Taming Abundance: Doing Digital Archival Research (as Political Scientists)

Diana S. Kim, *Georgetown University, USA*

ABSTRACT Political scientists are increasingly using digitized documents from archives. This article is a practical introduction to doing digital archival research. First, it explains when and why political scientists use evidence from archival research. Second, it argues that the remote accessibility of digitized records provides new opportunities for comparative and transnational research. However, digital archival research also risks aggravating five types of biases that pose challenges for qualitative, quantitative, interpretive, and mixed-methods research: *survival*, *transfer*, *digitization*, and *reinforcement bias* at the level of record collection and *source bias* at the level of record creation. Third, this article offers concrete strategies for anticipating and mitigating these biases by walking readers through the experience of entering, being in, and leaving an archive, while also underscoring the importance of learning the structure of an archive. The article concludes by addressing the ethical implications to archival research as a type of field research for political scientists.

Among political scientists, the practice of using digitized documents from archives has become increasingly common. This article is a practical introduction to doing digital archival research. First, it explains when and why political scientists use evidence based on archival research. Second, it argues that the remote accessibility of digitized records provides new opportunities for comparative and transnational research. However, digital archival research also risks aggravating five types of biases that pose challenges for scholarship relying on qualitative, quantitative, interpretive, and mixed-methods: *survival*, *transfer*, *digitization*, and *reinforcement bias* at the level of record collection and *source bias* at the level of record creation. Third, the article offers concrete strategies for anticipating and mitigating these biases by walking readers through the experience of entering, being in, and leaving an archive, while also underscoring the importance of learning the structure of an archive. The article concludes by addressing the ethical implications to archival research as a type of field research for political scientists.

Diana S. Kim (ORCID) *is assistant professor at Georgetown University in the Edmund A. Walsh School of Foreign Service and a core faculty member of the Asian Studies Program. She can be reached at diana.kim@georgetown.edu.*

WHEN AND WHY POLITICAL SCIENTISTS TURN TO ARCHIVES

For empirically oriented subfields, "doing archival research" often means collecting data from historical records for a quantitative or a mixed-methods approach to causal inference or primary sources for qualitative case studies aimed at theory testing (American Political Science Association 2019, 2021). Political scientists also do archival research to gather original evidence for descriptive inference, case studies for theory building, and interpretive analyses of causal processes and concept histories. Despite methodological differences, political scientists tend to collect archival records as a means to an end. We seek records with words and numbers that may support, confirm, refine, or rule out answers to questions formulated *before* entering the archive. Unlike historians who explore questions *in* the archives, political scientists pursue theory-driven evidence.[1]

Specifically, archives beckon political scientists for several reasons.[2] First, the study of formal institutions occupies a privileged place in the discipline. The reasons why bureaucracies, legislatures, courts, the military, and the police behave the way they do—making decisions and enforcing them, extracting and distributing resources, and exercising coercive and symbolic power—are difficult to ascertain directly. Declassified records from the archives of relevant agencies can yield information about institutional behavior. Internal correspondence, draft policy memos, and minutes of meetings and debates are especially

helpful for theorizing about why leaders make choices; how coalitions and rivalries in politics and society emerge and evolve; and the role of ideology, ideas, beliefs, and other factors relating to human agency on outcomes of interest (Blaydes 2018; Lawrence 2013; Mendoza 2021; Saunders 2011; Subotic 2019).

Second, and relatedly, the archives of formal institutions can provide indirect information about social actors who do not generate their own written records or lack resources to conserve them. For instance, reports from Truth and Reconciliation Commissions contain interview transcripts with survivors of major atrocities; police records include statements made by citizens arrested or under surveillance; and legal and court records include people's testimonies and petitions (Hussin 2016; Leiby 2009; Luft 2020; Nako 2019). Informal archives—namely, "unmapped, non-systematized collections of materials kept by individuals and groups in the spaces under study"—may yield more proximate records into the lived past of social actors (Auerbach 2018, 345; Balcells and Sullivan 2018; Davenport 2010).

Third, political scientists do archival research to find original evidence for case studies using process tracing to identify causal mechanisms and for theory-testing and theory-building purposes. Most process tracing requires precise sequencing of independent, dependent, and intervening variables, as well as careful descriptions of each step in a chronological trajectory or causal narrative (Collier 2011; Faletti 2006; Ricks and Liu 2018). Archival research helps scholars address two key challenges that arise in this line of research: confirmation bias and imperfect counterfactuals. One may inadvertently "cherry-pick" evidence that supports a hypothesis without observing evidence that contradicts it or lends credibility to a rival hypothesis. There also are risks of teleological explanations because it is difficult to reconstruct plausible alternative trajectories through which a given outcome is obtained (i.e., the "paths not taken"). Both problems are more likely when case studies rely heavily on secondary sources. Given authors' own biases and the types of historiographical or methodological debates in which they are engaged, there is a risk of deliberately or unconsciously focusing on specific factors or events, giving them greater visibility that can be mistaken for greater causal importance (Lustick 1996; Møller and Skaaning 2018). When indebted solely to secondary sources, there also is a risk of mistaking an historical narrative for a social process, conflating "what happened" and "that which is said to have happened" (Trouillot 1995, 2). Political scientists seek to reduce these problems through archival research of primary sources.

Fourth, a turn to archives coincides with a distinctive turn to history in political science (American Political Science Association 2019; Mahoney and Thelen 2015). For those studying the long-run consequences of institutions or events, archival records can yield granular time-series data amenable to rigorous quantitative causal analyses that also are descriptively valuable for identifying empirically puzzling or theoretically surprising patterns (Guardado 2018; Suryanarayan and White 2021).

Fifth, interpretive analyses of historical events, processes, and concept histories may rely on archives for texts that capture, mediate, and represent the ideas, linguistic communities, and worldviews of actors in context (Grant 2015; Kim 2020; Mackinnon 2019).

PROMISES AND PITFALLS OF DIGITAL ARCHIVAL RESEARCH

The large-scale digitization of archival records has affected research practices in both positive and negative ways (Trivellato 2019, 8–10; Turnbull 2014). On the one hand, the online availability of documents digitized by archives has reduced economies of scale for identifying, accessing, and collecting records that once required significant time investment, resources, and country-specific and regional expertise (Putnam 2016, 389). The full-text searchability of archival records and improved quality of optical character recognition (OCR) enables multisite, multilanguage archival research. Skimming through large quantities of records has become efficient. Neural machine translation interfaces (e.g., Google Translate) make it possible to collect sources in foreign languages without knowing the language itself. To paraphrase the historian Lara Putnam (2016, 383), political scientists also are "able to find without knowing where to look."

On the other hand, the availability of digitized records that are too easily accessed remotely may generate problems of excess abundance. It becomes more difficult for researchers to ascertain how the documents they consult fit within an archive's general structure. Without knowing how representative a subset of documents is of which universe of records, it is difficult to make broader inferences about the empirical reality they capture.

Excess abundance aggravates four types of biases at the level of data collection.[3] *Survival bias* occurs when records are missing and destroyed in a nonrandom way. *Transfer bias* occurs when the records that an archive acquires (i.e., "accession") and catalogues reflect asymmetries of power, wealth, and privilege that favor certain agencies and individuals or the archive's own institutional interests. *Digitization bias* can amplify transfer bias when archives are selective about which records are digitized and made accessible remotely. *Reinforcement bias* occurs when researchers focus on collecting a subset of records that confirm their hypotheses without consulting other record groups. At the level of record creation, digital archival research also faces greater risks of *source bias*, which reflects the extent to which governments and the powerful tend to be those who write records in the first place.[4]

To be sure, these challenges have always plagued onsite archival research. What has changed with more digitization and remote access is a disruption in the ways that researchers are able to discern and address biases through the practical experience of being in an archive and the physical tasks of requesting and accessing documents. Political scientists have always relied on humanistic solutions to problems of abundance in archival research.[5] The repetitive act of using call/shelf numbers to order documents in a reading room is also an act of tacit learning about the record hierarchy in which a document is embedded. Locating the "right" document that serves as evidence for (or counterevidence to) one's hypothesis requires browsing and skimming through a large quantity of seemingly irrelevant

This article is a practical introduction to doing digital archival research.

records. This serves as a quasi-forced check against reinforcement bias and generates serendipitous encounters with evidence that one does not necessarily know to look for. Finite physical capacity also forces political scientists to make deliberate choices about what to consult and collect. When doing archival research in-person, there are only so many documents that one can request and copy in a day. This limit is less so at an "infinite archive" with digitized records.

Born-digital records are items that are created originally in digital form, such as emails, social media posts, and other types of electronic records. *Digitized records* generally refer to scanned copies of an original analog record; they are a type of *access derivative.* Just as paper-based records are fragile and can experience wear over time, digitized and born-digital records also face risks of data degradation (i.e., "bit rot"), losses in the process of transcoding and compression, and obsolescent formats.

Political scientists have always relied on humanistic solutions to problems of abundance in archival research.

How may political scientists doing archival research anticipate and address challenges of excess abundance from digitized records? The following subsections offer several concrete strategies that commonly center on ways to tacitly learn the structure of an archive.

Before Entering an Archive

Imagine that you are planning to travel to a new country.[6] Learn the basic language that people at the archives use. *Provenance* (also known as *respect des fonds*) refers to the original creator of a record and its history ownership. It is a type of principle for arranging records in a way that preserves their integrity based on how they originated. It also informs the practice of *original order*, by which archives maintain records according to how creators arranged them originally.[7] This is why records are not necessarily found chronologically, alphabetically, or according to geography. The organizing categories are those created by the individual, family, or agency from whom the archive acquired the documents.

Many archives arrange records according to a hierarchy. *Collections* are a general grouping of records that do not necessarily share the same provenance. Within a collection, the highest level of description is a *fonds* (or "record group") in which records share provenance. A fonds is subdivided into *series* (and *subseries*), which are further subdivided into *files.* The lowest level of the hierarchy is an *item*, which is a record that is indivisible. The item-level record is what we usually understand as an archival document—the piece(s) of paper for a surveillance file of an individual, a court-proceedings transcript, or a tax record (figure 1).[8]

Finding aids are one of the most important tools for navigating an archive (figure 2). They are detailed inventories of the records in a collection, containing metadata of a collection's provenance, summary of contents and organization, administrative history and biographical notes, and size (e.g., number of boxes and linear feet of records). Some finding aids may include a file-by-file, item-by-item list of the collection's contents. An *index* (or *catalogue*) is a list of records with shelf/call numbers. *Research guides* provide descriptive explanations of how to explore an archive's holdings and often are written by an archivist or subject specialist. Think of each tool as a different genre of storytelling about an archive. Finding aids and indexes are often cryptic and not necessarily meant to be read from cover to cover. Rather, consult them selectively. Often hidden within their flat prose is invaluable contextual information about a collection. Research guides are more reader friendly, with rich narratives that should be consulted discerningly, not least because they are the products of another's interpretation of an infinite archive.

Entering an Archive

Now you are at the archive, in the (virtual) reading room (figures 3 and 4).[9] How do you begin to find documents? For political scientists in search of theory-driven evidence, a first helpful step is to develop a list of search keywords relating to their research question and tentative argument: X causes Y; A influences B. What are your concepts, words, and proper nouns relating to X and Y, A and B?

A surprising amount of archival research in the twenty-first century, whether in person or remote access, is time spent doing reiterative keyword searches. The search box, whether on an archive's closed intranet terminal or public website, always mediates access.

However, digital archival research is not a Google search. To effectively use the search box, consider creating a two-layered set of keywords: (1) words that reflect how the archive labels and categorizes records relating to your X/A and Y/B; and (2) words in

Figure 1
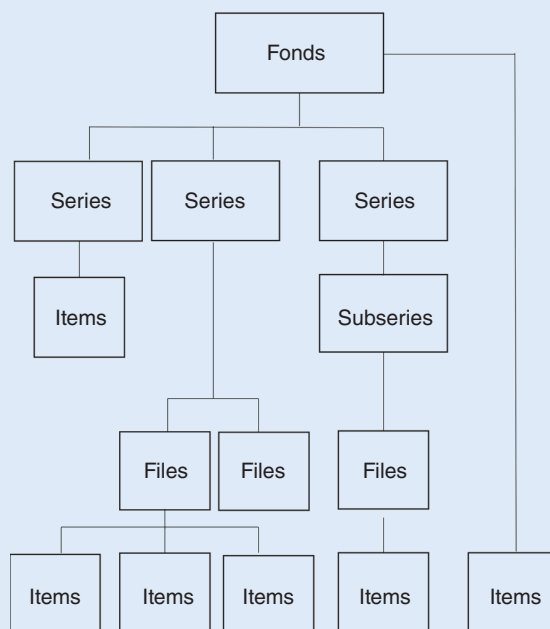### Diagram of Levels of Archival Arrangement

Figure 2
Example of a Virtual Finding Aid

The Library of Congress > Researchers > Search Finding Aids > Patsy T. Mink papers, 1883-2005

Patsy T. Mink papers, 1883-2005

Search this Finding Aid [_____] [all words ▾] [Search]   Search All Finding Aids | Help | Contact Us

Overview | Contents List | Index Terms | Using this Collection | Search Results | Print/Download

Title Page | Collection Summary | Biographical/Organizational Note | Scope and Content | Arrangement

⚠ Access restrictions apply.   ⚠ Some or all content stored offsite.

Collection Summary

Title	Patsy T. Mink papers, 1883-2005
Span Dates	1883-2005
Bulk Dates	(bulk 1953-2002)
ID No.	MSS84957
Creator	Mink, Patsy T., 1927-2002
Extent	880,600 items
Extent	2,638 containers plus 71 oversize and 1 classified
Extent	1,530 linear feet
Extent	1 microfilm reel
Language	Collection material in English
Location	Manuscript Division, Library of Congress, Washington, D.C.
Finding Aid Permalink	Cite or bookmark this finding aid as: https://hdl.loc.gov/loc.mss/eadmss.ms010008
LCCN Permalink	LC Catalog record: https://lccn.loc.gov/mm2003084957
Summary	Lawyer, public official, and United States representative from Hawaii. Correspondence, memoranda, legislative files, speeches and writings, testimony, statements, press releases, appointment books and scheduling files, travel itineraries, campaign files and political ephemera, notes, casework, law practice client files, court documents, maps, card files, biographical material, interviews, questionnaires, awards and honors, photographs, student papers, family papers, scrapbooks, news clippings, and printed matter documenting Mink's public service career.

Source: Patsy T. Mink Papers, 1883–2005, US Library of Congress.
Notes: The "Using This Collection" tab includes information on provenance. The "Scope and Content Note" tab summarizes the content of the 14 series that comprise this collection: nine series on Mink's professional and political career and four series including family papers and classified records. The "Overview/Collection Summary" tab provides information on the collection's size; the "Index Terms" tab provides search keywords (i.e., names, places, occupations, organizations, and subjects) used to index the collection's description. As a PDF document, this finding aid is 532 pages, available at https://findingaids.loc.gov/exist_collections/ead3pdf/mss/2010/ms010008.pdf.

context—that is, what past actors and institutions would have called your X/A and Y/B. To figure out the former, consult a finding aid, research guide, or index. To ascertain the latter, consult a seminal history or empirical study relating to your research.

For instance, suppose your X/A relates to street-level bureaucrats. The concept itself is an academic term of art. Which alternative words might capture the presence of those actors in the archive? Commonsense may suggest "local administrators" and "municipal officials." Now think back to Lipsky's (1980, 17–18) canonical study of street-level bureaucrats in the United States, which refers to them as "public service workers," "public employees," and "low-level workers." Add these three keywords to your list. Perhaps urban Pakistan in the 2000s is your context; Hull's (2013, 57–59) *Government of Paper* guides you toward this word: "clerks." Perhaps the historical context of the early-twentieth-century British empire is relevant: "district officers" abound in Lugard's (1922) *The Dual Mandate in British Tropical Africa* and Mamdani's (1996) *Citizen and Subject.* Now turn to your Y/B. It includes the word "opium." A digitized finding aid for the India Office Records at the British Library indicates that "opium" is cross-referenced with terms such as "Abkaree" and "Separate Revenue." You already are starting to identify the controlled vocabulary of the archive, which has grouped together records according to their provenance of the British Indian

colonial government for Burma's Excise Department (Kim 2020).[10]

Archival research is a reiterative process of discovery. By refining a list of search keywords rooted in both historical and archival contexts, political scientists learn tacitly about how "their" documents fit within the broader structure of an archive's records, and they are better able to identify survival and transfer bias. Reiteration is an investigative process, not dull repetition, especially in a digital archive. Allow yourself to be distracted, especially by unexpected words, unfamiliar concepts, and odd proper nouns that pop up while scrolling and browsing. These are the moments when chance encounters that lead to new discoveries may occur.

Being in an Archive

There is something both exhilarating and disorienting about being in an archive, physically or virtually. You have just found a 140-page archival document that seems promising, and there are so many more. What do you do? Taking notes systematically and organizing them according to original order are two seemingly mundane yet powerful strategies for addressing challenges of excess abundance.

First, design a consistent template for taking notes on each item that you consult.[11] There is no right or wrong approach to notetaking. The template only needs to be one that can be

Figure 3

In-Person Reading Room at France's National Archives for Overseas Territories (Aix-en-Provence)

Source: Photograph by the author.

Virtual Reading Room of the US State Department Archives Online

Note: See https://foia.state.gov/Search/Search.aspx.

repeated over and over again. A minimalist may include only the file's call number, title, and a brief summary of its contents. A maximalist may add the date accessed, date of original creation, copyright restrictions, a more detailed item-by-item description of its contents, and transcribed notes. Consistent and systematic notetaking at the item level is tedious and, at first, time-consuming. However, it becomes a habit that saves time in the long run. Crucially, it establishes a cumulative record of not only what types of records and information you selected to include in the final analysis but also what was *not* incorporated, which helps to mitigate reinforcement bias.

Second, systematic notetaking goes hand in hand with systematically organizing those notes. Consider mimicking the original order of the archive when storing notes as well as digitized copies of original documents. For instance, your 140-page document is from the UK National Archives in Kew. You located a digitized copy online and the original reference number is CO/885/1/20. First, create four nested files. Then, under "UK National Archives, Kew (digital)," create a second file labeled CO/885, followed by a third file labeled CO/885/1, and finally "20." You have just replicated the record hierarchy for this specific document: Item No. 20 in the Miscellaneous Files (1) in the War and Colonial Department and Colonial Office Series (885), within the Colonial Office Records (CO) (figure 5).

Leaving an Archive

A realization hits you as you leave an archive, whether walking out of the building or closing your web browser for the day. The documents that you consulted and the records that you accessed bear traces of the lives of others. Inevitably, the archival research you have just done is an encounter with people in history.

Digital archival research is inextricably tied to ethical considerations for political scientists. Survival, transfer, digitization, and reinforcement biases are products of how archives are not impartial repositories but rather institutions shaped by power, politics, and privilege. Remote accessibility makes it easier to forget how selective and partial an archive's holdings are, not least because it eliminates the many inconveniences that remind researchers of its arbitrariness and incompleteness. Thus, there is a greater risk of overrepresenting the coherence of past events, processes, and human experiences because it is easier to presume the digitized
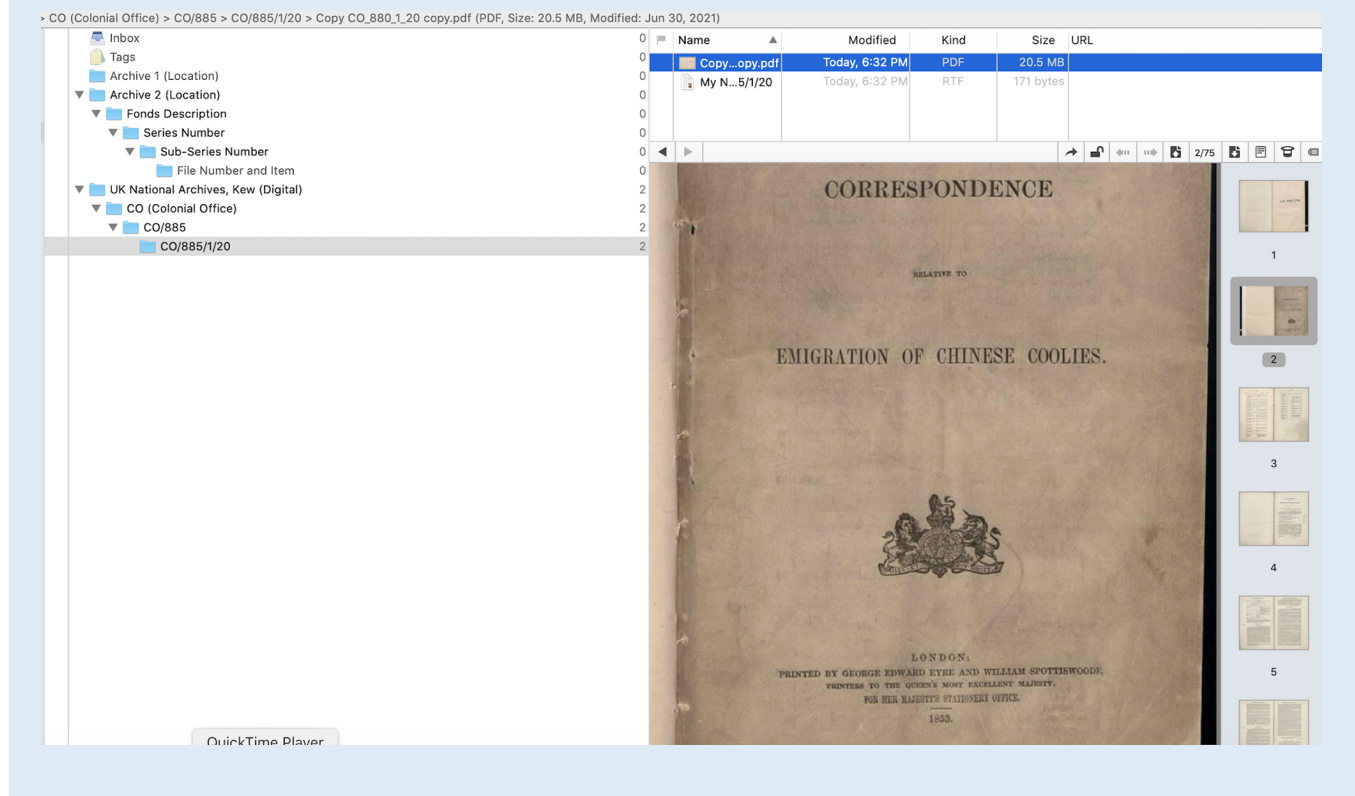
Digital archival research is inextricably tied to ethical considerations for political scientists.

available documents capture a greater share of historical reality than actually warranted. The technology of digitization also generates new considerations that bring the ethics of historical representation squarely into the ambit of political science's theory-driven evidence seeking. OCR errors can result in the erasure of an individual's trace in the archives and, conversely, digitizing microlevel data may inadvertently "find" people with vexed histories unknown to their descendants. The boundaries of

Figure 5

Example of How to Mimic an Archive's Original Order When Storing Notes (Using the Data Storage App Devonthink)

copyrights also are blurry for digitized records extending to photographs.

A keen sensitivity to these issues animates emerging scholarship on archival and historical research for political science (American Political Science Association 2019, 2021; Balcells and Sullivan 2018). To bring ethical considerations to the forefront, a basic and necessary question for researchers to ponder is: *Why* do we choose to do what we do with archival records? Although there is no right answer, there are many different types of thoughtful responses. For some scholars, there is an impulse of social justice, of advocacy on behalf of the weaker, the voiceless—not least to rescue them from "the enormous condescension of posterity" (Thompson 1963, 12). Other scholars may seek *not* to speak for other people. "The intention here isn't anything as miraculous as recovering the lives of the enslaved or redeeming the dead," Hartman (2008, 11) wrote of her approach to the archives of the eighteenth-century transatlantic slave trade. Rather, she explained, it is about "laboring to paint as full a picture of the lives of the captives as possible" (Hartman 2008, 11). Different epistemologies may counsel archival research as a way to "better understand how local history and context can be leveraged to inform the design of better policy" or, alternatively, to gain a richer menu of counterfactual empirical realities for rigorous social-scientific inquiry (Fouka 2020; Nunn 2020, 1). In doing archival research, choices are already being made, with stakes that are amplified in the process of remotely accessing digitized records.

Digital archival research gains ethical import when political scientists are able to recognize and explicitly articulate these choices.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIALS

To view supplementary material for this article, please visit http://doi.org/10.1017/S104909652100192X.

CONFLICTS OF INTEREST

The authors declare no ethical issues or conflicts of interest in this research. ∎

N O T E S

1. On epistemological and methodological differences between social scientists and historians when approaching archival evidence, see Gaddis (2002), Lemercier and Zalc (2019), and Sewell (2005). On archival methods in field research for political scientists, see Kapiszewski, MacLean, and Read (2015). On how COVID-19–related interruptions are reshaping field research, see MacLean et al. (2020).

2. Here, I refer mainly to official archives in the form of material records, both textual and nontextual, that either are preserved at the producing institution or collected by an external repository. This is only one among a vast variety of entities that scholars recognize as archives, ranging from informally kept records to nontangible embodiments of collective memory, past ideas, and lived experiences.

3. On survival and transfer bias, see Lee (2017, 5–6) and, more generally, on how biased sampling from archival research can affect hypothesis testing.

4. On selection bias as a type of source bias, see Lustick (1996). On problems of scale that influence biases, see Kumekawa (Forthcoming).

5. For exemplars of such humanistic approaches among historians, see Farge (2015) on the role of tacit learning and tactile experience in archival research; Trivellato (2019) on recovering a "lost canon" through deep contextualized readings of sources identified through digital libraries and data-mining tools; and Rothschild (2021) on reconstructing complex transnational family and social ties using large-scale network visualization, combined with micro-histories of individuals. For an illuminating example of these visualizations, see www.infinitehistory.org/en/networks.html.

6. For valuable "how-to" guides detailing step-by-step processes from identifying archival sites to gaining access and technologies for digitizing records, to writing and analysis, see Redman (2013) and Abbott (2014).

7. For more on provenance and original order, see Schellenberg (1951). On the modern history of provenance as an archival principle, see Sweeney (2008).

8. For a lucid guide to archive terminology, see the online glossary of the University of Cambridge, Kings College Archive Centre, at www.kings.cam.ac.uk/archive-centre/introduction-to-archives/glossary#item.

9. One of the first virtual reading rooms for a US academic repository was established for Richard Rorty's papers at the University of California, Irvine (Light 2014).

10. On a controlled vocabulary, see Abbott (2014, 39–50).

11. See online appendix 1 for a sample template.

REFERENCES

Abbott, Andrew. 2014. *Digital Paper: A Manual for Research and Writing with Library and Internet Materials*. Chicago: University of Chicago Press.

American Political Science Association. 2019. "Comparative Politics and History." *Comparative Politics Newsletter* (7) 1: 1–128. Washington, DC: American Political Science Association.

American Political Science Association. 2021. "Roundtable on 'The Theory and Practice of Archival Research in International History and Politics.'" *International History and Politics Newsletter* 29 (2): 2–13. Washington, DC: American Political Science Association.

Auerbach, Adam. 2018. "Informal Archives: Historical Narratives and the Preservation of Paper in India's Urban Slums." *Studies in Comparative International Development* 53:343–64.

Balcells, Laia, and Christopher Sullivan. 2018. "New Findings from Conflict Archives: An Introduction and Methodological Framework." *Journal of Peace Research* 55 (2): 137–46.

Blaydes, Lisa. 2018. *State of Repression: Iraq under Saddam Hussein*. Princeton, NJ: Princeton University Press.

Collier, David. 2011. "Understanding Process Tracing." *PS: Political Science & Politics* 44 (4): 823–30.

Davenport, Christian. 2010. "Data from the Dark Side: Notes on Archiving Political Conflict and Violence." *PS: Political Science & Politics* 43 (1): 37–41.

Faletti, Tulia. 2006. "Theory-Guided Process Tracing in Comparative Politics: Something Old, Something New." *Newsletter of the Organized Section in Comparative Politics of the American Political Science Association* 17 (1): 9–14.

Farge, Arlette. 2015. *The Allure of the Archives*. Trans. Thomas Scott-Railton. New Haven, CT: Yale University Press.

Fouka, Vasiliki. 2020. "The 1920s and H1B Visas." *Broadstreet*, September 2. https://broadstreet.blog/2020/09/02/the-1920s-and-h1b-visas.

Gaddis, John L. 2002. *The Landscape of History: How Historians Map the Past*. New York: Oxford University Press.

Grant, Daragh. 2015. "The Treaty of Hartford (1638): Reconsidering Jurisdiction in Southern New England." *William and Mary Quarterly* 72 (3): 461–98.

Guardado, Jenny. 2018. "Office-Selling, Corruption, and Long-Term Development in Peru." *American Political Science Review* 112 (4): 971–95.

Hartman, Saidiya. 2008. "Venus in Two Acts." *Small Axe* 12 (2): 1–14.

Hull, Matthew. 2013. *Government of Paper: The Materiality of Bureaucracy in Urban Pakistan*. Oakland: University of California Press.

Hussin, Iza. 2016. *The Politics of Islamic Law: Local Elites, Colonial Authority, and the Making of the Muslim State*. Chicago: University of Chicago Press.

Kapiszewski, Diana, Lauren M. MacLean, and Benjamin L. Read. 2015. *Field Research in Political Science: Practices and Principles*. New York: Cambridge University Press.

Kim, Diana. 2020. *Empires of Vice: The Rise of Opium Prohibition across Southeast Asia*. Princeton, NJ: Princeton University Press.

Kumekawa, Ian. Forthcoming. "Historical Network Analysis." In *SAGE Handbook of Social Network Analysis*, second edition, ed. John McLevey, Peter J. Carrington, and John Scot. London: SAGE Publishing.

Lawrence, Adria. 2013. *Imperial Rule and the Politics of Nationalism: Anti-Colonial Protest in the French Empire*. New York: Cambridge University Press.

Lee, Alexander. 2017. "The Library of Babel Problem: Hypothesis Testing with Archival Sources." Unpublished manuscript, last modified November 24. www.rochester.edu/college/faculty/alexander_lee/wp-content/uploads/2017/11/archives3.pdf.

Leiby, Michele. 2009. "Digging in the Archives: The Promises and Perils of Primary Documents." *Politics and Society* 37 (1): 75–100.

Lemercier, Claire, and Claire Zalc. 2019. *Quantitative Methods in the Humanities: An Introduction*. Trans. Arthur Goldhammer. Charlottesville: University of Virginia Press.

Light, Michelle. 2014. "Managing Risk with a Virtual Reading Room: Two Born-Digital Projects." In *Reference and Access: Innovative Practices for Archives and Special Collections*, ed. Kate Theimer, 17–35. Lanham, MD: Rowman & Littlefield Publishers.

Lipsky, Michael. 1980. *Street-Level Bureaucracy: Dilemmas of the Individual in Public Services*. New York: Russell Sage Foundation.

Luft, Aliza. 2020. "How Do You Repair a Broken World? Conflict(ing) Archives after the Holocaust." *Qualitative Sociology* 43: 317–43.

Lugard, Frederik. 1922. *The Dual Mandate in British Tropical Africa*. London: Blackwood and Sons.

Lustick, Ian. 1996. "History, Historiography, and Political Science: Multiple Historical Records and the Problem of Selection Bias." *American Political Science Review* 90 (3): 605–18.

Mackinnon, Emma S. 2019. "Declaration as Disavowal: The Politics of Race and Empire in the Universal Declaration of Human Rights." *Political Theory* 47 (1): 57–81.

MacLean, Lauren M., Robin Turner, Nabila Rahman, and Jack Corbett. 2020. "Disrupted Fieldwork: Navigating Innovation, Redesign, and Ethics during an Ongoing Pandemic." *Qualitative and Multi-Method Research* 18 (2): 1–8.

Mahoney, James, and Kathleen Thelen. 2015. *Advances in Comparative-Historical Analysis*. New York: Cambridge University Press.

Mamdani, Mahmood. 1996. *Citizen and Subject: Contemporary Africa and the Legacy of Late Colonialism*. Princeton, NJ: Princeton University Press.

Mendoza, Mary Anne. 2021. "When Institutions Reinforce Regional Divides: Comparing Christian and Muslim Colonial Education Policies in the Philippines." *Asian Politics & Policy* 13 (1): 90–104.

Møller, Jørgen, and Svend-Erik Skaaning. 2018. "The Ulysses Principle: A Criterial Framework for Reducing Bias When Enlisting the Work of Historians." *Sociological Methods & Research* 50 (1): 103–34.

Nako, Nontsasa. 2019. "The Live Witness in the Archive: Analysing Live-Witness Testimony in the South African Truth and Reconciliation Commission's Archival Project." *Australian Feminist Studies* 34 (100): 216–31.

Nunn, Nathan. 2020. "The Historical Roots of Economic Development." *Science* 367 (6485). DOI:10.1126/science.aaz9986.

Putnam, Lara. 2016. "The Transnational and the Text-Searchable: Digitized Sources and the Shadows They Cast." *American Historical Review* 121 (2): 377–402.

Redman, Samuel J. 2013. *Historical Research in Archives: A Practical Guide*. Washington, DC: American Historical Association.

Ricks, Jacob I., and Amy H. Liu. 2018. Process-Tracing Research Designs: A Practical Guide. *PS: Political Science & Politics* 51 (4): 842–46.

Rothschild, Emma. 2021. *An Infinite History: The Story of a Family in France over Three Centuries*. Princeton, NJ: Princeton University Press.

Saunders, Elizabeth N. 2011. *Leaders at War: How Presidents Shape Military Interventions*. Ithaca, NY: Cornell University Press.

Schellenberg, Theodore R. 1951. "Principles of Arrangement." *Staff Information Paper Number 18*. Washington, DC: National Archives and Records Administration.

Sewell, William H., Jr. 2005. *Logics of History: Social Theory and Social Transformation*. Chicago: University of Chicago Press.

Subotic, Jelena. 2019. *Yellow Star, Red Star: Holocaust Remembrance after Communism*. Ithaca, NY: Cornell University Press.

Suryanarayan, Pavithra, and Steven White. 2021. "Slavery, Reconstruction, and Bureaucratic Capacity in the American South." *American Political Science Review* 115 (2): 568–84.

Sweeney, Shelley. 2008. "The Ambiguous Origins of the Archival Principle of 'Provenance.'" *Libraries & the Cultural Record* 43 (2): 193–213.

Thompson, Edward Palmer. 1963. *The Making of the English Working Class*. New York: Pantheon Books.

Trivellato, Francesca. 2019. *The Promise and Peril of Credit: What a Forgotten Legend about Jews and Finance Tell Us about the Making of European Commercial Society*. Princeton, NJ: Princeton University Press.

Trouillot, Michel-Rolph. 1995. *Silencing the Past: Power and the Production of History*. Boston: Beacon Press.

Turnbull, Paul. 2014. "Margins, Mainstreams and the Mission of Digital Humanities." In *Advancing Digital Humanities: Research, Methods, Theories*, ed. Katherine Bode and Paul L. Arthur, 258–73. Basingstoke, Hampshire, UK: Palgrave Macmillan.