe the involvement of civil society in this move, noting a much more active and vocal civil society engagement in the passage of FOI law in Kenya. Each case study then analyses a dataset published through the portals. In each case, the dataset concerns land because land has been the subject of datafication and record-making since ancient times, it was recorded and datafied across the empire using the same techniques, it was central to the colonial project and is now central to the work of post-colonial restitution and reparation, it is the subject of contracts, and it is coterminous with the nation states within which social contracts are enacted. Land information management in Kenya and Australia is explained before a particular dataset is analysed for signs of its documented custodianship and systematic control through metadata.

In Chapter 9, a comparison was made of the Kenyan and Australian datasets across the three categories of metadata, custodianship and systems. This comparison reiterated the significance of the three controls for assurances of authenticity. The case studies effectively demonstrated the consequences for the trustworthiness of open government data when one or more of these controls is weak or missing. Where metadata was missing or contradictory, doubts were raised about the identity and integrity, and therefore the authenticity, of the data. When custodianship throughout the life of the dataset could not be seen, there was room for suspicion about the curation of the data, and where the functioning of the systems involved in the curation and publication of the data could not be

observed, the trustworthiness of the data was open to question. The case studies, then, demonstrated the direct and impactful correlation between record-keeping controls and data trustworthiness.

The case studies suggested two further findings. Firstly, that data and records exist on the same scales of quality, richness and evidential weight: that data can build up into a record and that a record can be broken down into data through the addition or subtraction of metadata about context. The richer the metadata, the more like a record the data becomes, and that this, as the comparison of the case studies shows, means that open data can be made more like records. It should be possible to improve the trustworthiness of open government data by applying the kinds of record-keeping controls seen in the Bengal Records Manual and ICA-Req. Secondly, these controls must be visible to the user in order to effect a sense of trustworthiness. MacNeil describes photostats of records excluded as evidence because they were not a 'full rendition of the original material'.⁸³² It may not be possible to develop universally applicable guidance on what constitutes a full rendition of data, given that each dataset and curation process has parameters shaped by more or less unique requirements, uses and applications, but the principle of publicity of metadata, systems and custodianship improves trustworthiness. This helps to redress information asymmetry under the social contract.

Publicity in data curation also unlocks the political potential of open data beyond the immediate use of the information in question. It makes government information systems sites of dissensus. The dissensual government system, where information asymmetry is contested in specific cases and practical ways, opens new meanings and applications for information in the social contract. These are not necessarily dialogic, as in the recursivity of the record-as-command, but like the record-as-command, they have socio-political ramifications both for the individual and the body politic.

10.3. Conclusions

⁸³² MacNeil, H., *Trusting Records: Legal, Historical and Diplomatic Perspectives* (Springer, 2000) p.50.

As I have established, the individual body and the body politic are juridical entities and they are datafied. The notion of the juridical person is very clearly articulated by Duranti:

In a diplomatic context, as well as in a legal one, persons are the subject of rights and duties; they are entities recognized by the juridical system as capable of or having the potential for acting legally. Just as collections or successions of human beings (juridical persons) derive their qualification of persons from a legislative authority, individual human beings are not automatically persons because they are human, but become such if rights and duties are ascribed to them. Every full citizen is a person; subjects may be persons; but human beings having no rights are not persons. The person is the legal subject or substance of which the rights and duties are attributes.⁸³³

This personhood and its attendant rights and duties are created through documentation, as established in Chapter 2. As legal and political systems have evolved and become more complex, the documentation of the person and their rights and duties has become richer, technologies of record-making and record-keeping responding to the informational needs of power and opening up new possibilities for control in a recursive cycle of development. Open government data is among the phenomena, with big data, biometrics, surveillance and counterintelligence technologies, that reveal the recent acceleration of the historical process of the datafication of persons. This is only an extension of past practices. Writing about examinations, Foucault claimed that they ‘introduced individuality into the field of documentation... The examination that places individuals in a field of surveillance also situates them in a network of writing; it engages them in a whole mass of documents that capture and fix them’.⁸³⁴ As information technologies have developed, this kind of documentation has become more thorough and more easily aggregated, as Roberts explained when he wrote that

⁸³³ Duranti, L., *Diplomatics: New Uses for an Old Science*, (Lanham; London: Scarecrow Press, Inc., 1998) p.83.

⁸³⁴ Foucault, M., *Discipline and Punish: The Birth of the Prison*, trans by Sheridan, A., 2nd ed. (New York: Vintage Books, 1995) p.189.

much personal data has been public for many years, but ‘the ability to collect and use this information has been constrained because traditionally it has been recorded on paper and stored by a large number of state and local governments’.⁸³⁵ The increasing richness of the data available from records, government and other systems, and technologies, and the new data that can be created by their combination and comparison, together with the data individuals push into the public sphere through social media platforms and other forms of online content creation – the richness of this data is widely recognised by data subjects, businesses and governments alike. This richness means that our data doubles are taking shape more clearly, particularly where there are moves to aggregate, share and combine datasets, as established in Chapter 2.

As our data doubles move through the digital environment, our physical bodies move through datafied space. The infrastructure for Guattari’s ‘monitored city’ is being realised through the use of ICT and the Internet of Things to build ‘smart cities’; the ODI has argued that beyond technology and infrastructure, open data is a vital component of smart cities.⁸³⁶ Virilio sees electronic tagging of prisoners and the use of mobile telephones to blur the line between work and private life as the datafication of everyday life that signals a transition to Deleuze’s societies of control.⁸³⁷ When every site, person and activity can be conceived of as a data point, element or set, governance becomes datafied and public sector information becomes the primary apparatus of sovereign power to an extent that exceeds the written orders of kings (after Vismann) or the information gathering of the colonial and totalitarian regimes of the modern period. Access to that government data has and will have significant political consequences.

Datafication recreates the problem of the individual versus the general will in the digital environment. In the same way that Anaxagoras of Clazomenae cannot opt out of the general will in Athens, it is not possible to opt out digitally, or as Raley puts it ‘...should we not also be able to opt out of all advertising databases or indeed out of the whole system of “cybernetic capitalism” itself? But it is arguably

⁸³⁵ Ibid., p.208.

⁸³⁶ Open Data Institute, *Smart Cities*, <https://theodi.org/smart-cities> [last accessed 18 October 2018].

⁸³⁷ Virilio, P., *The Information Bomb*, 2nd ed. (London; New York: Verso, 2005) p.66.

the case that exit in the form of forgetting or genuine anonymity is no longer possible, that disappearance itself has disappeared'.⁸³⁸ Notwithstanding the attempt of the General Data Protection Regulations to inaugurate a 'right to be forgotten', opting out of 'cybernetic capitalism' or in any case out of the datafied digital environment is to move outside of the *demos*, outside of participation. Brunton and Nissenbaum argue that '...obfuscation is a tool particularly suited to the "weak" – the situationally disadvantaged, those at the wrong end of asymmetrical power relationships. It is a method, after all, that you have reason to adopt if you can't be invisible – if you can't refuse to be tracked or surveilled, if you can't simply opt out or operate within professionally secured networks'.⁸³⁹ Obfuscation is a critical response to control through datafication, but one that requires an understanding of the 'rules of the game' or the technologies of information management and control, which returns to the problem of information asymmetry. When the individual and public space are both datafied, a control environment is established; one that is asymmetrical. The power dynamic that operates through the record-as-command comes to operate through data-as-command, which is more granular and blurs the already fuzzy boundaries between subjects and supersubjects, states, bureaucracies and polities. Diverse technologies enable this datafied social contract, and civic technologies in particular deserve consideration as dissensual technologies because of their promise of participatory government.

In 1973, Michael Spence wrote that if the 'incentives for veracity in reporting anything by means of a conventional signaling code are weak, then one must look for other means by which information transfers take place'.⁸⁴⁰ Open government data is an alternative to earlier means of information transfer, such as the archival access and Freedom of Information described in Chapter 4, though each mechanism has its own features and applications. The use of that data opens another set of alternatives.

⁸³⁸ Raley, R., 'Dataveillance and Countervailance', in Gitelman, L. (ed.) *Raw Data is an Oxymoron*, (Cambridge: MIT Press, 2013) p.130.

⁸³⁹ Brunton, F. & Nissenbaum, H., *Obfuscation: A User's Guide for Privacy and Protest*, (Cambridge: MIT Press, 2015) p.9.

⁸⁴⁰ Spence, M., 'Job Market Signaling,' *The Quarterly Journal of Economics*, 87.3 (1973) pp.355–74.

As I have noted earlier, information asymmetry extends to the ‘rules of the game’. Many of those who engage with government may not necessarily understand its modes of operation. As Roberts observes, even ‘in countries with long-established disclosure laws, making a request for information requires knowledge about the bureaucratic routine by which information requests are processed and about the legal provisions that should govern decision on the release of information’.⁸⁴¹ As established in Chapter 4, the open government movement conceives of participation in government as fundamental to its mission: the Open Government Declaration commits signatories to support civic participation, for instance.⁸⁴² Hazell, Worthy and Glover studied the documents and events that were crucial in the development of the UK’s FOI Act to discern the key aims of the Act. They found that the aims most frequently cited in official documentation were:

- Increasing the openness and transparency of government
- Increasing the accountability of government
- Improving the quality of government decision-making
- Improving public understanding of government
- Increasing public trust in government
- Increasing public participation in government.⁸⁴³

These aims correspond with those found in Mendel’s 2008 international study of over 70 pieces of FOI legislation.⁸⁴⁴ Participation is linked with information access, accountability and trust throughout. The political significance of open information is set out plainly by Hetherington: ‘If authoritarianism was built on the state’s control of information, then democracy is to be built on by giving citizens

⁸⁴¹ Roberts, A., *Blacked Out: Government Secrecy in the Information Age* (New York: Cambridge University Press, 2008) p.117.

⁸⁴² Open Government Partnership, *Open Government Declaration* <https://www.opengovpartnership.org/open-government-declaration> [last accessed 18 October 2018].

⁸⁴³ Hazell, R., Worthy, B. & Glover, M., *The Impact of the Freedom of Information Act on Central Government in the UK: Does FOI Work?* (London: Palgrave MacMillan, 2010) p.18.

⁸⁴⁴ Mendel, T., *Freedom of Information: A Comparative Legal Study*, UNESCO, 2008, <http://unesdoc.unesco.org/images/0015/001584/158450e.pdf> [last accessed 18 October 2018].

access to, and indeed control over and responsibility for, all information.⁸⁴⁵ The use of government information in fostering participatory democracy has been acted out through civic technologies. I have previously defined civic technologies as:

...technology that intersects with public life, enabling engagement or participation of the public with the government for a variety of purposes, from making government more transparent and accountable, to enhancing civic communities, and impacting policy decisions.⁸⁴⁶

Open government data has been used to fuel these technologies. Civic technologies draw the public into a closer and ideally more responsive relationship with government in a variety of ways. These technologies allow users to engage directly with governments (fixmystreet.com), to monitor agents (theyworkforyou.com) and bypass government altogether (the work of Chief Kariuki, described below). Like any information system that allows dialogue or user created content, civic technologies are platforms for dissensus and for co-option, depending on how they are used. Feldman wrote about the way that procedure-based systems of filing curtailed alternatives to a controlled form of dialogue, writing that 'Reiterative authority relies on regularity and on an expansive view of the bureaucratic domain. Files and civil servants accrued and deployed authority in part through a process that seemed to offer no alternative, no other place to go.'⁸⁴⁷ Civic technologies, especially those designed by non-government actors, open up new informational spaces for political action.

Edward Higgs noted that democratic participation eroded 'something of the mystique of public authority'⁸⁴⁸ and civic technologies, and open data in general, have the potential to erode the 'secrecy and mystery' Marx attributed to

⁸⁴⁵ Hetherington, K., *Guerrilla Auditors: The Politics of Transparency in Neoliberal Paraguay*, (Durham: Duke University Press, 2011) p.6.

⁸⁴⁶ Lowry, J., *Open Government Data Literature Review*, InterPARES Trust https://interparestrust.org/assets/public/dissemination/EU02_20151210_OpenGovernmentDataLiteratureReview_FinalReport.pdf [last accessed 18 October 2018].

⁸⁴⁷ Feldman, I., *Governing Gaza: Bureaucracy, Authority, and the Work of Rule, 1917–1967*, (Durham: Duke University Press, 2008) p.220.

⁸⁴⁸ Higgs, E., *The Information State in England: The Central Collection of Information on Citizens since 1500* (Palgrave Macmillan, Basingstoke, UK, 2004) p.169.

bureaucracy. This effect is predicated on a critical civil society. The power of civil society in opening government is clear in Kenya, where non-government organisations and individual advocates lobbied government for greater access to information, while establishing outlets and applications for the use of data. Roberts stated that 'The question in countries that are now adopting disclosure laws is not only whether government has the capacity to fulfill the law, but whether the nongovernmental sector will have the capacity to use the law effectively'.⁸⁴⁹ There are numerous examples around the world of non-government actors working to use access laws for the public good, such as the Transactional Records Access Clearinghouse at Syracuse University, which uses US Freedom of Information legislation to compile and publish data on government spending. At a more grass roots level, the guerilla auditors documented by Hetherington use rights afforded by the law to analyse government documents, through which they 'unleash undisciplined interpretations and hence novel possibilities into the situations that the documents purport to describe'.⁸⁵⁰ Free and widely accessible education that supports critical data literacy, then, is essential for giving effect to the laws and technologies that support information access. In Kenya, following the achievements around KODI and access legislation, civil society is turning its support to implementation, and a large part of that is educating the public of its rights and the 'rules of the game'.

An extension of the participatory government enabled by civic technologies and open government data, is a kind of parallel governance fuelled by crowdsourced data that is acted on at the community level. Of the alternatives made available by open government data, perhaps the most radical departure from the controlled speech of the file is found in the kind of work being coordinated by Chief Kariuki, described in Chapter 7. Community data collection in Lanet Umoja produced baseline data on access to clean water that Kariuki used to secure a donation of water filters that were installed in approximately 3000 households. This initiative, while coordinated by Kariuki, both a leader in the traditional system of social organisation and an administrative chief formally part of the nation-wide system of government, occurred outside of the established, legally mandated citizen /

⁸⁴⁹ Roberts, *Blacked Out*, p.119.

⁸⁵⁰ Hetherington, *Guerrilla Auditors*, p.9.

state dynamic. Rather than engaging in separate and different dialogues channeled through government files, in the hope that the bureaucracy would fix each problem or identify a communal problem across the hypothetical corpus of complaint letters, the people of Lanet Umoja documented the water problem independently. This data informed conversations with the donors of the water pumps, rather than informing government budget-setters and procurement officers and setting in motion new chains of correspondence, the creation of new files and the accumulation of an archive. It is unclear what forms of documentation were created by this process that ran in parallel to formal government procedures, and beyond the original crowdsourced data, the contracts, duties and rights inaugurated at each conversation, order, gift or installation may or may not be documented. In this case, then, metering, rather than monitoring, is possible in that the resolution of the problem is demonstrable, but perhaps un(der)documented.

Parallel forms of governance and civic activity, then, explode the formal dialogue of the file, unleashing new possibilities for participation, information symmetry and ultimately power. These new constellations of information resources, conversations, datasets and processes may be radical and generative, but they may also be problematic in environments where legal rights are at issue and the law does not have as evidence 'what it has, itself, written down'. There is a role here for information professionals and technologists to work in these emergent and evolving information ecosystems to ask when records are becoming, where the theoretical construct of the file might be constituted, and how the metadata, custodianship and systems that link the pieces together can be implemented or improved. One example is in the need to link the identity of each person or instrument that gathers data during crowdsourcing and introduce a non-repudiation function so that the reliability of the data can be linked to the reliability of the person or thing that introduced it into the dataset. In the decentralised environment of the modern office, actions by staff members can be documented through ID numbers and passwords. In more complex settings, or in data gathering practices that span organisations, sectors, actors, nations, etc., linking activities to persons is less straightforward. There must be a system that enables digital signing or sealing. MacNeil reported how early thinking within the

InterPARES project recognised that procedural controls should stand in for more explicit attestations, such as signatures.⁸⁵¹ What procedural controls are in place for the curation of open data? Revisiting Raffel's four measures for controlling record-writers, it is plain to see that crowdsourcing projects might easily address, as methodological questions, 1) restricting the privilege of record-writing, 2) requiring signatures from participants, 3) giving responsibility to each participant for reporting only a portion of a fact, and/or increasing the number of those who report the same fact, and 4) making the same record serve a number of different users, so that the writer cannot tailor the message to the audience.⁸⁵² This suggests that MacNeil's assertion, prior to the advent of open data, that 'the limits of contemporary archival diplomatics as a body of concepts and methods for ensuring the trustworthiness of electronic records are inseparable from the limits of the Weberian model of bureaucracy on which those concepts and methods are built' is no longer accurate.⁸⁵³ Conceivably, archival-diplomatic techniques are applicable across all forms of social and informational organisation.

Where the role for information professionals and technologists becomes problematic is in circumstances where custodianship is not politically apposite. Crowdsourced data can be weaponised against states and forgery, Gupta has posited, '...is a creative, positive act (one forges ahead) that challenges the right of the state to authenticate, to adjudicate between the true and the false. Forgery can be a subaltern strategy to undermine the authority of state writing.'⁸⁵⁴ Perhaps this is the clearest statement of the principle that data curation is implicitly political, that to the 'extent that stamp papers, degrees from certified educational institutions, signatures of notaries, government seals, and watermarks authenticate, they create a parallel economy of the inauthentic: the forged and the counterfeit, which constantly threaten, simultaneously, to undermine the authoritativeness of state writing and to multiply its effects and effectiveness'.⁸⁵⁵ This echoes Monroe's assertion that systems always contain

⁸⁵¹ MacNeil, H., 'Providing Grounds for Trust: Developing Conceptual Requirements for the Long-Term Preservation of Authentic Electronic Records,' *Archivaria* 50.1 (2000) p.67.

⁸⁵² Duranti, *Diplomatics*, p.71.

⁸⁵³ MacNeil, *Trusting Records*, p.111.

⁸⁵⁴ Gupta, *Red Tape*, p.227.

⁸⁵⁵ *Ibid.*, p.230.

the means of their own demise.⁸⁵⁶ Crowdsourced data, false and obfuscating data, forged records can all be used for, against or by the state. It is the tactical deployment of record-keeping systems that makes them spaces for dissensus.

This is also true of the archive, which has traditionally been seen as a site of and surrogate for authority. The legal and archival theory that has developed around questions of reliability and authenticity over hundreds of years anticipated a certain kind of record-making. Record-making processes were often linear, formulaic, transactional or narrative, and as Vismann has shown, dialogic. In newer processes of data gathering, compilation, curation and publication, individuals acting alone or in concert across variously structured and changeable networks, using any number of tools and systems from paper forms to scientific instruments, may aggregate data that may or may not be considered a record, or may hover somewhere near that conceptual boundary. Do the traditional approaches to understanding the trustworthiness of information still work?

In discussing rules for assessing the facts of a case in trials in England in the twelfth and thirteenth centuries, MacNeil noted that the ‘jurors were men of the neighbourhood and it was assumed that they would know the facts and incorporate their own knowledge in their verdict’.⁸⁵⁷ This is essentially the premise of open data, where judgments of validity are left to the people ‘of the neighbourhood’. In the case of crowdsourcing, it may be the authors of the data who must judge its reliability. Duranti has argued that reliability has traditionally hinged on ‘(i) the degree of control exercised on the procedure of creation, (ii) the degree of control exercised on the authors, and (iii) the degree of completeness of the documents themselves’.⁸⁵⁸ In the case of a crowdsourced dataset, these considerations are still valid, though ‘control’ ought to be conceived of as an agreement over (i) processes, (ii) roles and (iii) the parameters of the dataset. Similarly, authenticity continues to be assessable against the controls discussed in Chapters 5 and 6. How do the tools of data collection and processing work? What have the curators of the data done with it? If it is stored somewhere, has it

⁸⁵⁶ Monroe, A., *Interrogation Machine: Laibach and NSK* (Cambridge; London: MIT Press, 2005) p.8.

⁸⁵⁷ MacNeil, *Trusting Records*, p.7.

⁸⁵⁸ Duranti, L., ‘Archives as Place’, *Archives and Manuscripts* 24.2 (1996) pp.243-244.

changed over time, and how? These things, the people 'of the neighbourhood' can only judge if the controls are transparent, and they know how to read them. This recalls the guerilla auditors of Paraguay: 'To be sure, they read the content, looking for inconsistencies, lies, or elisions in the text. But they also read as auditors might, analyzing the traces of documentary transactions, precisely the stamps and signatures I found so uninteresting. How, they always asked, was this document made?'⁸⁵⁹

It has been argued that authentication of information occurs at the crossing of the 'archival threshold'. Citing Roman law and the *ius archivi* theorists of seventeenth century Germany, Duranti defines the archival threshold as 'the space where the officer of the public authority takes charge of the documents, identifies them by their provenance and class, associates them intellectually with those that belong in the same aggregation, and forwards them to the inside space'.⁸⁶⁰ The discourse about the trustworthiness of records has led to a narrative about the archival threshold, the line where authentication took place, and the idea of a reliance on archival institutions for the validation of records. MacNeil writes that

As a trusted custodian of records, ancient archival institutions sustained and lent credibility to contractual relationships between citizens. They also lent credibility to the implicit social contract between citizens and the state by preserving the records of the state's past actions on the basis of which the state could be held to account.⁸⁶¹

Beyond that threshold, the systematic and documented custodianship of information continues in perpetuum. In bureaucracies before computerisation, one might find this transmission between separate entities of creator and keeper within organisations in the forms of the author / officer and the registry clerk, and within governments between ministries, departments and agencies, and state archives. In today's digital environment, this manifests within organisations in the author / officer and capture into a digital records management system, where the

⁸⁵⁹ Hetherington, *Guerilla Auditors*, p.148.

⁸⁶⁰ Duranti, L., 'Archives as Place', p.244.

⁸⁶¹ MacNeil, H., 'Providing Grounds for Trust', p.71.

'Declare as Record' function serves as the non-repudiation function and links the character of the author / submitter to the reliability of the record, and within governments through the transmission of digital records from ministries, departments and agencies to the state archives' digital repository. Instead of existing at the door of an institution, the archival threshold may be found in numerous places within offices, systems or networks.

Records created in more or less traditional forms according to established processes of command and response may still follow this trajectory. However, after open data, the declaration of authenticity is no longer made by the crossing of the threshold into archival space. Instead, information can only be authenticated by the crossing out of the institution and into public space. This is the reversal of the archival threshold. The archive is inverted. Authentication is now done in public. *Fides publica* is not a state of being but a continual process of inspection in public space.

Relatedly, information that has passed out of official custody and into the public sphere through unauthorised means constitute a distinct category of government information. Records made public under access laws are often redacted, and as Blouin and Rosenberg have suggested, the 'black ink of the redactor explicitly reduces the legalities of state control to a relationship of trust.'⁸⁶² It is not only the absence of redactions that is suggestive of trustworthiness in leaked documents: the forms, stamps and signatures that are used to authenticate documents in formal contexts are made available for public inspection. It is this spectacle of bureaucracy laid bare that inclines us to accept the validity of leaked documents without question. The public is much less likely to demand to see audit trails and system reports for leaked records not only because there is no expectation of having access to them, but because the illicit nature of the release is suggestive of having broken into that archival space, imbued, as it is, with authenticity. A leaked document is read as an exposure: '...the leak offers a mockery of state power by reproducing its words verbatim in the public arena...'⁸⁶³ Gitelman

⁸⁶² Blouin, F. & Rosenberg, W., *Processing the Past: Contesting Authority in History and the Archives* (New York: Oxford University Press, 2011) p.180.

⁸⁶³ Gitelman, L., *Paper Knowledge: Toward a Media History of Documents*, (Durham: Duke University Press, 2014) p.106.

argues that the recontextualisation of the document that happens through the leaking process makes the document into a kind of parody: ‘...a leaked copy has the potential not only to transgress or leak across the inside-outside boundary but also, importantly, to mirror – one might say technologically to reproduce – its iteration as a form of critique’.⁸⁶⁴ This critical recontextualisation, the exposure of the authors, is a kind of witnessing of a secret speech – speech previously confined to the file but now made public, along with the speaker. This is essentially a performance of the concept that underpins the US business records exception to the hearsay rule, which holds that records can be deemed self-authenticating if they have been used by an organization as the basis of its operations, decisions, transactions, etc. In other words, if the organisation has treated the record as authentic, it may be read that way by the law. ‘In both archival theory and jurisprudence, records that the creator relies on in the usual and ordinary course of business are presumed authentic’.⁸⁶⁵ We trust leaked records because they appear to have been in use by bureaucracies. Can we presume that datasets are authentic on this basis? If erroneous data is a defence for poor policy decisions, the documentation of the provenance of government data becomes legally significant.

The transition from the society of discipline to the society of control, according to Deleuze, requires a ‘crisis of the institutions’ that, as discussed in Chapter 2, is already in process. For Deleuze, this is seen across systems, for example with respect to the corporate system, he wrote about ‘new ways of handling money, profits and humans that no longer pass through the old factory form’ which may now be with us in the form of distributed ledger technology, zero hour contracts and the tools of the so-called ‘sharing economy’. If there is, or recently has been, a crisis of the institutions, how has it manifested in the informational system, and in particular what does it mean for the institution of the archive? Weld has pointed to some of the things that state archives are, claiming that archival thinking ‘demands that we see archives not only as sources of data... but also as more than the sum of their parts – as instruments of political action, implements of

⁸⁶⁴ Gitelman, *Paper Knowledge*, p.95.

⁸⁶⁵ Authenticity Task Force, *Requirements for Assessing and Maintaining the Authenticity of Electronic Records* (InterPARES, 2002) p.2
http://www.interpares.org/book/interpares_book_k_app02.pdf [last accessed 18 October 2018].

state formation (“technologies of rule”), institutions of liberal democratization, enablers of gaze and desire, and sites of social struggle’.⁸⁶⁶ The idea of the crisis of the institutions concerns not so much what the institutions are, though, but how they work. For government information to be trusted by the public, the inner workings of the institution – conceived of as the sum of the archival or curatorial operations around data – must be exposed. The reversal of the archival threshold is the sign not only of a crisis of public trust in government, but a sign of the crisis of the institutions, which must transition to new modes of working. The reversal is merely a new mode of performing the same control function: abandoning *fides publica* in the institution of the archive and dragging the records out into the agora. This supports Hetherington’s assertion that:

...the practices of representation that go into creating transparency are saturated at every turn with precisely that aspect of social life that they are meant to get rid of: politics. Indeed, far from stabilizing information, these larger technical networks create new spaces for disagreement and contestation.⁸⁶⁷

Beyond the changed requirements for authentication, the archive is affected by the transition to the control society through its repurposing as a distributed databank that connects to an increasingly interlocking set of systems. The panoptical potential of technology requires the datafication of all actors, objects, environments and events, as described above. ‘The conception of a control mechanism, giving the position of any element within an open environment at any given instant... is not necessarily one of science fiction’.⁸⁶⁸ The crisis of the institutions, then, for the archive, is a promotion from the authenticating repository of documents of public acts to a databank on which all other operations draw. In 2006, Roberts observed persistent limits to openness in context, structure, cultures and technologies of government.⁸⁶⁹ Roberts argues that a networked approach to governance, which aims to share information across agencies, can

⁸⁶⁶ Weld, K., *Paper Cadavers: The Archives of Dictatorship in Guatemala*, (Durham: Duke University Press, 2014) p.13.

⁸⁶⁷ Hetherington, *Guerilla Auditors*, p.7.

⁸⁶⁸ Deleuze, G., ‘Postscript on the Societies of Control’, *October*, Vol. 59 (Winter, 1992) p.7.

⁸⁶⁹ Roberts, *Blacked Out*, p.18.

strengthen external barriers to disclosure. This makes system design and interoperability political, like data practices. Open government data, analysis and correction by the public, and even data practices that seem to bypass state control all potentially enrich this databank. Control of that databank and access to it are essential for power, and the object of a permanent dissensual dialogue.

In the time that it has taken to write this thesis, the open data community has started to recognise the significance of record-keeping principles for the quality of open data. This recognition has largely been of the value of the documentation of provenance, but the engagement with the concept has been limited. In May 2018, the Open Data Institute published a report called *Registers and Collaboration: Making Lists We Can Trust*.⁸⁷⁰ It provides advice on preparing and maintaining registers for public inspection. In some respects, the report accords with record-keeping understandings of concepts such as authority, completeness and accuracy. In other ways, the report reinvents concepts well established in record-keeping. For instance, custodianship is called 'stewarding' in the report. The report concludes, as this dissertation has, that transparency in custodianship is vital to establishing trustworthiness in information:

Trust in registers is built by high-quality stewarding practices. While much of the activity of a custodian comprises ensuring that the register is kept up to date and accurate, it is also the responsibility of the custodian to make the governance processes surrounding a register transparent and open.⁸⁷¹

What the open government data community has not yet identified is the mechanisms for giving effect to its nascent aspirations for data trustworthiness. This thesis has shown how concepts and practices developed throughout a long history of record-keeping can be introduced into data curation for that purpose.

This research has presented new analyses of open government data, looking in detail at issues of trustworthiness in the data presented through government

⁸⁷⁰ Miller, D. & Roe, S., *Registers and Collaboration: Making Lists We Can Trust*, Open Data Institute, May 2018 <https://theodi.org/article/registers-and-collaboration-making-lists-we-can-trust-report/> [last accessed 22 October 2018].

⁸⁷¹ *Ibid.*, p.21.

portals. The research has brought this open data work together, for the first time, with detailed analyses of record-keeping practices to show how the two lines of work could be brought together to improve assurances of trustworthiness in government information. As the first application of archival-diplomatics to open government data, the research has found that comprehensive metadata, documented custodianship and auditable systems could be employed to make data curation practices more robust and the resultant datasets more trustworthy. It has elucidated links across systems and processes that are contingent, though they result in seemingly unproblematic datasets, or as Ribes and Jackson put it, ‘...the work of sustaining massive repositories reveals only a thin slice in the long chain of coordinated action that stretches back directly to a multitude of local sites and operations through which data in their “raw” form get mined, minted, and produced’.⁸⁷² The research has investigated the technical issues in such sequences of curatorial and custodial interventions, and proposed an archival-diplomatic approach to fixing them. The research has also set these technical issues in their broader socio-political context, showing that there is a direct relationship between data curation and information asymmetry, which, problematically, characterises the social contract. As Brunton and Nissenbaum have written, ‘What we have called power asymmetries map closely onto traditional vectors of power – wealth, social class, education, race, and so forth. In today’s data-driven societies, epistemic or information asymmetries are highly consequential’.⁸⁷³ In order for information asymmetry to be corrected in the social contract, critical approaches to data curation are needed. So too are critical approaches to reading data for signs of appropriate curation, so that information systems can be dissensual. This would constitute an informational kind of impertinent dialectic, ongoing and permanently correcting.

10.4. Critical Data Literacy: The Need for Further Research

⁸⁷² Ribes, D. & Jackson, S., ‘Data Bite Man: The Work of Sustaining a Long-Term Study’ in Gitelman, L. (ed.), *Raw Data is an Oxymoron* (Cambridge: MIT Press 2013) p.152.

⁸⁷³ Brunton, F. & Nissenbaum, H., *Obfuscation*, (Cambridge: MIT Press, 2015) p.79.

...facts are ontological, evidence is epistemological, data is rhetorical. A datum may also be a fact, just as a fact may be evidence. But, from its first vernacular formation, the existence of a datum has been independent of any consideration of corresponding ontological truth. When a fact is proven false, it ceases to be a fact. False data is data nonetheless.⁸⁷⁴

These findings show that the introduction of record-keeping controls into data curation practices has the potential to radically improve the trustworthiness of open government data. Additionally, it has found that *critical* data literacy is essential to interrogating data practices to make informed judgments about trustworthiness. This is a rich area for further research.

Critical data literacy can be conceived of as an ability to interrogate data for its accuracy, reliability and authenticity, but firstly, it is the ability to understand how to access data. Higgs attributed widening public concern for official information from the 1960s onwards in part to the 'sheer proliferation of contacts between citizens and civil servants that took place in the twentieth century. Given the amount of form filling that now became part of everyday life, it was impossible for citizens to be unaware of the amount of material held by government officials'.⁸⁷⁵ Even if government data gathering has gathered pace alongside information technologies, the visibility of that data gathering may not necessarily have also increased. This form-filling and data retention is arguably obscured by some data gathering techniques and technologies, where data is derived from disparate systems, processed algorithmically and stored in undisclosed locations. As the information economics literature shows (Chapter 3), information asymmetry often extends to knowledge of the 'rules of the game'. 'The act of inspecting records at the town office or courthouse imposed its own costs: the knowledge needed to determine the location of the records; the time needed to visit the office; the nerve often needed to request even information that is, as a settled matter of law, publicly accessible'.⁸⁷⁶ Critical data literacy, then, begins with this basic knowledge of rights over data, locations of data, and the procedures and

⁸⁷⁴ Rosenberg, D., 'Data before the Fact' in Gitelman, L. (ed.) *Raw Data is an Oxymoron* (Cambridge: MIT Press, 2013) p.18.

⁸⁷⁵ Higgs, E., *The Information State in England*, p.168.

⁸⁷⁶ Roberts, *Blacked Out*, p.226.

technologies of access. Information professionals have a role to play in raising awareness of these rights, locations, procedures and technologies: each one of these needs to be opened up to use.

Following this, critical data literacy concerns the interrogation of data, which requires an understanding of the curation practices involved in producing the data, and an awareness of the importance of the presence of record-keeping controls in those practices. This involves not only record-keeping, but a range of professions and activities that are often disparate and siloed. The work of statisticians needs to be better understood by data users and those who would teach them critical data literacy. As Gupta suggested, 'if one does not critically analyze how those statistics are collected and employed, then the relationship between biopower and issues like sovereignty and violence is likely to be misunderstood'.⁸⁷⁷ Statistics may be derived from records, and illuminating the chain between the statistic, the dialogical and valorising record, and the individual, Gupta suggested that, if 'one may derive from Foucault the suggestion of an inherent affinity between the biopolitical and the statistical, then forms are the critical modality by which one is converted into the other'.⁸⁷⁸ As the case studies have shown, these intersections (the completion of the form, the extraction of the form's data) can be weak points in the double chain of custody and documentation. Critical data literacy involves understanding record-making methods and statistical methods, or knowing to enquire about them. And there are other intersections. Statistics derived from scientific instruments open up another field of knowledge that becomes relevant to critical data use; design justice is implicated.

Critical data literacy also involves a whole set of skills employed by data analysts, journalists and activists to make use of the data in ways that allow open data to open government. A literature exists around information literacy, and there is relevant work in librarianship and research data management, such as work around data citation practices. These strands have yet to be pulled together in ways that prepare individuals for the critical use of open government data. The

⁸⁷⁷ Gupta, *Red Tape*, p.42.

⁸⁷⁸ *Ibid.*, p.145.

cross-disciplinary nature of data curation means that effective data literacy programmes must introduce concepts from a range of fields; record-keeping is one of these, but there has been little engagement in open government data from that professional community. Records professionals must not only seek to build controls into data curation practices, they must teach others to look for the presence or absence of these controls when using data. The social and political benefits of open data cannot be enjoyed without this literacy and its application, which is why further research is needed in this area.

10.5 Conclusion

In conclusion, I have shown how the record-keeping profession has defined trustworthiness in information and that record-keeping controls, particularly those associated with archival-diplomatics in the three broad categories of 1) custodianship documented in 2) metadata by 3) auditable systems, can be applied to data curation practices to improve the trustworthiness of open government data, thereby positively affecting information asymmetry in the social contract. This is evident in my analysis of open government data initiatives in Kenya and Australia. In addition to this primary finding, I have: shown that information asymmetry is the defining feature of the principal-agent problem that is the social contract, because of the agency of the record-as-command; expanded the record-as-command into the concept of data-as-command, to show how the same power dynamic operates in the datafied environment, in a manner that further blurs the boundaries between the subject and the supersubject of the state; shown that this imbues data, and information management systems and platforms, with a political potentiality – they become sites of disensus and a means for an impertinent dialectic; shown that, in this context, trustworthiness in open government data becomes important for disensus and the social contract, since it deeply affects information asymmetry; suggested that open government data and the associated civic technologies that make use of it have the potential to further affect the social contract by enabling parallel governance, where communities use data to inform or effect change; further suggested that the same principles from archival-diplomatics used in this research could be used to

strengthen crowd-sourced and community-managed data; shown that the premise of open government data reverses the archival threshold so that the authenticating function is performed by the *demos*, not by the public institution of the archive, which challenges the orthodox archival theory that enabled this finding; and identified the necessity of critical data literacy amongst data users if the sociopolitical power of open government data is to be used in the ongoing correction of information asymmetry.

BIBLIOGRAPHY

- Abioye, A., 'Milestones in archives administration in Nigeria', *Comma*, 2012.1 (2013) pp.15-26.
- Abuya, E., 'Towards Promoting Access to Information in Kenya', *African Network of Constitutional Lawyers* (April 2011)
https://spaa.newark.rutgers.edu/sites/default/files/files/Transparency_Research_Conference/Papers/Abuya_Edwin.pdf
- Acker, A., 'When is a Record: A Research Framework for Locating Electronic Records in Infrastructure', in Gilliland, A. J., McKemmish, S. & Lau, A. J. (eds.) *Research in the Archival Multiverse* (Clayton, Australia: Monash University Publishing, 2017).
- Adams, J. 'Probing the East German State Security Archives,' *International Journal of Intelligence and Counter Intelligence*, 13.1 (2000) pp.21–34.
<http://doi.org/10.1080/088506000304925>
- Advisory Group on Reform of Australian Government Administration, *Ahead of the Game: Blueprint for the Reform of Australian Government Administration* (2010) <http://apo.org.au/system/files/20863/apo-nid20863-24401.pdf>
- Akerlof, G. A., 'The Market for 'Lemons': Quality Uncertainty and the Market Mechanism,' *The Quarterly Journal of Economics* 84.3 (1970) pp.488–500.
doi:10.2307/1879431
- Alt, J., E. B. de Mesquita, and S Rose 'Disentangling Accountability and Competence in Elections: Evidence from U.S. Term Limits,' *Journal of Politics* 73.1 (2011) pp.171–86, doi:10.1017/S0022381610000940
- Apaza, C. R., & Chang, Y., 'Effective Whistleblowing Conceptual Framework', in Apaza, C. R. & Y Chang, Y., (eds.) *Whistleblowing in the World: Government Policy, Mass Media, and the Law* (Cham: Palgrave Macmillan, 2017).

- Arrow, K., 'The Economics of Information: An Exposition,' *Empirica* 23.2 (1996) pp.119–28. doi:10.1007/BF00925335
- Arrow, K., 'Information and the Organization of Industry' in Chichilnisky G., (ed.) *Markets, Information and Uncertainty: Essays in Economic Theory in Honor of Kenneth J. Arrow* (Cambridge: Cambridge University Press, 1999).
- Arrow, K. J., & Debreu, G., 'Existence of an Equilibrium for a Competitive Economy,' *Econometrica* 22.3 (1954) pp.265–90 doi:10.2307/1907353
- Arthur, C., & Cross, M., 'Give us back our crown Jewels,' *The Guardian* (9 March 2006) <https://www.theguardian.com/technology/2006/mar/09/education.epublic>
- Article 19, *Eastern Africa*, (n.d.) <https://www.article19.org/regional-office/eastern-africa/>
- Ashworth, S., 'Reputational Dynamics and Political Careers,' *Journal of Law, Economics, and Organization* 21.2 (2005) pp.441–66 doi:10.1093/jleo/ewi015
- Awortwi, N., 'Contracting Out Local Government Services to Private Agents: An Analysis of Contract Design and Service Delivery Performance in Ghana,' *International Journal of Public Administration* 35.13 (2012) pp.886–900 doi:10.1080/01900692.2012.686033
- Australian Bureau of Statistics, *ABS.Stat beta*, (n.d.) <http://stat.data.abs.gov.au/Index.aspx>
- Australian Bureau of Statistics, *Land Management Practices (2011-12)* <https://data.gov.au/data/dataset/land-management-practices>
- Australian Government, Australian Law Reform Commission, *Secrecy Laws and Open Government in Australia: Report 112* (December 2009) <https://www.alrc.gov.au/sites/default/files/pdfs/publications/ALRC112.pdf>

Australian Government, Department of the Prime Minister and Cabinet, *Midterm Self-Assessment Report for Australia's First Open Government National Action Plan 2016-2018*

https://www.opengovpartnership.org/sites/default/files/Australia_Midterm-Self-Assessment_2016-2018.pdf

Australian Government, Office of the Australian Information Commissioner, *About Us* <https://www.oaic.gov.au/about-us/>

Australian Government, Office of the Australian Information Commissioner, *Australian Information Commissioner Act* <https://www.oaic.gov.au/about-us/who-we-are/australian-information-commissioner-act>

Australian Government, Office of the Australian Information Commissioner, *De-identification Decision-Making Framework* <https://www.oaic.gov.au/agencies-and-organisations/guides/de-identification-decision-making-framework>

Australian Government, Office of the Australian Information Commissioner, *FOI Guide*, <https://www.oaic.gov.au/freedom-of-information/foi-resources/foi-guide/part-a-freedom-of-information-past-and-present>

Australian Government, Office of the Australian Information Commissioner, *History of the Privacy Act* <https://www.oaic.gov.au/about-us/who-we-are/history-of-the-privacy-act>

Australian Government, Office of the Australian Information Commissioner, *Our Information Publication Scheme* <https://www.oaic.gov.au/about-us/access-our-information/our-information-publication-scheme/>

Australian Government, Office of the Australian Information Commissioner, *Principles on Open Public Sector Information* <https://www.oaic.gov.au/resources/information-policy/information-policy-resources/principles-on-open-public-sector-information.pdf>

Australian Government, 'Use Cases', *data.gov.au* <https://data.gov.au/showcase>

Australian Open Government Civil Society Network, *About*,
<https://opengovernment.org.au/about/>

Bakunin, M., *Selected Writings*. (London: Jonathan Cape, 1973)

Banakar, R., 'In Search of Heimat: A Note on Franz Kafka's Concept of Law,' *Law & Literature*, 22 (2010) pp.463-490

Banisar, D., *Whistleblowing International Standards and Developments*. Paper presented at the Primera Conferencia Internacional Sobre Corrupcion y Transparencia, Ciudad de Mexico (23-25 March 2006).

Barrett, M. & Ling Kent, J., 'Inside the government agency designing tech to fight fake news', *NBC*, (19 April 2018). <https://www.nbcnews.com/tech/tech-news/inside-government-agency-designing-tech-fight-fake-news-n865586>

Bastian, J. A., 'Taking Custody, Giving Access: A Postcustodial Role for a New Century', *Archivaria*, 53.1 (2002) pp.76-93.

Bastiat, F. *Selected Essays on Political Economy*. (Irvington-on-Hudson: Foundation for Economic Education, 1995)

Bannister, J., Appleby, G., Olijnyk, A., & Howe, J., *Government Accountability: Australian Administrative Law* (Cambridge: Cambridge University Press, 2015)

BBC News, 'The Interview: A guide to the cyberattack on Hollywood', 29 December 2014 <https://www.bbc.co.uk/news/entertainment-arts-30512032>

Bearman, D., 'An Indefensible Bastion: Archives as a Repository in an Electronic Age,' *Archives and Museum Informatics Technical Report 13* (1991) pp.14-24.

- Bedos-Rezak, B., 'Seals and Sigillography, Western European,' in Strayer, J. (ed.) *Dictionary of the Middle Ages*, vol. 11 (New York: Charles Scribners Sons, 1988)
- Besley, T., *Principled Agents?: The Political Economy of Good Government* (Oxford: Oxford University Press, 2007)
- Birkinshaw, P., *Reforming the Secret State* (Milton Keynes: Open University Press, 1990)
- Blouin, F. X., 'Convergences and Divergences in Archival Tradition: A North American Perspective,' *Second European Conference on Archives, Proceedings* (Paris: International Council on Archives, 1989)
- Blouin, F. X., & Rosenberg, W., *Processing the Past: Contesting Authority in History and the Archives* (New York: Oxford University Press, 2011)
- Board of Revenue, Government of Bengal, *The Bengal Records Manual 1943*, (Alipore: Bengal Government Press, 1943)
- Bowcott, O., 'High court rules data retention and surveillance legislation unlawful', *The Guardian* (17 July 2015)
<http://www.theguardian.com/world/2015/jul/17/data-retention-and-surveillance-legislation-ruled-unlawful>
- Brevini, B., 'Metadata Laws, Journalism and Resistance in Australia', *Media and Communication*, 5.1 (2017) pp.76–83
- Brooke, C. N. L., 'The Teaching of Diplomatic,' *Journal of the Society of Archivists*, 4.1 (1970) pp.1-9 doi:10.1080/00379817009513930
- Brown, G., *Why Kenya's open data portal is failing – and why it can still succeed*,
<https://sunlightfoundation.com/2013/09/23/why-kenyas-open-data-portal-is-failing-and-why-it-can-still-succeed/>

Brunton, F. & Nissenbaum, H., *Obfuscation: A User's Guide for Privacy and Protest* (Cambridge: MIT Press, 2015)

Burke, C., *America's Information Wars: The Untold Story of Information Systems in America's Conflicts and Politics from World War II to the Internet Age* (Lanham: Rowman and Littlefield, 2018).

Butler, J., *Precarious Life: The Powers of Mourning and Violence* (London; New York: Verso, 2004).

Cabinet Office, United Kingdom, *UK Open Government National Action Plan 2011 to 2013*, <https://www.gov.uk/government/publications/uk-open-government-national-action-plan-2011-to-2013>

Cabinet Office, United Kingdom, *UK Open Government National Access Plan 2016-2018*, <https://www.gov.uk/government/publications/uk-open-government-national-action-plan-2016-18/uk-open-government-national-action-plan-2016-18>

Calland, R., & Dehn, G. (eds.) *Whistleblowing Around the World: Law, Culture, and Practice* (London: Public Concern At Work and the Open Democracy Advice Centre, 2004).

Campaign for Freedom of Information, *About Us*, <https://www.cfoi.org.uk/about/>

Canes-Wrone, B., & Shotts, K. W., 'When Do Elections Encourage Ideological Rigidity?' *American Political Science Review* 101.2 (2007) pp.273–88.
doi:10.1017/S0003055407070256

Carranza, J., *East Africa's Open Data Revolution*
<http://www.data4sdgs.org/master-blog/2016/9/7/east-africas-open-data-revolution>

Central Land Council, Australia, *The History of the Central Land Council*,
<https://www.clc.org.au/articles/cat/our-history/>

Central Land Council, Australia, *Functions and Responsibilities*,
<https://www.clc.org.au/index.php?/articles/info/what-land-councils-must-do/>

Chancellor of the Duchy of Lancaster, *Open Government White Paper* (London:
Her Majesty's Stationery Office, 1993)

Chandler, A., *Strategy and Structure: Chapters in the History of the American
Industrial Enterprise* (Washington: Beard Books, 1962)

Chapman, R. A. & Hunt, M. (eds.) *Open Government* (New York: Croom Helm,
1987)

Clanchy, M. T., *From Memory to Written Record: England 1066-1307* (London:
Edward Arnold Publishers, 1979)

Coats, J. C., 'Applications of Principal-Agent Models to Government Contracting
and Accountability Decision Making,' *International Journal of Public
Administration* 25.4 (2002) pp.441–61. doi:10.1081/PAD-120013250

Cook, J., *The Price of Freedom* (Sevenoaks: New English Library, 1985).

Cook, T., "'We Are What We Keep; We Keep What We Are": Archival Appraisal
Past, Present and Future,' *Journal of the Society of Archivists*, 32.2 (2011)
pp.173–189 <http://doi.org/10.1080/00379816.2011.619688>

Cook, T. & Schwartz, J. M., 'Archives, records, and power: From (postmodern)
theory to (archival) performance,' *Archival Science*, 2.3 (2002) pp.171–185.
<http://doi.org/10.1007/BF02435620>

Darch, C. & Underwood, P., *Freedom of Information and the Developing World:
The Citizen, the State and Models of Openness* (Oxford: Chandos, 2010)

Data Documentation Initiative (n.d.) <http://www.ddialliance.org>

Data Science Africa (n.d.) <http://www.datascienceafrica.org/>

Davies, W., 'Weaponising Paperwork,' *London Review of Books*, Vol 40 Number 9 (10 May 2018).

Davies, T., *Open Data Policies and Practice: An International Comparison*, (European Consortium for Political Research, 2014)
<https://ecpr.eu/Filestore/PaperProposal/d591e267-cbee-4d5d-b699-7d0bda633e2e.pdf>

Defence and Security Media Advisory (DSMA) Committee, *DA Notices*,
<http://www.dsma.uk/danotices/index.htm>

Deleuze, G., 'Postscript on the Societies of Control,' *October*, 59, (1992) pp.3–7.
<http://doi.org/10.2307/778828>

Deleuze, G., *Foucault* (London: Bloomsbury Academic, 2006)

Deleuze, G. & Guattari, F., *Kafka Toward a Minor Literature*, 9th ed. (Minneapolis: University of Minnesota Press, 1986)

The Democratic Society, *Developing a Model for Open Government in the EU*, (2016) http://www.demsoc.org/wp-content/uploads/2016/12/Demsoc.OSEPI_EUOpenGovModel.pdf

Densley, J., 'Street Gang Recruitment. Signaling, Screening, and Selection,' *Social Problems* 59.3 (2012) pp.301–21

Department of Planning, Transport and Infrastructure, Government of South Australia, *What is the Torrens Title System?* (2008)

[https://www.landservices.sa.gov.au/1public/Education -
Careers/ Documentary Movies/4What is the Torrens Title System.asp](https://www.landservices.sa.gov.au/1public/Education-_Careers/_Documentary_Movies/4What_is_the_Torrens_Title_System.asp)

Derrida, J., *Archive Fever: A Freudian Impression* (Chicago: University of Chicago Press, 1998)

DLM Forum <http://moreq.info>

DLM Forum, *Other Specifications*, <http://moreq2.eu/other-specifications>

Dryden, J., 'Standards: News, Progress Reports and Reviews', *Journal of Archival Organization*, 8.3 (2010) pp.174-266

Drucker, J., *What Is? Nine Epistemological Essays* (Berkeley: Cuneiform Press, 2013)

Duff, W. M. & Harris, V., 'Stories and names: Archival description as narrating records and constructing meanings,' *Archival Science*, 2.3-4 (2002) pp.263–285 <http://doi.org/10.1007/BF02435625>

Duncan, G., *The High Court of Delegates* (Cambridge: Cambridge University Press, 1971)

Duranti, L., 'Archives as Place,' *Archives and Manuscripts* 24.2 (1996) pp.242-255

Duranti, L., *Diplomatics: New Uses for an Old Science*, (Lanham; London: Scarecrow Press, Inc., 1998)

Duranti, L., Eastwood, T., & MacNeil, H., *Preservation of the Integrity of Electronic Records* (Dordrecht: Kluwer Academic Publishing, 2002)

Duranti, L. & Franks, P. C. (eds.), *Encyclopedia of Archival Science* (Lanham: Rowman and Littlefield, 2015)

- Duranti, L. & MacNeil, H., 'The Protection of the Integrity of Electronic Records: An Overview of the UBC-MAS Research Project', *Archivaria*, 42 (Fall 1996) pp.46-67
- Eisenhardt, K. M., 'Agency Theory: An Assessment and Review,' *The Academy of Management Review* 14.1 (1989) pp.57–74 doi:10.2307/258191
- Executive Board of the United Nations Development Programme, the United Nation Population Fund and the United Nations Office for Project Services, *Management response to the evaluation of the role of UNDP in supporting national achievement of the Millennium Development Goals DP/2015/20* (2015) <http://undocs.org/en/DP/2015/20>
- Feldman, I., *Governing Gaza: Bureaucracy, Authority, and the Work of Rule, 1917–1967* (Durham: Duke University Press, 2008)
- Finn, P. & Horwitz, S., 'US charges Snowden with espionage', *The Washington Post*, 21 June 2015 https://www.washingtonpost.com/world/national-security/us-charges-snowden-with-espionage/2013/06/21/507497d8-dab1-11e2-a016-92547bf094cc_story.html?utm_term=.6e05592680a7
- Freedom of Information Independent Review Panel, *The Right to Information: Reviewing Queensland's Freedom of Information* (2008) <http://unpan1.un.org/intradoc/groups/public/documents/apcity/unpan032299.pdf>
- Foucault, M., 'The Subject and Power,' *Critical Inquiry*, 8.4 (1982) pp.777-795.
- Foucault, M., *Power/Knowledge: Selected Interviews and Other Writings, 1972-1977*, trans. by Gordon, C., Marshall, L., Mepham, J. & Soper, K. (New York: Pantheon Books, 1980)

- Foucault, M., *Discipline and Punish: The Birth of the Prison*, trans by Sheridan, A., 2nd ed. (New York: Vintage Books, 1995)
- Fox, P. M., 'The story behind the Torrens System,' *Australian Law Journal* 489 (1950).
- Fulton Committee, *The Civil Service, Vol. 1: Report of the Committee 1966-1968, Cmnd. 3638*, (London: Her Majesty's Stationery Office, 1968)
<http://www.civilservant.org.uk/library/fulton/fulton8.pdf>
- Funder, A, *Stasiland* (London: Granta, 2003)
- Galloway, A. R., *Protocol: How Control Exists After Decentralization* (Cambridge: MIT Press, 2004)
- George, A. & Bennett, A., *Case Studies and Theory Development in the Social Sciences* (Cambridge: MIT Press, 2005).
- Giddens, A., *The third way and its critics* (Cambridge: Polity Press, 2000)
- Gilliland, A. J., 'The Wink that's Worth a Thousand Words' in, Smit, F., Glaudemans, A. & Jonker, R., (eds.), *Archives in Liquid Times* (Stiching Archiefpublicaties, 2017) pp.212-227.
- Gilliland, A. J. & McKemmish, S., 'The Role of Participatory Archives in Furthering Human Rights, Reconciliation and Recovery', *Atlanti: Review for Modern Archival Theory and Practice*, 24.1 (2014) pp.78-88
- Gilliland, A. J. & McKemmish, S., 'Rights in Records as a Platform for Participative Archiving' in Cox, R.J., Langmead, A. & Mattern, E. (eds.) *Archival Education and Research: Selected Papers from the 2014 AERI Conference* (Sacramento: Litwin Press, 2015) pp.355-385.

Giry, A., *Manuel de Diplomatique*, trans. and quoted in Wigmore, J. H., *Evidence in Trials at Common Law, Volume 7*, ed. and rev. by Chadbourn, J., (Boston: Little, Brown, 1978) para. 2161.

Gitelman, L. & Jackson, V., 'Introduction', in Gitelman, L., (ed.) *"Raw Data" is an Oxymoron*, (Cambridge: MIT Press, 2013).

Gitelman, L., *Paper Knowledge: Toward a Media History of Documents* (Durham: Duke University Press, 2014).

Global Partnership for Sustainable Development Data, *Data4SDGs Toolbox*
<http://www.data4sdgs.org/initiatives/data4sdgs-toolbox>

Goody, J., *The Logic of Writing and the Organization of Society* (Cambridge: Cambridge University Press, 1986)

Government 2.0 Taskforce, Australia, *Engage Getting on with Government 2.0* (2009)
<https://www.finance.gov.au/sites/default/files/Government20TaskforceReport.pdf?v=1>

Government Digital Service, *data.gov.uk*, (n.d.) <https://data.gov.uk/about>

Government of Estonia, *E-Estonia*, (n.d.) <https://e-estonia.com>

Government of South Australia, Department of Planning, Transport and Infrastructure, *South Australian Integrated Land Information System: Access to Land Property Information* (n.d.)
http://www.sa.gov.au/data/assets/pdf_file/0003/156567/Access-Land-and-Property-Information.PDF

Greene, N., 'Salo: The Refusal to Consume' in Rumble, P. & Testa, B. (eds.) *Pier Paolo Pasolini: Contemporary Perspectives* (Toronto: University of Toronto Press, 1994) pp.232–42

Gregory, A., 'Open Data and Metadata Standards: Should We Be Satisfied with "Good Enough"?', Open Data Foundation (16 June 2011)

<http://odaf.org/papers/Open%20Data%20and%20Metadata%20Standards.pdf>

The Guardian, *Free Our Data*, (n.d.) <http://www.theguardian.com/technology/free-our-data>

Guilbert, K., 'Smartphone-wielding civilians to combat war crimes – with an app', Thomson Reuters Foundation (7 June 2015)

<http://news.trust.org/item/20180607151711-ymjan/>

Gunasekara, G., 'Whistle-blowing: New Zealand and UK solutions to a Common Problem', *Statute Law Review*, 24.1 (2003) pp.39-62

Gunnlaugsdottir, J., 'Government secrecy: public attitudes toward information provided by the authorities,' *Records Management Journal*, Vol. 25 Iss 2 (2015) pp.197–222

Gupta, A., *Red Tape: Bureaucracy, Structural Violence, and Poverty in India*, (Durham: Duke University Press, 2012)

Haggerty, K. & Ericson, R., 'The Surveillant Assemblage', *British Journal of Sociology*, 51:4 (2000) pp. 605-622.

Half, G. & Gregory, A., 'What Is Public Relations to Society? Toward an Economically Informed Understanding of Public Relations,' *Public Relations Review* 41.5 (2015) pp.719-725 doi:10.1016/j.pubrev.2015.06.007

Ham, F. G., 'Archival Strategies for the Post-Custodial Era', *The American Archivist*, 44:3 (1981) pp.207-216.

Harris, V., 'The archival sliver: Power, memory, and archives in South Africa,' *Archival Science*, 2.1 (2002) pp.63–86 <http://doi.org/10.1007/BF02435631>

- Hazell, R., Worthy, B. & Glover, M., *The Impact of the Freedom of Information Act on Central Government in the UK: Does FOI Work?* (London: Palgrave MacMillan, 2010)
- Head, R., 'A Comparative Case-Study Approach to Historical Archives in Europe,' in Gilliland, A., McKemmish, S. & Lau, A. (eds.) *Research in the Archival Multiverse* (Melbourne: Monash University Publishing, 2016) pp.433-455.
- Head, R., 'Documents, Archives, and Proof around 1700', *The Historical Journal*, 56.4 (2013) pp.909-930
- Healy, S., 'ISO 15489 Records Management: its development and significance,' *Records Management Journal*, Vol. 11 No. 2 (2001) pp.133-42
- Heemsbergen, L. & Daly, A., 'Leaking Boats and Borders: The Virtue of Surveilling Australia's Refugee Population,' *Surveillance & Society*, 15.3-4 (2017) pp.389-396
- Held, D., *Models of Democracy* (Stanford: Stanford University Press, 2006)
- Herbst, J., *States and Power in Africa: Comparative Lessons in Authority and Control* (Princeton: Princeton University Press, 2000)
- Hetherington, K., *Guerrilla Auditors: The Politics of Transparency in Neoliberal Paraguay* (London: Duke University Press, 2011)
- Higgs, E., *The Information State in England: The Central Collection of Information on Citizens since 1500* (Basingstoke: Palgrave Macmillan, 2004)
- Hobbes, T., *Leviathan* (Harmondsworth: Penguin Classics, 1982)

- Hooper, D., *Official Secrets: The Use and Abuse of the Act*, (London: Secker and Warburg, 1987)
- Hull, M. S., *Government of Paper: The Materiality of Bureaucracy in Urban Pakistan* (Berkeley, Los Angeles, London: University of California Press, 2012)
- Hurley, C., 'Parallel Provenance: What, if Anything, is Archival Description?' *Archives and Manuscripts*, 33.1 (2005) pp.110-145
- Hurley, C. 'The Australian 'Series' System: An Exposition' in McKemmish, S. & Piggott, M. (eds.) *The Records Continuum: Ian Maclean and Australian Archives First Fifty Years* (Clayton: Ancora Press, 1994)
- Information and Data Manager, *TRIM Does It Again* (2006)
<https://idm.net.au/blog/002207tower's-trim-does-it-again>
- International Commission of Jurists, *About*, <https://www.icj.org/about/>
- International Council on Archives, *Code of Ethics*, (Paris: International Council on Archives, 1996) https://www.ica.org/sites/default/files/ICA_1996-09-06_code%20of%20ethics_EN.pdf
- International Council on Archives, *Principles and Functional Requirements for Records in Electronic Environments - Module 2: Guidelines and Functional Requirements for Electronic Records Management Systems* (Paris: International Council on Archives, 2008)
- International Council on Archives, *Principles and Functional Requirements for Records in Electronic Office Environments: Overview of Implementation Guidance and Training Products* (Paris: International Council on Archives, 2013)

International Records Management Trust, *Kenya Country Report* (August 2011)
<http://www.irmt.org/portfolio/managing-records-reliable-evidence-ict-e-government-freedom-information-east-africa-2010-%e2%80%93-2011>

International Standardisation Organisation, *ISO 14721:2012 Space data and information transfer systems – Open archival information system (OAIS) – Reference model* (Geneva: International Standardisation Organisation, 2012)

International Standardisation Organisation, *ISO 16175:2012 - Information and documentation - Principles and functional requirements for records in electronic office environments*, (Geneva: International Standardisation Organisation, 2012)

International Standardisation Organisation, *SIO/IEC 27037:2012 Information technology – Security techniques – Guidelines for identification, collection, acquisition, and preservation of digital evidence*. (Geneva: International Organization for Standardisation, 2012)

Internet Society, *Africa Internet History: Highlights* (n.d.)
https://www.internetsociety.org/wp-content/uploads/2017/09/history_internet_africa.pdf

InterPARES Project <http://interpares.org>

InterPARES, *The InterPARES Glossary: A controlled vocabulary of terms used in the InterPARES Project* (University of British Columbia, 2002)
<http://www.interpares.org/documents/InterPARES%20Glossary%202002-1.pdf>

InterPARES 1 Authenticity Task Force, 'Appendix 2 - Requirements for Assessing and Maintaining the Authenticity of Electronic Records', in InterPARES 1 Project, *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project* (InterPARES Project, 2002)
http://www.interpares.org/book/interpares_book_k_app02.pdf

InterPARES 1 Preservation Task Force, 'Appendix 6 – How to Preserve Authentic Electronic Records', in *InterPARES 1 Project, The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project* (InterPARES Project, 2002)
http://www.interpares.org/book/interpares_book_o_app06.pdf

InterPARES 2 Terminology Database,
(http://www.interpares.org/ip2/ip2_terminology_db.cfm)

Isaacs, D., 'Are healthcare professionals working in Australia's immigration detention centres condoning torture?', *Journal of Medical Ethics* 42 (2016) pp.413-415

Jeitschko, T. D. & Kim, B. C., 'Signaling, Learning, and Screening Prior to Trial: Informational Implications of Preliminary Injunctions,' *The Journal of Law, Economics, & Organization* 29.5 (2013) pp.1085–1113
doi:10.1093/jleo/ews015

Jenkinson, H., *The English archivist: a new profession: being an inaugural lecture for a new course in archive administration delivered at University College London, 14 October, 1947* (London: HKLewis, 1948)

Jenkinson, H., *A Manual of Archive Administration*, (London: Percy Lund Humphries, & Co LTD., 1937)

Jensen, M. C. & Meckling, W. H., 'Theory of the firm: Managerial behavior, agency costs and ownership structure,' *Journal of Financial Economics*, 3.4 (1976) pp.305–360. [http://doi.org/10.1016/0304-405X\(76\)90026-X](http://doi.org/10.1016/0304-405X(76)90026-X)

Jimerson, R. C., *Archives Power: Memory, Accountability, and Social Justice* (Chicago: American Library Association, 2009)

Jouvenel, B., *On Power* (Carmel: Liberty Fund, 1993)

Kang, C., 'A Tweet to Kurt Eichenwald, a Strobe and a Seizure. Now, an Arrest', *New York Times*, <https://www.nytimes.com/2017/03/17/technology/social-media-attack-that-set-off-a-seizure-leads-to-an-arrest.html>

Kenny, J. C., *Secrecy Provisions: Policy and Practice*, (2011)
<http://www.fedcourt.gov.au/digital-law-library/judges-speeches/justice-kenny/kenny-j-20110324>

Kenya Open Data Initiative (KODI) <http://www.opendata.go.ke>

Ketelaar, E., 'Cultivating archives: meanings and identities,' *Archival Science*, 12.1 (2011) pp.19–33 <http://doi.org/10.1007/s10502-011-9142-5>

Kitchin, R., *The Data Revolution: Big Data, Open Data, Data Infrastructures and their Consequences* (London: Sage, 2014)

Kittler, F., *Discourse Networks 1800/1900*, (Stanford: Stanford University Press, 1990)

Knight Foundation, *The Emergence of Civic Tech: Investments in a Growing Field* (December 2013)
https://www.knightfoundation.org/media/uploads/publication_pdfs/knight-civic-tech.pdf

Knowles, D. D., *The Historian and Character: And Other Essays* (Cambridge: Cambridge University Press, 2008)

Koo, J. & Chou, C. H., 'PDF to PDF/A: Evaluation of Converter Software in Digital Repository Workflow', *New Review of Information Networking*, 18.1 (2013) pp.1-15.

Kosinski, M., Stillwell, D., & Graepel, T., 'Private traits and attributes are predictable from digital records of human behavior,' *Proceedings of the*

- National Academy of Sciences*, 110.15 (2013) pp.5802–5805
<http://doi.org/10.1073/pnas.1218772110>
- Koucky, J. (ed.) *Second European Conference on Archives: Proceedings* (Paris: International Council on Archives, 1989)
- Krauss, R., 'Two Moments from the Post-Medium Condition,' *October*, 116 (2006) pp.55-62.
- Laffont, J-J. & Martimort, D., *The Theory of Incentives: The Principal-Agent Model* (Princeton; Oxford: Princeton University Press, 2002)
- Lamothe, M., 'Redesigning the Hollow State: A Study of Florida Child Welfare Service Reform through the Lens of Principal-Agent Theory,' *International Journal of Public Administration* 34.8 (2011) pp.497–515
doi:10.1080/01900692.2011.582622
- Lange, T., 'A Theory of the Firm Only a Microeconomist Could Love? A Microeconomist's Reply to Lubatkin's Critique of Agency Theory,' *Journal of Management Inquiry* 14.4 (2005) pp.404–6 doi:10.1177/1056492605280419
- Lappin, J., 'What will be the next records management orthodoxy,' *Records Management Journal*, 20.3 (2010) pp.252-264
- Laserfiche, *Why You Need to Care About DoD 5015.2*,
<https://www.laserfiche.com/ecmblog/why-you-need-to-care-about-dod-5015-2/>
- Latour B., *Aramis, Or, The Love of Technology*, trans. by Porter, C. (Cambridge: Harvard University Press, 1996) <http://hdl.handle.net/2027/heb.01146>
- Law, J. & Smullen, J. (eds.) *A Dictionary of Finance and Banking*, 5th ed., (Oxford: Oxford University Press, 2014)

- Lebo, T., Sahoo, S. & McGuinness, D. *PROV-O: The PROV Ontology*. World Wide Web Consortium (30 April 2013). <https://www.w3.org/TR/prov-o/>
- Lemieux, V., *Competitive Viability, Accountability and Record Keeping: A Theoretical and Empirical Exploration Using a Case Study of Jamaican Commercial Bank Failures*, doctoral thesis, University College London, 2001.
- Lesperance, G., 'Beginning to Think about Salò' in Iannucci, A. (ed.) *Dante, Cinema, and Television*, pp.97-105 (Toronto: University of Toronto Press, 2004)
- Library of Congress, *Sustainability of Digital Formats: Planning for Library of Congress Collections*
<http://www.digitalpreservation.gov/formats/fdd/fdd000125.shtml>
- Lihoma, P., 'The Influence of the One Party Regime on Archives in Malawi: 1964 to 1994' in Lowry, J. & Wamukoya, J. (eds.) *Integrity in Government through Records Management* (Farnham: Ashgate, 2014)
- Lovering, T., 'British Colonial Administrations' registry systems: a comparative study of Northern Rhodesia and Nyasaland', *Archival Science* 10.1 (2010) pp.1-23.
- Lowry, J., 'Correlations between ICT and records policy integration and court case management system functionality', *Records Management Journal*, 23.1 (2013) pp.51–60
- Lowry, J., 'Opening Government: Open Data and Access to Information' in Lowry, J. & Wamukoya, J. (eds.) *Integrity in Government through Records Management: Essays in Honour of Anne Thurston* (Farnham: Ashgate, 2014)
- Lowry, J., *Open Government Data Literature Review* (InterPARES Trust, 2015)
https://interparestrust.org/assets/public/dissemination/EU02_20151210_OpenGovernmentDataLiteratureReview_FinalReport.pdf

Lubatkin, M. H., 'A Theory of the Firm Only a Microeconomist Could Love,'
Journal of Management Inquiry 14.2 (2005) pp.213–16
doi:10.1177/1056492605275352

Lyotard, J.-F., *The Postmodern Condition: A Report on Knowledge*, trans. by
Bennington, G. & Massumi, B. (Manchester: Manchester University Press,
1984)

Mabillon, J., *De re diplomatica* (1681)

Mabo v Queensland (No 2) [1992] HCA 23 [http://www8.austlii.edu.au/cgi-
bin/viewdoc/au/cases/cth/HCA/1992/23.html](http://www8.austlii.edu.au/cgi-bin/viewdoc/au/cases/cth/HCA/1992/23.html)

MacAskill, E., Borger, J., Hopkins, N., Davies, N., & Ball, J., 'GCHQ taps fibre-
optic cables for secret access to world's communications,' *The Guardian* (21
June 2013) [http://www.theguardian.com/uk/2013/jun/21/gchq-cables-secret-
world-communications-nsa](http://www.theguardian.com/uk/2013/jun/21/gchq-cables-secret-world-communications-nsa)

MacCormack, P., *Posthuman Ethics* (Farnham: Ashgate, 2012)

MacNeil, H., *Trusting Records: Legal, Historical and Diplomatic Perspectives*
(Dordrecht; London: Kluwer Academic, 2000)

MacNeil, H., 'Providing Grounds for Trust: Developing Conceptual Requirements
for the Long-Term Preservation of Authentic Electronic Records,' *Archivaria*
50.1 (2000) pp.52-78

MacNeil, H., 'Trusting Records in a Postmodern World', *Archivaria*, 51.1 (2001)
pp.36-47.

MacNeil, H., 'Providing Grounds for Trust II: The Findings of the Authenticity Task
Force of InterPARES', *Archivaria*, 54 (2002) pp.24-58

MacNeil, H., 'Trusting Description: Authenticity, Accountability, and Archival Descriptive Standards', *Journal of Archival Organization*, 7.3 (2009) pp.89-107

MacNeil, H., 'Trust and professional identity: narratives, counter-narratives and lingering ambiguities', *Archival Science*, 11.3-4 (2011) pp.175-192

MacNeil, H. & Mak, B., 'Constructions of Authenticity,' *Library Trends*, 56.1 (2007) pp.26-52

Maer, L. & Gay, O., *Official Secrecy* (London: House of Commons Library, 2008)
<https://fas.org/irp/world/uk/secrecy.pdf>

Maftai, M. & McAndrew, I., *Lineage of Elements Included in the Template for Analysis (pre-InterPARES): From Traditional Diplomatics to Contemporary Archival Diplomatics* (Vancouver: University of British Columbia, 2000).

Magaya, M. & Lowry, J., 'The National Archives of Tanzania Fifty Years After Wright's Report on the Government Records and Public Archives of Tanganyika', *Comma: Journal of the International Council on Archives*, 2012.1 (2013) pp.47-57

Manning, C. E., 'The years since I was jailed for releasing the 'war diaries' have been a rollercoaster,' *The Guardian* (27 May 2015)
<https://www.theguardian.com/commentisfree/2015/may/27/anniversary-chelsea-manning-arrest-war-diaries>

Map Kibera, *Methods*, <http://mapkibera.org/work/methods/>

Marsh, N. (ed.), *Public Access to Government-Held Information* (London: Stevens and Sons, 1987)

Marsh, N., 'Public Access to Government-Held Information: A Select Comparative Survey', in Marsh, N. (ed.) *Public Access to Government-Held Information*, (London: Stevens and Sons, 1987) pp.259-6.

- Marx, K., *Critique of Hegel's "Philosophy Of Right."*, trans by O'Malley, J. (Cambridge: Cambridge University Press, 1977)
- Maskin, E. & Tirole, J., 'The Politician and the Judge: Accountability in Government' *American Economic Review* 94.4 (2004) pp.1034–54.
doi:10.1257/0002828042002606
- May, A. & Rowan, K. (eds.) *Inside Information: British Government and the Media* (London: Constable and Company, 1982).
- McDonald, A., 'Archives and Open Government', in McDonald, A. & Terrill, G. (eds.) *Open Government: Freedom of Information and Privacy* (Basingstoke: Macmillan Press, 1998)
- McKemmish, S., Upward, F. & Reed, B., 'Records Continuum Model', *Encyclopedia of Library and Information Sciences* (3rd ed.) (CRC Press, 2010) pp.4447–8.
- Mendel, T., *Freedom of Information: A Comparative Legal Study* (Paris: UNESCO, 2008) <http://unesdoc.unesco.org/images/0015/001584/158450e.pdf>
- Merifield, M., *Loving it when a plan come together*, (London: The National Archives) <https://.blog.nationalarchives.gov.uk/blog/loving-plan-comes-together/>
- Michael, J., *The Politics of Secrecy: Confidential Government and the Public Right to Know* (London: Penguin, 1982)
- Michael, J., 'Fol in Other Countries', in Wilson, D. (ed.) *The Secrets File: The case for freedom of Information in Britain today* (London: Heinemann Education Books, 1984).

- Miller, D. & Roe, S., *Registers and Collaboration: Making Lists We Can Trust*, (Open Data Institute, May 2018) <https://theodi.org/article/registers-and-collaboration-making-lists-we-can-trust-report/>
- Miller, G. J., 'The Political Evolution of Principal-Agent Models,' *Annual Review of Political Science*, 8 (2005) pp.203–25
doi:10.1146/annurev.polisci.8.082103.104840
- Miller, J., 'Settling Accounts with a Secret Police: The German Law on the Stasi Records,' *Europe-Asia Studies*, 50.2 (1998) pp.305–330
- Milligan, S., 'The Problem of *Publicite* in the Archives of Second Empire France', in Blouin, F. & Rosenberg, W. (eds.) *Archives, Documentation and Institutions of Social Memory: Essays from the Sawyer Seminar* (Ann Arbor: University of Michigan Press, 2010)
- Milward, H. B. & Provan, K. G., 'Governing the Hollow State', *Journal of Public Administration Research and Theory*, 10.2 (2000) pp.359–380.
- Ministry of Justice, United Kingdom, *Lord Chancellor's Code of Practice on the management of records issued under section 46 of the Freedom of Information Act 2000* <http://www.nationalarchives.gov.uk/documents/information-management/foi-section-46-code-of-practice.pdf>
- Mitnick, B. M., 'The Theory of Agency: The Policing "Paradox" and Regulatory Behavior', *Public Choice*, 24 (Winter 1975) pp.27-42
- Mitnick, B. M., *Origin of the Theory of Agency: An Account By One of the Theory's Originators* (Rochester: Social Science Research Network, 2013)
<http://papers.ssrn.com/abstract=1020378>
- Mnjama, N., 'Access to Records and Archives in Kenya', *African Research and Documentation*, 106 (2008) pp.59-75

- Mnjama, 'Anne Thurston and Record-keeping Reform in Commonwealth Africa' in J Lowry and J Wamukoya, (eds.) *Integrity in Government Through Records Management: Essays in Honour of Anne Thurston*, (Farnham: Ashgate, 2014)
- Monroe, A., *Interrogation Machine: Laibach and NSK* (Cambridge; London: MIT Press, 2005)
- Moran, C., *Classified: Secrecy and the State in Modern Britain*, (Cambridge: Cambridge University Press, 2013)
- Moreau, L. & Missier, P. (eds.) *PROV-DM: The PROV Data Model* (World Wide Web Consortium, 30 April 2013) <http://www.w3.org/TR/2013/REC-prov-dm-20130430/>
- Moss, M. 'What is the same and what is different', in Moss, M., Endicott-Popovsky, B. & Depuis, M. J. (eds.) *Is Digital Different?* (London: Facet Publishing, 2015)
- Muigai, A., *A Look Back at Kenya's Open Data Journey* (Open Institute, 2014) <http://www.openinstitute.com/a-look-back-at-kenyas-open-data-journey/>
- Musembi, M., 'Development of Archive Services in East Africa', in *Historical Development of Archival Services in Eastern and Southern Africa: Proceedings of the 9th Biennial General Conference* (Mbabane; Roma: ESCARBICA, 1986)
- Mussolini, B., *Discorsi Del 1927*, 2nd ed. (Milan: Alpes, 1928)
- Mutuku, L. & Mahihu, C., *Open Data in Developing Countries: Understanding the Impacts of Kenya Open Data Applications and Services* (Nairobi: iHub Research, 2014) <http://www.opendataresearch.org/sites/default/files/publications/ODDC%20Report%20iHub.pdf>

Nairne, P., 'Policy Making in Public', in Chapman, R. A. & Hunt, M., *Open Government: A Study of the prospects of open government within the limitations of the British political system* (Abingdon: Routledge, 2011)

The National Archives, *PRO2002* (Kew: The National Archives, 2002)
<http://www.nationalarchives.gov.uk/documents/referencefinal.pdf>

The National Archives (UK), *Generic requirements for sustaining electronic information over time: 1 Defining the characteristics for authentic records*, (Kew: The National Archives, 2003)

The National Archives (UK), *Reviewing Digital Records Management in Government* (Kew: The National Archives, 2014)
<http://www.nationalarchives.gov.uk/information-management/manage-information/policy-process/reviewing-digital-records-management-government/research/>

The National Archives (UK), *DROID: User Guide* (Kew: The National Archives, 2017) <http://www.nationalarchives.gov.uk/documents/information-management/droid-user-guide.pdf>

The National Archives (UK), *Making a Freedom of Information Request*
<http://www.nationalarchives.gov.uk/about/freedom-of-information/making-a-freedom-of-information-request/>

National Archives of Australia, *Australian Governments' Interactive Functions Thesaurus* <https://data.naa.gov.au/def/agift/AGIFT.html>

National Archives of Australia, *Our Organisation* <http://www.naa.gov.au/about-us/organisation/>

National Archives of Australia, *Check-up PLUS*,
<http://www.naa.gov.au/information-management/check-up/>

National Assembly of the Republic of Kenya, *Hansard*, 5 July 2000

Near, J. P. & Miceli, M. P., 'Organizational dissidence: The case of whistleblowing', *Journal of Business Ethics*, 1.4 (1982) pp.1-16

Nesmith, T., 'Archives from the Bottom Up: Social History and Archival Scholarship', *Archivaria* 14 (Summer 1982) pp.5–26

Nethery, A. & Holman, R., 'Secrecy and human rights abuse in Australia's offshore immigration detention centres', *The International Journal of Human Rights*, 20.7 (2016) pp.1018-1038

Nietzsche, F., *Beyond Good and Evil: Prelude to a Philosophy of the Future* (Oxford: Oxford Paperbacks, 2008)

Nobel Media AB, *The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2001* (2000)
http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/2001/

Notes on the Lord Chancellor's Security and Intelligence Instrument, United Kingdom,
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/219905/notes-security-intelligence-instrument.pdf

Nursing and Midwifery Order 2001, part III.8, United Kingdom,
<https://www.legislation.gov.uk/ukxi/2002/253/contents/made>

O'Keefe, C. M., Otorespec, S., Elliot, M., Mackey, E. & O'Hara, K., *The De-Identification Decision-Making Framework: CSIRO Reports EP173122 and EP175702* (CSIRO, 2017)
<https://publications.csiro.au/rpr/download?pid=csiro:EP173122&dsid=DS2>

Offentlig Elektronisk Postjournal (OEP) <https://www.oep.no/?lang=en>

The Office of the Hon Lindsay Tanner MP, *Media Release 41/2010: Declaration of Open Government* (2010)

http://parlinfo.aph.gov.au/parlInfo/download/media/pressrel/AKCX6/upload_binary/akcx60.pdf;fileType=application%2Fpdf#search=%22media/pressrel/AKCX6%22

Open Data Institute, *Smart Cities*, <https://theodi.org/smart-cities>

Open Data Institute, *Business Plan 2012-2017* <http://e642e8368e3bf8d5526e-464b4b70b4554c1a79566214d402739e.r6.cf3.rackcdn.com/odi-business-plan-may-release.pdf>

Open Duka, *About Us*, <http://www.openduka.org/index.php/about>

Open Government Partnership, *About OGP*

<https://www.opengovpartnership.org/about/about-ogp>

Open Government Partnership, *How To Join*

<https://www.opengovpartnership.org/how-join>

Open Government Partnership, *Open Government Declaration*,

<https://www.opengovpartnership.org/open-government-declaration>

Open Government Partnership, *Australian Letter of Intent to Join OGP*,

<https://www.opengovpartnership.org/documents/australia-letter-of-intent-join-ogp>

Open Government Partnership, *OGP Response Policy Case Review and*

Resolution of Christopher Snow's Letter of Concern, Dated September 6, 2016 Against the Government of Australia

https://www.opengovpartnership.org/sites/default/files/Australia_RP-Report_August2017.pdf

Open Government Partnership Australia, *Australia's first Open Government National Action Plan: Commitment 3.3 – Improve the discoverability and accessibility of government data and information*
<https://ogpau.pmc.gov.au/australias-first-open-government-national-action-plan-2016-18/commitments/iii-access-government-2>

Open Institute, *Chief Kariuki presenting Lanet Umoja's data experiences* (29 September 2016) <https://www.youtube.com/watch?v=7EXkkYzUyLU>

Open Knowledge International, *The Open Definition* <https://opendefinition.org>

Ormerod, D., Murphy, P. & Phillips, J. (eds.), *Blackstone's Criminal Practice*, (Oxford: Oxford University Press, 2008)

Oxford English Dictionary, 'Diplomatics' (n.d.).
<http://www.oed.com/view/Entry/53206?redirectedFrom=diplomats#eid6620476>

Parikka, J., *Digital Contagions: A Media Archaeology of Computer Viruses* (Oxford: Peter Lang Publishing, 2007)

Pasolini, P. P., *Salò o Le 120 giornate di Sodoma* (London: British Film Institute, 1975)

Pasquinelli, M., 'To Anticipate and Accelerate: Italian Operaismo and Reading Marx's Notion of the Organic Composition of Capital,' *Rethinking Marxism*, 26.2 (2014) pp.178–192 <http://doi.org/10.1080/08935696.2014.888833>

Peng, Y., 'Chinese Villages and Townships as Industrial Corporations: Ownership, Governance, and Market Discipline,' *American Journal of Sociology* 106.5 (2001) pp.1338–70 doi:10.1086/320820

Pickard, A., *Research Methods in Information* (London: Facet Publishing, 2007)

Property Exchange Australia, *About e-Conveyancing*

<https://www.pexa.com.au/lawyers-conveyancers>

Proudfoot, W. B., *The origin of stencil duplicating* (London: Hutchinson & Co. Ltd., 1972).

Public Concern at Work, *Whistleblowing Case Studies* (2003)

<http://www.pcaw.org.uk/law-policy/legal-updates/whistleblowing-case-summaries>

Quill, L., *Secrets and Democracy: From Arcana Imperii to WikiLeaks* (New York: Palgrave MacMillan, 2014)

Radner, R., 'Hierarchy: The Economics of Managing', *Journal of Economic Literature*, Vol. 30, No. 3 (Sep. 1992) pp.1382-1415.

Ragin, C., *The Comparative Method: Moving beyond Qualitative and Quantitative Strategies*. (Berkeley: University of California Press, 1987).

Raley, R., 'Dataveillance and Countervailance,' in Gitelman, L. (ed.) *Raw Data is an Oxymoron*, (Cambridge: MIT Press, 2013)

Ranciere, J., *On the Shores of Politics* (London; New York: Verso, 2007)

Ranciere, J., *Dissensus: On Politics and Aesthetics* (London: Bloomsbury, 2010)

Rasmussen, E. & Magnus, G., 'Government Support Programmes to Promote Academic Entrepreneurship: A Principal-Agent Perspective,' *European Planning Studies* 20.4 (2012) pp.527-46. doi:10.1080/09654313.2012.665035

Republic of Kenya, *OGP National Action Plan 2012-2013*

<https://www.opengovpartnership.org/documents/kenya-action-plan-2012-2013>

Republic of Kenya, *OGP National Action Plan 2*

http://www.opengovpartnership.org/sites/default/files/Kenya_AP2_2016_0.pdf

Republic of Kenya, *Records Management Procedures Manual for the Public Service*, (Nairobi: Office of the Prime Minister, Ministry of State for Public Service, May 2010)

Reuters, 'Hacker posts client emails from Swiss bank BCGE' (9 January 2015)
<https://uk.reuters.com/article/us-bc-geneve-hacker/hacker-posts-client-emails-from-swiss-bank-bcge-idUKKBN0KI1MK20150109>

Ribes, D., & Jackson, S. J., 'Data, Bite, Man, The Work of Sustaining a Long-Term Study', in Gitelman, L. (ed.) *"Raw Data" is an Oxymoron* (Cambridge: MIT Press, 2013)

Richards, T., *The Imperial Archive: Knowledge and the Fantasy of Empire*, (London; New York: Verso, 1993)

Ridener, J., *From Polders to Postmodernism: A Concise History of Archival Theory* (Duluth: Litwin Books, 2009)

Roberts, C., 'The Theatrical Satanism of Self-Awareness Itself: Religion, Art and Anarchy in Pasolini's Salò,' *Angelaki*, 15.1 (2010) pp.29–43.

Roberts, A., *Blacked Out: Government Secrecy in the Information Age* (New York: Cambridge University Press, 2008)

Robertson, K. G., *Public Secrets: A Study in the Development of Government Secrecy* (London: MacMillan Press, 1982)

Robertson, C., *The Passport in America: The History of a Document* (Oxford: Oxford University Press, 2012)

- Roeder, J., Eppard, P., Underwood, W., & Lauriault, T., 'Part Three—Authenticity, Reliability and Accuracy of Digital Records in the Artistic, Scientific and Governmental Sectors: Domain 2 Task Force Report' in Duranti L. & Preston, R. (eds.) *International Research on Permanent Authentic Records in Electronic Systems (InterPARES 2): Experiential, Interactive and Dynamic Records*, (Padova: Associazione Nazionale Archivistica Italiana, 2008)
- Roper, M., 'Access to Public Records', in Chapman, R. A. & Hunt, M., *Open Government: A Study of the prospects of open government within the limitations of the British political system* (Abingdon: Routledge, 2011)
- Rosenberg, D., 'Data before the Fact', in Gitelman, L. (ed.) "*Raw Data*" is an Oxymoron (Cambridge: MIT Press, 2013) pp.15-40.
- Ross, S. A., 'The Economic Theory of Agency: The Principal's Problem,' *The American Economic Review* 63.2 (1993) pp.134–39
- Rousseau, J-J., *A Treatise on the Social Contract* (London: T. Becket and P.A. De Hondt, 1764)
- Rowat, D. (ed.) *Administrative Secrecy in Developed Countries* (London: MacMillan Press, 1979).
- Runji Njeru, G., *Independent Reporting Mechanism Kenya Progress Report 2012-13* (Open Government Partnership, 2013)
http://www.opengovpartnership.org/sites/default/files/Kenya_final_2012_0.pdf
- Sade, D.A.F., *The 120 Days of Sodom* (London: Arrow Books, 1990).
- Sarangi, S. & Slembrouck, S., *Language, Bureaucracy, and Social Control* (New York: Routledge, 2014)
- Schalkwyk, F. van, Willmers, M. & Schonwetter, T., *Embedding Open Data Practice: Developing Indicators on the Institutionalisation of Open Data*

- Practice in Two African Governments*, (Cape Town: University of Cape Town Intellectual Property Unit, June 2015)
<http://webfoundation.org/docs/2015/08/ODDC-2-Embedding-Open-Data-Practice-FINAL.pdf>
- Sconce, J., 'On the Origins of the Origins of the Influencing Machine' in Huhtamo, E. & Parikka, J. (eds.) *Media Archaeology: Approaches, Applications, and Implications* (Berkeley: University of California Press, 2011)
- Searle, J., 'A Taxonomy of Illocutionary Acts', *Language, Mind and Knowledge*, 7 (1975) pp.344–69
- Serres, M., *The Natural Contract* (Ann Arbor: University of Michigan Press, 1995)
- Sharma, A., 'Professional as Agent: Knowledge Asymmetry in Agency Exchange,' *The Academy of Management Review* 22.3 (1997) pp.758-798.
- Shepherd, E., *Archives and Archivists in 20th Century England*, (Farnham: Ashgate, 2009)
- Siegel, G., *Forensic Media: Reconstructing Accidents in Accelerated Modernity*, (Durham: Duke University Press, 2014)
- Sigmund, J. P., "Form, Function and Archival Value", *Archivaria*, 33.2 (1991) pp.141-147.
- Skemer, D. C., 'Diplomatics and Archives,' *The American Archivist*, 52.3 (1989) pp.376–82.
- Smith, A., *The Wealth of Nations: Books I-III*, (New York: Penguin Classics, 1982).
- Smith, A., *The Theory of Moral Sentiments*, (New York: Penguin Classics, 2010).

- Sorrell, S. et al, *The Economics of Energy Efficiency: Barriers to Cost-Effective Investment* (Cheltenham: Edward Elgar, 2004)
- Spence, M., 'Job Market Signaling,' *The Quarterly Journal of Economics* 87.3 pp.355–74.
- Spencer, B. J., 'Asymmetric Information and Excessive Budgets in Government Bureaucracies: A Principal and Agent Approach' *Journal of Economic Behavior and Organization* 3.2 (1982) pp.197–224. doi:10.1016/0167-2681(82)90018-X
- Stake, R., *The art of case study research* (Thousand Oaks: Sage Publications, 1995).
- Stanley, M., 'Where is that Moon, Anyway? The Problem of Interpreting Historical Solar Eclipse Observations', in Gitelman, L. (ed.) *Raw Data is an Oxymoron*, (Cambridge: MIT Press, 2013) pp.77-88.
- Statistical Data and Metadata Standard (SDMX), *SDMX Standards: Section 1; Framework for SDMX Technical Standards, Version 2.1* https://sdmx.org/wp-content/uploads/SDMX_2-1_SECTION_1_Framework.pdf
- Stewart, D., 'Assessing Access to Information in Australia: The impact of freedom of information laws on the scrutiny and operation of the Commonwealth Government', in Wanna, J., Lindquist, E. A. & Marshall, P. (eds.) *New Accountabilities, New Challenges* (Canberra: ANU Press, 2015)
- Stigler, G J., 'Information in the Labor Market,' *Journal of Political Economy* 70.5 (1962) 94–105 doi:10.2307/1829106
- Stiglitz, J. E., 'The Contributions of the Economics of Information to Twentieth Century Economics,' *The Quarterly Journal of Economics* 115.4 (2000) pp.1441–78.

- Stirner, M., *The Ego And His Own: The Case Of The Individual Against Authority* (London: Verso, 2014)
- Stoler, A. L., *Along the Archival Grain: Epistemic Anxieties and Colonial Common Sense* (Princeton; Oxford: Princeton University Press, 2009)
- Stolowy, H., Baker, R., Jeanjean, T. & Messners, M., 'Information, trust and the limits of "intelligent accountability" in investment decision making: Insights from the Madoff case,' *Les Cahiers de Recherche* 956 (2011)
<https://EconPapers.repec.org/RePEc:ebg:heccah:0956>
- Stryker-Rodda, H., *Understanding Colonial Handwriting*, (Baltimore: Genealogical Publishing Co. Inc., 1986)
- Stubbs, R., 'Freedom of Information and Democracy in Australia and Beyond,' *Australian Journal of Political Science*, 43.4 (2008) pp.667-684
- Subramaniam, V., *Transplanted Indo-British Administration*, (New Delhi: Ashish House Publishing, 1977)
- Supreme Court of the United States, *Federal Rules of Evidence* (2016)
https://www.law.cornell.edu/rules/fre/rule_803
- Sweeney, S., 'The Ambiguous Origins of the Archival Principle of "Provenance"', *Libraries & the Cultural Record*, Vol. 43, No. 2 (2008) pp.193-213.
- Tanner, L., 'Declaration of Open Government', *Australian Government Department of Finance* (2010)
<http://www.finance.gov.au/blog/2010/07/16/declaration-open-government>
- Taylor, H., 'My Very Act and Deed': Some Reflections on the Role of Textual Records in the Conduct of Affairs', *The American Archivist*, 51.4 (1988) pp.456–69.

Terrill, G., ‘A bit fast money, a bit white shoe brigade’? Freedom of Information and Australian History,’ *Twentieth Century British History*, 12.2, (2001) pp.231–242

Think-tank for Action on Social Change (TASC) *Setting the Record Straight: Record Keeping for Good Governance*, (TASC, 2016)
https://www.tasc.ie/download/pdf/policybrief_settingtherecordstraightfinal.pdf

Thompson, M., *A Paper House* (New York: Pantheon, 1992)

Thurston, A., ‘Trustworthy Records and Open Data’, *The Journal of Community Informatics*, 8:2, 2012.

Thurston, A., *Managing Records and Information for Transparent, Accountable, and Inclusive Governance in the Digital Environment: Lessons from Nordic Countries (English)* (Washington: World Bank Group, 2015)
<http://documents.worldbank.org/curated/en/222041468189536195/Managing-records-and-information-for-transparent-accountable-and-inclusive-governance-in-the-digital-environment-lessons-from-Nordic-Countries>

Thüsing, G. & Forst, G., ‘Whistleblowing Around the World: A Comparative Analysis of Whistleblowing in 23 Countries’ in Thüsingm G. & Forst, G. (eds.) *Whistleblowing: A Comparative Study* (Dordrecht: Springer, 2016) pp.3-30.

Tough, A. & Lihoma, P., ‘The development of recordkeeping systems in the British Empire and Commonwealth, 1870s–1960s’, *Archives and Manuscripts*, 40.3 (2012) pp.191-216

Transparency and Accountability Initiative, *Open Gov Guide*, now at
<https://www.opengovpartnership.org/resources/open-government-guide>

Transparency International, *Open Data: Promise, but not enough progress from G20 countries* (2017)

https://www.transparency.org/news/feature/open_data_promise_but_not_enough_progress_from_g20_countries

Travers, R., *Ideology and Empire in Eighteenth-Century India: the British in Bengal* (Cambridge: Cambridge University Press, 2007)

Treadwell, G., 'FOI scholarship reflects a return to secrecy', *Pacific Journalism Review*, 22.1 (2016) pp.121-136.

Turner, R., 'Speech and the social contract' *Inquiry*, 28.1-4 (1985) pp.43–53.
<http://doi.org/10.1080/00201748508602047>

UK Open Government Civil Society Network,
<https://www.opengovernment.org.uk/>

UNECA, African Union, African Development Bank & United Nations Development Programme, *Africa Data Consensus*, (29 March 2015)
http://www.uneca.org/sites/default/files/PageAttachments/final_adc_-_english.pdf

United Nations Economic Commission for Africa (UNECA) *Statisticians converge on African Data Revolution* (28 March 2014)
<http://www.uneca.org/stories/statisticians-converge-african-data-revolution>

United Nations' High-Level Panel of Eminent Persons on the Post-2015 Development Agenda, *A New Global Partnership: Eradicate Poverty and Transform Economies through Sustainable Development* (New York: United Nations Publications, 2013) <http://www.post2015hlp.org/wp-content/uploads/2013/05/UN-Report.pdf>

Vincent, D., *The Culture of Secrecy: Britain, 1832-1998* (Oxford: Oxford University Press, 1998)

Virilio, P., *The Information Bomb*, 2nd ed. (London; New York: Verso, 2005)

- Vismann, C., *Files: Law and Media Technology*, trans. by Winthrop-Young, G. (Stanford: Stanford University Press, 2008)
- Wamukoya, J., 'Records Management and Governance in Africa in the Digital Age' in Lowry, J. & Wamukoya, J. (eds.), *Integrity in Government Through Records Management: Essays in Honour of Anne Thurston* (Ashgate: Farnham, 2014)
- Waruru, M., 'To arm against drought, Kenya maps its water resources', *Thomas Reuters Foundation* <http://news.trust.org/item/20180610090015-agcuf/>
- Weiss, A. & Stiglitz, J., 'Sorting Out the Differences Between Signaling and Screening Models' (Cambridge: National Bureau of Economic Research, Inc., 1990) <http://www.nber.org/papers/t0093.pdf>
- Weld, K., *Paper Cadavers: The Archives of Dictatorship in Guatemala* (Durham: Duke University Press, 2014)
- White, P., 'Official Secrets and Government Openness in Britain', *The Australian Library Journal*, 29:1 (1980) pp.20-28.
- Wigmore, J. H., *Evidence in Trials at Common Law, Volume 7*, ed. and rev. by Chadbourn, J. (Boston: Little, Brown, 1978)
- Wilkinson, N., *Secrecy and the Media: The Official History of the United Kingdom's D-Notice System* (London: Routledge, 2009)
- Wilkinson, D., 'Open Government: The Development of Policy in the United Kingdom in the 1990s' in McDonald, A. & Terrill, G. (eds.) *Open Government* (Basingstoke: Macmillan, 1998).
- Williams, C., 'Diplomatic Attitudes: From Mabillon to Metadata', *Journal of the Society of Archivists*, 26.1 (2005) pp.1-24

Wilson, D. (ed.), *The Secrets File: The Case for Freedom of Information in Britain Today* (London: Heinemann Educational Books, 1984).

Wood, B., 'Principal-Agent Models of Political Control of Bureaucracy,' *The American Political Science Review* 83.3 (1989) pp.965–78.
doi:10.2307/1962069

Wood, S., 'Making Secret(s): The Infrastructure of Classified Information',
doctoral thesis, University of California, Los Angeles, 2017.

World Wide Web Consortium, *Data on the Web Best Practices*
<http://www.w3.org/TR/dwbp/>

Worthy, B., 'Does FOI Create a Chilling Effect? Evidence from the UK', *Open Data Study: Research on Open Data and Transparency*,
<https://opendatastudy.wordpress.com/2017/09/28/does-foi-create-a-chilling-effect-evidence-from-the-uk/>

Worthy, B., *Independent Reporting Mechanism (IRM): United Kingdom End of Term Report 2013-2015*, (Open Government Partnership, 2016)
https://www.opengovpartnership.org/sites/default/files/UK_EOTR_2013-2015_0.pdf

Worthy, B., *The Politics of Freedom of Information: How and Why Governments Pass Laws that Threaten their Power* (Manchester: Manchester University Press, 2017)

World Wide Web Foundation, *Open Data Barometer, Third Edition, Regional Report, Africa* (May 2016) <http://opendatabarometer.org/doc/3rdEdition/ODB-3rdEdition-AfricaReport.pdf>

World Wide Web Foundation, *Open Data Barometer: ODB Methodology - v1.0 28 April 2015* <http://opendatabarometer.org/doc/3rdEdition/ODB-3rdEdition-Methodology.pdf>

Wraith, R. E., 'United Kingdom', in Rowat, D.C. (ed.) *Administrative Secrecy in Developed Countries* (London: MacMillan, 1979)

Wright, P., Mukherji, A. & Kroll, M. J., 'A Reexamination of Agency Theory Assumptions: Extensions and Extrapolations,' *The Journal of Socio-Economics* 30.5 (2001) pp.413–29 doi:10.1016/S1053-5357(01)00102-0

Wrong, M., *It's Our Turn to Eat* (London: HarperCollins, 2009)

Yates, J., *Control through Communication: The Rise of System in American Management* (Baltimore: Johns Hopkins University Press, 1993)

Yeo, G., *Records, Information, and Data: Exploring the Role of Record Keeping in an Information Culture* (London: Facet Publishing, 2018)

Yin, R., *Case study research: Design and methods* (5th ed.) (Thousand Oaks: Sage Publications, 2014)

Young, A. & Verhulst, S., *Kenya's Open Duka: Open Data for Transactional Transparency*, (OMIDYAR Network, January 2016) <http://odimpact.org/case-kenyas-open-duka.html>

Zukin, S. *Beyond Marx and Tito: Theory and Practice in Yugoslav Socialism* (Cambridge: Cambridge University Press, 1975)

Legislation

Access to Information Act 2016, Republic of Kenya

<http://kenyalaw.org/lex/actview.xql?actid=No.%2031%20of%202016>

Constitution of Kenya Act 1969, s79(1), Republic of Kenya

Freedom of Information Act 2000, United Kingdom,

<https://www.legislation.gov.uk/ukpga/2000/36/contents>

Freedom of Information (Fees and Charges) Amendment Regulations (2010)

No.1., Commonwealth of Australia

<https://www.legislation.gov.au/Details/F2010L02781>

Official Secrets Act (No. 11 of 1968) section 3.1(c), Republic of Kenya

Public Interest Disclosure Act 1998, S 43B(1) United Kingdom

<https://www.legislation.gov.uk/ukpga/1998/23/section/1>