Life on File: Archival Epistemology and Theory

A. K. M. Skarpelis [1]

Abstract

History's epistemological dilemma equally applies to sociology: how can we make claims about persons, events, and processes on the basis of archival records? This article develops a framework called Life on File that combines sociological strengths in qualitative methodologies and an interest in how states shape populations, with library and information sciences' attention to documentary production and anthropological and historical insights into the intersections of archives, knowledge, and power. The framework has three components: the act of recovery of life, processes, and events; the turning of life into a record; and the movement of a file from collection to preservation and use. The result is a methodologically rigorous and globally mobile theory. Empirically, the piece draws on comparative historical fieldwork on state-led racial classification and naturalization practices preserved in Japanese and German archives. While the framework is grounded in research on state archives, its utility extends across types of archives and records. In addition, it provides sociologists with a roadmap on how to use archives in single case as well as comparative and transnational research.

Keywords Sociology of knowledge · Epistemology · Inference · Theory · Archives · Racial classification

> *The archive is like a raw material, which is not the same as saying that it is an originary material or an unworked-upon material; rather it is what has been made available, what has been thus presented to us, a kind of gift, which is to say also – for future constituencies, future publics – a kind of debt. (*Osborne 1999, 57)

> *Cast in the image of the humans who create them, archives are mortals with aspirations to immortality. (*Daston 2017, 329)

✉ A. K. M. Skarpelis
   askarpelis@fas.harvard.edu

1   Weatherhead Center for International Affairs, Harvard University, 1737 Cambridge Street, Cambridge, MA 02138, USA

Introduction: Sociology's Archival Dilemma

Each sociological subfield struggles with the methodological challenge of how to make claims about the social world. What the historian Reinhart Koselleck called history's fundamental epistemological dilemma—making statements of truth value based on archival records, all the while recognizing that any claim made by the historian will be relative—is just as pertinent for sociological inquiry (Koselleck 2004). Tilly parsed this concern into three components: 1) How does the phenomenon of interest leave traces? 2) How can a researcher elicit or observe these traces? And 3) how can we reconstruct whatever is of interest of the phenomenon and understand a cause/effect relationship (Tilly 2002)? For sociologists working with and through archives, this means asking about how we can make claims about the social world based on archival records, often textual documents stored in institutional repositories. If history is produced from what the archive offers (Fuentes 2016), sociologists drawing on archives need to better understand the epistemological implications surrounding the collection, analysis, and interpretation of archival documents—the process of "production of knowledge about social life" (Reed 2010, 20). Charles Tilly in turn asserted that "all social research rests (…) [on] two theories: a theory explaining the phenomenon under study, another theory explaining the generation of evidence concerning the phenomenon" (Tilly 2002, 248).

Here, I argue analogously that epistemological considerations—issues pertaining to "the creation and dissemination of knowledge in particular areas of inquiry" (Steup 2017)—around archives fall as much under the purview of sociological theory as they do under research methods. Consequently, this article seeks to render manifest qualitative sociologists' epistemological practices by following life *on* file as well as the life *of* the file, all the way from the creation of documentary reality to a final state of archivization and analysis.[1] I ask how life—whether the life of individual persons, interactions, organizational activities, or bureaucratic processes—is captured on file and in archives, and how both in turn shape the production of sociological knowledge.

The materiality of documents often takes the shape of concrete, flat, ink-on-paper text-bearing objects, but information scientists have long argued that anything can be a document, as long as it transports information (Briet 1951; Buckland 1997). Other research distinguishes documents from records by tying the former to format, and the latter to their proximity to action (Yeo 2011). This article purposely collapses the distinction between records and documents and uses the terms file, document, and record almost interchangeably, and defines a file functionally as *intentionally generated and organized physical evidence that was purposefully processed in order to record or transmit information*.

Empirically, the article draws on examples from historical comparative archival fieldwork into racialization processes in the German and Japanese multi-ethnic annexationist empires, collected over some years in a dozen archives in Asia, Europe, and the United States. While the records are specific—historical rather than contemporary, state rather than private or organizational, authoritarian rather than democratic, person files rather than organizational reports, paper files rather than videos—the framework of life on file applies equally to sociological work based on archival records in other time periods and geographic locations. In its focus on questions of knowledge, power, the state, and what we can know about the social world, the project aspires to speak to a number of sociological subfields: to

---

[1] Of course, archives are not solely the prerogative of qualitative sociologists; quantitative and computational sociologists increasingly draw on archives and also benefit from incorporating a less restrictive understanding of our data sources into our analysis, rather than relegating it to a section on the work's unavoidable limitations.

organizational sociology by treating the creation of documentary reality as an organizational and thus traceable and legible process; to cultural and political sociology by unboxing how the meanings we can recover are contingent on specific power structures, organizational processes, and the agency of professional archivists; to historical comparative sociology by providing a framework for thinking about comparison of distinct datasets; to ethnography by tracing forms of inscription and performativity; and to sociological methods in strengthening how we can think about what it means to conduct "good" research based on inevitably "bad" records.

Inferences from Relics?

Assessing the evidentiary value of already existing material records in making truth claims about the social world poses a significant methodological challenge. In his critique of grand historical sociology, Goldthorpe indicated that the ability to make statements about "historical facts" in large parts relies on making "inferences from the relics" (Goldthorpe 1991, 213). Compared to scholars who can generate evidence by designing and collecting their own data, historians in this view were stuck with "objectified communications," frequently textual documents, whose defining characteristics are finiteness and incompleteness. What Goldthorpe's methodological bifurcation ignores is that scholars working with archival records are by no means bound by the limits of one particular archive, nor are they restricted to using documents "as is." In fact, those taking recourse to archives as a main data source do frequently generate new data. Examples include the combination of existing documents into data sets that can be analyzed in ways inaccessible to those looking at each document in isolation;[2] the generation of new data altogether by interviewing those mentioned in the files in oral history and others forms of interviews; and the retrieval of supplementary data sources that are not confined to archival holdings. Thus, the incomplete relics charge posited by Goldthorpe, in contrast to ostensibly methodologically more satisfying and data-rich forms of generative sociological fact production, is untenable. More than that, it falsely suggests a dichotomy between scholars passively using archives and those actively generating data.

Still, Goldthorpe's well-taken critique of grand historical sociology as exemplified by Theda Skocpol and Barrington Moore that links evidence and argument in tenuous and arbitrary ways, generates methodologically and hermeneutically productive points of departure (Goldthorpe 1991, 222, Moore 1966, Skocpol 1979). For example, the critique of finiteness and incompleteness ignores that documents generated by others have a number of merits that any data we might generate do not have—from the fact that the documents hold immense amounts of information that transcend the informational value contained in the texts they store, to what historians see as the evidentiary gold standard: pulling meaning out of documents written for a different purpose. The extraction of supplementary information exceeding textual information as well as peripheral reading practices are what make for a sociological use of archival records, instead of confining the research process to a purely extractive and positivist one. A qualitative sociological take on archives draws on research on documents and the archive in sibling disciplines[3] and leverages sociology's strengths of foregrounding questions of power, knowledge, and subjectivity.

---

[2] This poses technical challenges because files rarely have the same document structure; see Salminen et al. (1997).

[3] These include but are in no way restricted to history, anthropology, the history of science, documentation studies, archaeology, and library and information studies.

Life on File attempts to contribute to the sociological foray into archival research by setting out an argument about epistemology in archival research that disentangles processes of knowledge construction and organization around the recovery, reconstruction, and depiction of events, processes, and persons. As such, the project stands in opposition to positivist ways of treating the archive as repository of facts and thus an uncomplicated, if frustratingly incomplete, source for information extraction. Availing ourselves of Tilly's tripartite division, a phenomenology of the archive and of files seeks out:

1. The *recovery of life*—whether persons, processes, or events—from files, where life is dislodged from paper, and persons and processes are reanimated for disparate sakes;
2. The creation of documentary evidence from social life, that is, *the turning of life into a record*;
3. The *archivization* of this documentary evidence, which follows the life of the file as it moves through institutional, physical, and intellectual spaces.

My analysis begins in many ways at the end, with big questions about power, politics, and knowledge. What political uses and abuses are documents and archives pulled into? And what are the resulting implications for the contemporary and asynchronous use of files by scholars? Second, what kind of an object is a document? As objects generated by persons and institutions, documents go beyond the mere text and have performative properties. Third, I explore archives as organizations with interests whose actions establish a second performative layer on top of the documents themselves. The conclusion spells out methodological implications for comparison, theorizing, inference, and case selection.

Life through File: State, Power, and Archives in the Making of Human Kinds

Lives are acted upon in a very real way when traces are left on paper. It is no understatement to suggest that the Nazis were preoccupied, if not obsessed, with technologies of person construction. In the National Socialist (NS) regime's use of archives, being found in the archive could be advantageous if you wanted to prove Aryan status, and not being found could be detrimental; on the other hand, being recorded in religious registers could spell deportation and death, as the regime could use the records for identification, deportation, murder (Adler 1974; Majer 2003). After all, the NS project of ethnic cleansing required the positive identification of those the government sought to enslave or murder, but also those it sought to assimilate as German in the hopes of resettlement, repatriation, and germanification. Perversely then, the giving of life—by creating a record that was preserved in the archives—could also lead to the taking of life, for those whose existence had been tallied on paper for deportation and murder.

States and their Archives

Files, record-keeping processes, and archives are crucial elements in the establishment of modern state power. Writing about the rational organization of authority relations in the form of democracy, Max Weber posited that the "management of the modern office is based upon written documents (the 'files')" (Weber 2013 [1922], 957), and that bureaucratized administration resulted in a "system of domination that was practically indestructible" (Weber 2013

[1922], 987). As sites for "authorized deposits" of documents (Ricoeur 2006 [1978]), archives are drawn upon by governments to surveil and to control, and by civil rights advocates as evidence in post-disaster circumstances to make claims against the state. Archival records can stabilize or topple modern states, sustain genocidal regimes, but also aid in the construction of new forms of citizenship after natural disasters or wars (Fassin and D'Halluin 2005, Hull 2012, Ogborn 2007, Petryna 2013, Raman 2012, Weber 2013 [1922]). The German Archives Administration of the General Government was extremely conscious of the importance of archival documents for governance of East Europe when it euphemistically remarked: "The historical importance of the German constructive work in the East requires that. .. the documents that accompany the events be carefully preserved and collected as the permanent witnesses of German cultural planning and creating" (Posner 1944, 226–227).

Although often presumed to be under full control of authoritarian states, archival agency was in practice more ambiguous and cannot be entirely understood as the expression of unadulterated state power. Thus, the connection between states and archives is not always as smooth as Weberian arguments of positive reinforcement effects between administrative forms of record keeping and the expansion of state power would suggest. Taking the case of post-war French archives and their treatment of Vichy-era bureaucratic documents, Steinlight found an unstable, rather than the expected mutually constitutive, relationship between archives and the state (Steinlight 2017), and the expansion of German power over Eastern European archives reveals similar elements of contention concerning record access. Many members of the Eastern European clergy and of the local and regional archives did not share the National Socialists' enthusiasm for uncovering genealogical evidence for ethnic Germans, and the communication records between the Central Immigration Office in Łódź and the Race and Settlement Main Office are full of complaints about the lack of collaboration of the local file possessors, who refused to collaborate under the guise of overwork. Inter-ministerial communication from the period reveals ongoing institutional quibbles between the Ethnic German Liaison Office, the Race and Settlement Main Office, and the Interior Ministry around the perceived failure of institutional cooperation and its effects on the efficiency and speed of racial classification and assimilation. In response, the NS regime invested unprecedented resources in the expansion of German and newly annexed archives, which gave the Germans unprecedented access to local documents that would prove valuable in governance matters to reward loyalists from the past, and govern in the present (Demeter 1969; Posner 1944).

Archives contain a versatile array of documents that can be repurposed for different political projects in the present. In that, their use temporally extends beyond their contemporaneous role in establishing and expanding state power. In a monograph on dictatorship archives, historian Kirsten Weld showed how archives are in and of themselves sites of political struggle, and thus more than simple building blocks of politics. In Guatemala, the identification cards that one archivist called "paper cadavers" in need of "resurrection" (Weld 2014, 2) transitioned from being a tool of surveillance to a tool in the pursuit of war crimes. This reverse use of the very same documents has historical precedents; one instance is the transition in the use made of records of the German People's List (DVL) from their wartime use in making naturalization decisions, to post-defeat use in the persecution of wartime collaborators. The very same archives used for genocide would become a crucial tool in identifying perpetrators in the post-war period.

But states also use historical records to rid themselves of moral culpability. Fritzsche described archival projects as "the assumption that artifacts and documents can be made to tell very special, if incomplete, stories about social identity," with institutions using records to

establish particular trajectories of the past (Fritzsche 2006, 186). States may make strategic uses of the archive in order to rid themselves of historical debts and anaesthetize the past—in instances where the archive poses a possible threat to the state, the state attempts to control the narratives that can emerge from the files. Mbembe calls this process chronophagy:

> The function of the archive is to thwart the dispersion of these traces and the possibility, always there, that left to themselves, they might eventually acquire a life of their own. Fundamentally, the dead should be formally prohibited from stirring up disorder in the present. (Mbembe 2013, 22)

In this scenario, the dead are returned to life in controlled manners in order to silence them, rather than allowing unauthorized manners of return that could expose the state to attack. This controlling of narrative is facilitated by the selective preservation of records; similarly, states' refusal to accord countable status to their minority populations has been a well-honed tactic to deny rights to entire populations, and not only scholars of slavery have pointed to archival silences and omissions (Fuentes 2016; Hartman 2007; Thomas et al. 2017).

Temporal Dimensions of Files

Files age in ways that impact their political and scholarly use. The cases of "paper cadavers" and of NS records show that the files' political use can be reversed to suit new political goals. Organizations in turn generate sequences of documents over time, and the organizational practices of continuous administrative updating that result in a clean or updated copy after various initial drafts often come into conflict with the researcher's asynchronous search for documentary paternity in the shape of the unmutilated original (Vismann 2008: 45). In the case of National Socialist bureaucratic race thinking, for example, having had access to a progression of drafts including handwritten edits on the definition of race concepts and inter-ministerial epistolary communication on the various drafts has allowed for explaining shifts in racial classification by recourse to geopolitical events (Skarpelis 2019).

In addition to asynchronous reversals in the political use of files and the chronological progression of document versions, documents age in a third way: archived files remain static, while the world they still inhabit moves on. Philosopher Ian Hacking described how classifying people into "human kinds" fundamentally changes the people so designated. Classification having altered how humans perceive themselves as belonging to a certain group "induces changes in self-conception and behavior of the people classified," which in turn leads to the need for new forms of classification (Hacking 1995, 370). He deployed the expression "looping effects" to denote these cycles of classification, behavioral adaptation, and modification of classification. And while he wrote about the impact of scientific experts generating human kinds, defined as "a kind of experience about which scientific knowledge is claimed" (Hacking 1995, 369), the looping model can be transposed to state acts of classification, especially where ethnic and racial classification is concerned. At one extreme, Quayshawn Spencer claims that the census classification of race is so commonplace "that the US meaning of 'race' is just its referent, specifically the referent of US census racial discourse" (Spencer 2014, 1027). Less expansively, I suggest that state forms of classification, often themselves an outcome of a multi-stakeholder process, have a strong, but not exclusive, impact on self-perception and the creation of population groups (Mora 2014; Morning 2005).

The Administered German

Let us briefly return to the question of how history shapes the archive, and how the archive shapes history by drawing on an empirical case, that of personal files racially classifying metropolitan Germans as well as ethnic German Eastern Europeans under NS rule. Between 1933 and 1945, a new kind of German was created through questionnaires, in petitions for upgraded citizenship or Aryan "proofs." The latter could only be provided through a lengthy and expensive genealogical process in which administrators drew on church books to establish who would be spared from mass killing. In order to prevent the modification of church book originals, the appraiser of racial science within the Kinship Office and later the Volk federation of German Kinship Association began issuing identification cards for kinship researchers in order to restrict the scope of those allowed to access the archives—the necessary identification card was officially called a "harmlessness badge."[4] Eventually, the quest for racial classification expanded to Eastern Europe, where those the German nation wanted to include and assimilate had to be identified through existing genealogical and new racial classification attempts. The Nazis brought with them highly trained personnel to run the archives of the Occupied East, as well as liaison bureaucrats in the local bureaus of the Race and Settlement Main Office (*RuSHA*) who would interface with local archives and churches to obtain "racial" genealogical records (Huener 2014).

But the existing documentary form of records and their management did not map neatly onto the exterminatory goals of National Socialist Germany. Consequently, the regime invested significant resources into what would today be referred to as information science, including having open calls and competitions on how to best record, store, and centrally administer information. Already in 1933, the Reich Kinship Office held a competition soliciting submissions for a new system of racial categorization. Entitled *Suggestions for Racial Research*, the file contains several handwritten and type manuscripts by university-affiliated researchers outlining novel ways of racially categorizing people, designing racial classification cards, and cataloguing this information in the most efficient, accessible, and centrally manageable manner through smart filing systems (Reichssippenamt 1933). The first entry was a handwritten memorandum spanning forty-one pages on the suggested creation of a *Quellenamt* (source office), authored by a university student of theology from the North German city of Kiel. The author's concern lies with the efficient collection and processing of racial files. Beginning with the entry of the file through the mail room, the file was to pass through ten different offices, each responsible for a different part of the file, including registration, legal, political, medical, psychological, economic, foreign, and ancestry file (Reichssippenamt 1933). Although never implemented, Weidemann's processing solution would have fundamentally altered the materiality of the original file by adding to it marginalia, assessments, and so on, turning the file into a folder and modifying the person it contains from petitioner into fully appraised and racialized human.

New technologies and infrastructures reshape the materiality of files and the fate of those whose lives are banished on file. Implemented technologies of recording, sorting, and preservation generated a whole new set of files—questionnaires, statistics, memoranda, or letters—that are now preserved in archives. Far from a monodirectional story of states extracting information from archives, the process of racial classification shows how an authoritarian regime with clearly stated goals of person recording and tracking actively solicited suggestions for technological innovation and the collaboration of existing institutions like archives and religious organizations, to pursue the twin goals of racial classification and ethnic cleansing. Archives as well as files act.

---

[4] For the NS government's decrees on archival preservation of church books that begun in 1933 already, see Kayser (1939).

Recovering Human Kinds

It is time to carve out the methodological implications of seeing archives as sites of sociologically meaningful action in the shaping of concepts, classifications, and groups for some core methodologies deployed in archival research: descriptive inference and process tracing (Mahoney 2004). One concern for descriptive inference is measurement validity, specifically establishing that indicators used to compare cases remain consistent over time, a concern that is challenged by classifications and concepts that change over time and in response to classification. In a comparative and theoretical paper, "Parenthetical Logic: Towards a Sociological Theory of Bracketing, Absence and Erasure," I showed how states engage in linguistic subterfuge in order to rid themselves of undesirable populations; by renaming them and changing legal criteria for belongingness, the boundaries of citizenship and access to other social rights are quickly altered (Skarpelis 2020). We can account for these changes by making use of wandering conceptual frameworks (Hacking) or altering levels of abstraction (intension/extension) in order to retrofit the concept to match altered circumstances (Sartori 1970). The three temporalities of documents and the archive—reversal of a document's political use over time, sequences of documents, and documents aging out of utility in reflecting actual social identification—in turn are crucial for process tracing, a powerful tool for making causal arguments that relies on appropriate sequencing.

Understanding the temporally variable creation of documents from organizational and social life prevents the accidental omission of a set of relevant precursor documents, the failure to notice conceptual framework wandering, or changes in a concept's extension and intension. So far, we have established that records in and of themselves act on people in a multitude of ways, and that archives are meaningful sites of power possessing distinct temporalities. The next two sections address how the nature of documents plays into how life comes to be on file and how archives shape possibilities for recovery.

Life on File: Evidentiary Promise and the Burden of Proof

Documents are not natural kinds; they have a social context on account of being almost always intentionally recorded (Finn 2018). This article does not engage in a prosopographic account of the file as media technology (Lemov 2015), but rather asks how life comes to be *on file*, which necessitates asking questions about context and process of documentary creation, including the materiality of files, and their capacity for action. Before spelling out the methodological and theoretical implications of thinking through files, I will address each of the three concerns in turn.

Files, Files, Files

To state the obvious, files are not literal transmissions of "what happened."[5] Files are traces of *something*, but in order to assess traces of what, we need to understand their materiality, their processes of creation, and their embeddedness in contemporary forms of available media technologies. Although we often consider files to be a medium of delayed transfer (Vismann

---

[5] Even when they profess to, rhetoric and actual social circumstance may significantly diverge—the *lettres de cachet* Foucault describes (Foucault 1979) use "grand rhetoric (…) to dress up those trifling affairs" (87) and deploy the language of Racine and Bossuet to clumsily signal erudition, class, and status in order to obtain desired ends.

2008, 49 ff.), their relationship to oral pronouncements and social action is a complex one that may follow any one among a set of different logics. Rarely a stenographic account of social action or a replacement for communicative acts, they alter events through selective preservation and transmission in the process of documentary creation (even when they are closest in distance to social action). But in their proximity to quotidian action, files often unwittingly record additional information, and the physical scars they bear of handling and use—marginalia, stamps, burn marks on the pages—allow insights into action a transcribed digital record no longer contains. Documents as vehicles of meaning transcend a written record's purely textual content.

Social life is turned into documentary reality through specific acts and processes committed by organizations; files are curtailed records of interaction that act through selection, transmission, and storage (Vismann 2008). Putting something on the record means fixing one set of meanings over another. *Quod non est in actis, non est in mundo:* this tenet of Roman Law, that what isn't in the records, isn't in the world, also describes one of the fundamental challenges facing comparativists (Taeger 2002). In this process of the documentary creation of reality, persons can go missing (when files are destroyed in disasters or are deleted in a bureaucratic mishap, leaving the not-so-deceased scrambling to prove that they exist), be silenced (the voices of colonial subjects or slaves in colonial and plantation archives), or misrepresented (as in court reporters' mis-transcription of African American English) (Jones et al. 2019; Finn 2018; Fuentes 2016). This can happen as a matter of general principle, or in the temporal progression of documents from early drafts to modified and cancelled drafts, to approved "clean" copies, where the completion of the clean copy effaces the urtext and thereby relegates it and its contents to the realm of unreadability (Vismann 2008, 27).

Exactly how life ends up on file crucially depends on infrastructure, which—both as organizational form and as technology—is fundamental in shaping what records can be generated. Simply put, archivization connects to field structure. Derrida connected the shape of media technologies to the field's development itself, and in a move of "retrospective science fiction" speculated about the impact of technology on archival record creation and preservation. Not only is technology altering "secondary recording"—the preservation of already-produced records— but may also alter the evolution of the entire field: an alternate set of communication technologies "would have transformed this history from top to bottom and in the most initial inside of its production, in its very *events*" (Derrida 1995, 17). Media form the infrastructural basis of understanding, or, in the unequivocal words of Friedrich Kittler, "media determine our situation" (Kittler 1987). This is consequential particularly for comparative work. If documentary practices are historically particular, we cannot infer the presence or absence of a phenomenon based on surviving documentation alone. We need to understand the ecosystem of record production, including the specific prerogatives of organizations generating and preserving them, to gain insight into normative and moral preoccupations.

Files in and through Organizations

Documents become authorized deposits once vetted and chosen for preservation in an archive (Ricoeur 2006 [1978]). In that, they are connected to two organizations—the organization or entity that created them, and the one that stores them. The organizations generating records have interests, and consequently organizational logics and practices contribute to (re)shaping documents—they "do not describe the [organizational] order, nor are they evidences of the order, but rather stand as representations of them" (Trace 2002, 152). What records do is

represent a version of organizational operations that may be independent of actual practice, "a socially derived, persuasive, and proper account of the organization as an orderly enterprise" (ibid).[6] Latour christened files the "most despised of all ethnographic objects," and yet they may be the object closest to recording social action of the historical past (Latour 1986: 26). They are representations of the social order, and at the same time become a source of power by virtue of reducing complexity through inscription: "By working on papers alone, on fragile inscriptions which are immensely less than the things from which they are extracted, it is still possible to dominate all things, and all people" (Latour 1986, 29).

In the case of NS bureaucracies, "organizational" goals included the racial categorization and identification of persons victimized by the regime's merciless eugenicist drive. Organizational targets were not always bluntly declared, and particularly bureaucracies in charge of killing frequently took recourse to linguistic euphemism to retain plausible deniability post-defeat. Eastern European ethnic Germans appear in the German Federal Archives because the NS regime had a strong interest in identifying and resettling this diaspora for purposes of ethnic cleansing; genealogical records that were mostly kept in church and city archives became a tool for implementing these racial policies. Being found in the archive could conversely also protect from this new form of violence, ranging from harassment to murder—if one could prove non-Jewishness through the Aryan Confirmation. In colonial Japan on the other hand, although colonial subjects were registered in "ethnic registries," the information contained within were much less multi-dimensional than the German archives, ostensibly because what was of interest to the Japanese state was a blanket ethnic designation. In practice, this leaves one with paper trails of up to one hundred pages per person for the German case, and little more than basic demographic data, tabulated in endless military booklets, on Japanese colonial subjects. Radically unalike people emerge, fleshed out in different ways so that the utopia of direct comparison has to be discarded. Analogical approaches to comparison like Vaughan's or Simmel's that structurally align and map between distinctive domains and their parts are helpful heuristics to deal with these issues, as they separate social content from form (Vaughan 2014, 64 ff.).

Dorothy E. Smith referred to the phenomenon of treating the text as internally determined structure of meaning as document time, an instance in which the text becomes fixed as a social accomplishment (Smith 1974). If files fix social life, they also help construct the person. It is interactional approaches within organization studies and sociological theory that reveal how documents act as technologies of reification: documents capture and fix people and do constitutional work through "the routine textually-mediated practices of people engaged in their daily activities" (Cahill 1998, 143, Kameo and Whalen 2015, 210). Documentary practice combines with archival policy—whether a person at all appears in the archive offers some clues about their status. That ethnic Germans emerge as much more three-dimensional characters through the files than ethnic Koreans do in Japanese archives has nothing to do with any characteristic pertaining to the persons we are seeking to reconstruct themselves, and everything to do with organizational practice, file creation, and archival preferences and practices. The resulting differential availability in quantity of files does not mean that the Japanese were unconcerned about racial classification, but that they in many cases drew blanket conclusions for entire populations. Interested in how fascist bureaucrats conducted their racial-anthropological fieldwork, I had to turn to scientists employed in folklore museums

---

[6] It isn't only organizations that create records—individuals record others and themselves as well, as the recent rise of the quantified self has shown. See Lemov (2017) and Gorichanaz (2019).

for more detailed and individualized descriptions, as well as to the few diaries of forcibly assimilated Taiwanese and Koreans that have been translated into Japanese—the state archive became point zero for my exploration and had to be complemented by research in private collections and museums, often outside of Japan.

But even a strong interest in population control does not necessarily lead to an actuality of individuation, to a population appearing in the archive in any detail. In *Dispossessed Lives,* Fuentes retells how the only voice she could find of the enslaved women of eighteenth century Bridgetown, Barbados, was court records' reference to their screams, which then became "the historical genre of the enslaved in the colonial archive" (Fuentes 2016, 143). But the subaltern and exploited are not the only ones missing from the archives; often, so are the super-rich and other elites, and their material situation and subjectivities have to be reconstructed indirectly. Whether we call them "the missing of archives" or "documentary orphans," what they have in common is a profound disconnect to their lives being found and reflected in archivally preserved files.[7]

Inference and Transcending the "Mere" Record

What makes for a "good" file? In "'Good' organizational reasons for 'bad' clinical records," Garfinkel described how medical case folders change in meaning depending on their operational context (Garfinkel 1984). Useless for actuarial purposes, the case file had the potential to become "good" when read as set of documents on which a potential therapeutic contract could be based. Here, the good use of files—a use in which they become both legible and useful—presumes a basic form of competent readership that rests on culturally specific knowledge "attuned to its ideological, symbolic, and metaphorical meanings" (Garfinkel 1984, 197 ff.). This skill is as relevant for scholars as for organizational actors, as we saw when Naomi Wolf misinterpreted Victorian court records noting "death recorded" as indicating a death sentence imposed on gay men, when in fact they denoted the opposite—the judge's recommendation of pardon (Adams-Campbell et al. 2015, 114, Cohen 2019, Sengal 2019).

Competent readership is a prerequisite for adequate hermeneutic access to meaning, but establishing evidentiary credibility to make truth claims about action and people past requires more if one wants to avoid a disembodiment of the source.[8] Transcending the mere record requires being cognizant of contexts of documentary production and dissemination. As sociologists, we like to deploy documents as material evidence and proof for something having happened. And where files record an event simultaneously, a good case can be made for treating them as convincing evidence. Milligan developed a roster of internal and external criteria by which to determine the evidentiary value of the source. After establishing authenticity of the source (Was it forged? How can we prove authorship?), internal criticism adjudicates content by asking about the context of the document's production: Was the author present at the event they record? If not, what was the physical and temporal distance? Who was the intended audience, and in what state of mind was the author (Milligan 1979)?

---

[7] The editors of this special issue helpfully pointed out that qualitative sociologists have found different workarounds to access information on these missing voices, whether through direct access, as is often the case in ethnographies of the poor, or indirectly for the rich (the moral implications of this differential treatment are significant and deserve more explicit treatment).
[8] Ragin's concerns about disembodiment and obscuring of cases in variable-oriented comparative work in his analysis of the sociological "case" has shaped how I think through how sociologists use files. See Ragin (1992).

Files in many ways act as protocols of reality that reverse the burden of proof—what isn't in the files needs to be otherwise proven to have happened. Where a literal reading of the file cannot provide this, the gold standard is finding documents that are, in Marc Bloch's felicitous expression, "witnesses in spite of themselves" (Bloch 1953). In this lucky circumstance, a document is used as evidence that far transcends the original purpose of the document; Milligan calls this peripheral reading, and Fuentes, reading *along the bias grain* that uses the sources "for contrary purposes" (Fuentes 2016; Milligan 1979). Returning to the beginning of this article and to Goldthorpe, one advantage of drawing on data one did not create oneself is that this double distance of authorship and of documentary intent provides an additional layer of credibility.

While documents constrict the circumference of what can be said through them, they "do not (…) prescribe what may be said. Historians have a negative obligation to the witnesses of past reality" (Koselleck 2004, 111). The recovery of truth is impossible—after all, social action in real time has passed, and reconstructions of events "live(s) on the fiction of actuality" (Koselleck 2004, 111). Sources control what may *not* be stated but provide little guidance beyond this as to how to recover an interpretation of the past. Koselleck uses the example of Louis XVI's beheading, in which the only certain "fact" was that a guillotine separated his head from his body. But whether he was executed, murdered, or punished, was a *historical* question, not one of facts. A more subtle case is described by East German author Christa Wolf, who came under severe public criticism after an opening of her Stasi files in the 1990s revealed that she had been a minor Stasi informant. Writing that the "perverse mountain of files has turned into a kind of negative grail, to which one makes a pilgrimage in order to experience truth, judgment, or absolution" she reflects that "Nothing better, really, could have happened to the Stasi after the fact: banal, narrow-minded file administrators and information fetishists turn state's evidence and receive once again, in some cases now truly for the first time, the power to judge the fates of human beings" (Gitlin 1993). Whether ample or scarce, archival records are pulled into a number of political projects. Where Christa Wolf's experience is one of a small number of ambiguous files given disproportionate weight, there are instances in which in spite of ample archival records, an engagement with the past is refused, ostensibly on account of the evidentiary limits of the files themselves.[9]

Let us return to the social context in the production of files. States of unrest and exception, like disasters or post-war reconstruction, make information orders—a heuristic used to denote the networks and institutions undergirding knowledge formation processes sustaining power—more easily visible than during ordinary times (Bayly 1999, 4–6). Information orders crucially inform how events are recorded; Megan Finn calls the resulting possibilities for knowledge that emerge in the interplay of infrastructures, institutions, technologies, and practices of document creation and sharing, event epistemologies (Finn 2018, 3). In the following, we will move on from files and the organizations that create them to look at a specific part of the information order: the archives themselves.

Life of the File: Reading the Archive along the Grain

Henry Rousso declared historical memory to be structured forgetfulness (Rousso 1994). Archives, growing in size and complexity with the nation-state, are part of "a whole epistemological complex"

---

[9] The history of German sociology under National Socialism as much as the American history of slavery until the 2000s, are cases in point: see Christ and Suderland (2014), Klingemann (1996), and Patterson (2018).

that serves to both classify and legitimate knowledge (Featherstone 2006). They are sites of power that influence the production of historical knowledge, "a set of knowledges in and of themselves" (Adams-Campbell et al. 2015, 111); in that, they are just as social as files that act as agents of collective memory. This section is concerned with archives as organizations peopled with professionals managing records; it concerns archivarial agency and the relationship between history and the archive. Archives and archivists hold a number of roles regarding the files: as triage agents who decide on which files get preserved; as intermediaries between the records and their users who shape what becomes collective memory; as collaborators and gatekeepers in their interactions with the state. Understanding how archives develop as actors and as organizations in their respective historical, social, economic, and cultural contexts calls for reading them *along* the grain, rather than against the grain; it means understanding their organizational logics as well as the larger informational landscape they are embedded in.

What Archives Do all Day

Archives provide material infrastructure to data – they assemble files into collections that render life comparable, and frequently transform objects-as-data into data-as-objects that can be transported and reused off-site (Strasser and Edwards 2017). Archives and their archivists are "intermediaries between documentary evidence and its readers" (Hedstrom 2002, 21). As organizations bound by the logics of their profession, they fall under national and regional laws that determine operating procedures at large but are also shaped by the objectives of archiving professionals. The fate of files hosted in archives is reliant on national laws governing what archives should preserve, as well as on access provisions and restrictions that render some files unseen for generations. The majority of potential files are denied institutional preservation, and while we may think of archivists as guardians of documents, the more crucial task of archiving is triage: identifying among the files brought to the archive those that do not need preservation and can be destroyed or kept indefinitely with only minimal cataloguing. What ends up in archives in the first place is thus already filtered multiple times.

If the archive produces history, history also produces the archive: Economic history work on the 1427 *catasto* (tax census) of Florence shows how the emergence of a new state capacity—taxation—was enabled through an archive that was itself the product of an emergent set of state capacities (Padgett and McLean 2011).[10] Similarly, the NS takeover of German state archives included requests to "Aryan" Germans to produce documents on their identification; this call for auto-archivization was part of an attempt at altering and fundamentally rebuilding the content of archives. Conversely, it is not only inclusion but also knowledge's removal that is governed by state interests and implemented through archival rules. The modern secrecy system that keeps US national security information safe is an instance of what Galison calls antiepistemology, a research field asking how knowledge can be *obscured* through classification (Galison 2004, 237).

Hide and Seek: New Archives out of Shambles

State archives are driven by archival projects that are bound up with specific and national interests; it is no coincidence that the consolidation of the nation state so often temporally

---

[10] I thank Lis Clemens for suggesting Padgett and McLean's work, and for suggesting to think about archives as informants.

coincides with the creation of state archives. In Germany, the coming of age of the country's first centralized state archives is tied to the quest of producing a national identity in the absence of formal national unity and against the threat of France. In the pursuit of an archival establishment of German identity, Germany's medieval past came to stand in for a common heritage and became the precursor of a united nation in the early to mid-nineteenth century. But it was only the joint events of the demise of Prussian authoritarianism and the advent of total war in World War I that pulled together distinct ministries' records in centralized state archives. Weimar Germany in turn again altered archival practice through instituting various parliamentary commissions to analyze the German revolution, culminating in the creation of the Reichsarchiv, the Imperial Archive, in 1919 (Fritzsche 2006). Beyond an investigation into the causes of World War I, a more general feeling of insecurity about what Germany was and who Germans were led historians to collect documents and artefacts from ostensibly unspoiled Germans (these turned out to be populations living in the German-Polish borderlands, perceived as untouched by urbanization).

World War I, with its drastic rise in epistolary communication, altered what was considered to be files appropriate for preservation, leading to the conservation of letters, poems, and other similar popular artefacts because archivists began considering these popular documents as historical evidence (Fritzsche 2006). What counted as a record was vastly expanded, and the absence of letters from ordinary people prior to WWI only tells us that they were not valued as records, not that they did not exist. Similarly, if Nazi killings and Japanese colonial atrocities are less frequently recorded than observations of genealogical presence like the notorious Aryan Confirmation, this is a sign of power and not a sign of an absence of the phenomenon in question. Archives are not "an objective representation of the past, but rather (…) a selection of objects that have been preserved for a variety of reasons (…)" (Manoff 2004, 14). In 2020, the Federal Archive in Berlin has a good number of records detailing Nazi logistics of toilet paper distribution in the Occupied and Annexed Eastern Territories, but only a handful of files of documentation on racial affirmation and denial letters.

The same type of file, say, racial examination records for colonial soldier recruits, will be preserved in one institutional context by the Occupation power, but not in another. Part of the reason why files on National Socialist governance are more easily accessible than those of other authoritarian mid-century nations is that the United States and other occupation powers had an interest in preserving documents in order to prosecute war criminals. As the Allied Powers were gearing up for military tribunals like Nuremberg, their interest in identifying individual war criminals led to the preservation and detailed filing by name of perpetrator record.[11] Supreme Court Justice Robert H. Jackson, appointed by then-US President Harry S. Truman as US Chief of Counsel for the prosecution of Nazi war criminals, decided to base 75% of the Nuremberg trials on documentary evidence, rather than on expert testimonials (Conference on Captured German and Related Records, Wolfe, and United States National Archives Records Service 1974). This has significant consequences on the collection and preservation of documents.

A decade after the Nuremberg Trial data collection, the American Historical Society in 1956 formed a collaborative committee on war documents with the US National Archives.

---

[11] Interestingly, national legislation impinges on whether or not researchers can get access to the very same files. In Germany, I was not allowed to see the SS files that had been preserved in the Berlin Document Center and were now housed in the National Archives and Records Administration in Maryland; the reason are German privacy laws. In the US, no such privacy laws obtained, and I was given free access to all files without any restrictions on their photographic reproduction.

Fearing that a deterioration in diplomatic relations between the US and Germany would impede access to the documents for future historians, a choice was made to microfilm the original documents before returning them to Germany. Due to budget constraints, the files were block filmed because triage decisions would have been too costly in terms of staff and working hours (Conference on Captured German and Related Records, Robert Wolfe, and United States National Archives Records Service 1974, 203). While the original files as well as microfilm copies were returned to the German Federal Archives, access to person files in Berlin is governed by local laws. I was able to access all person files freely and without permissions or reproduction restriction in College Park, Maryland, but was not allowed to access the same set of files in Germany in their entirety due to privacy legislation, even with special permission and after submitting a summary of my research project to vetting. This differential accessibility of records has implications for the production of historical comparative scholarship in different national academic fields.

Even where archival documents become historical sources directly relevant to state interests, they are often culled. After Germany's defeat, military personnel often took archival documents home to the United States as memorabilia. Similarly, some twenty years later, during the handover of the Berlin Document Center (BDC) archives by the United States to the German government, the staff in charge of microfilming the documents stole a significant number of files and sold them to collectors.[12] Despite the Nazis' obsessive archival practice, today's archival holdings have been minimized through various processes and have become what Fritzsche calls archives of loss, repositories whose defining characteristic is their incompleteness (Fritzsche 2005).

The Japanese National Diet Library's first deputy director, Nakai Masakazu, whose wartime activities had been followed and recorded by the secret police, tried to build the new library around the goal of creating a "vast memory apparatus for an independent citizenry" (Pincus 2011, 390). However, his attempt at "enlightening" the public was crushed when Nakai's term was cut short after a mere four years. The subsequent transformation of institutional goals was accompanied by the Occupation's failures at obtaining a significant number of wartime records. All in all, the 1946–1947 International Military Tribunal for the Far East charged only twenty-eight Japanese. This difference in how the two nations' citizens were put on trial had long-lasting effects on the preservation and accessibility of contemporary files. Whereas the Allied Powers seized millions of German documents—close to a million on SS members alone—an official Japanese estimate suggests that only about 470,000 items were seized and sent to the US for analysis (Muta, Shohei, Japan Center for Asian Historical Records, and National Archives of Japan 2007, 7).

Just like in Germany, where military sources and files were fairly quickly secured and collected, the Washington Documentation Center seized Japanese records quickly. However, a good number of military sources that were deemed lost and therefore unavailable for use during the International Military Tribunal for the Far East eventually re-emerged in the 1970s after having been hidden by the Navy General Staff Office (Muta, Shohei, Japan Center for Asian Historical Records, and National Archives of Japan 2007). After World War II and with the partition of Germany, East Germany used its records for leverage in order to pressure the United States for official recognition of East Germany as a separate country (Conference on Captured German and Related Records, Robert Wolfe, and United States National Archives Records Service 1974). This shows just how significantly the post-war fates of both Germany

---

[12] For details on the fate of BDC records, see Fehlauer (2010).

and Japan shaped what documents would survive and how accessible they would become to researchers.

Collusion, Complicity, and Other Forms of Archival Action

Archives are important historical actors in their own right, and "archivization produces as much as it records the event" (Manoff 2004, 12). Often entangled with the state through funding and archivists' educational requirements (Born 1950; Eckert 2007), archives reflect state interests even when they do not directly translate these into archival practice, whether due to resource constraints or alternative visions of how archives should be kept by archival leadership. What is certain is that between 1933 and 1945, identifying Jewish Europeans only became possible through surveillance and tracking that would take the form of mining church and local genealogical archives. This means that not only did archives become a complicit institutional actor in the Holocaust, but also that the Holocaust would not have been possible without archives. As Josef Franz Knöpfler, head of the Bavarian archival administration in 1936, succinctly put it: "There is no racial politics without archives, without archivists" (Ernst 1999).[13] In the deadly political economy of personal data, the Nazis generated data doubles, "chained to individuals" (Aronova et al. 2017, 9).

A straightforward lesson to take away from the different administrative histories of archives is that archives are actors in their own right that not only triage and preserve records, but are actively involved in the creation of data. Moreover, while regional and national legislation shapes the bulk of how archives work, external shocks—like a foreign occupation—introduces a transnational element into how archives do their work. Often, this means that preservation follows specific transnational or foreign interests, as was the case in the postwar trials of both countries (as we have seen, with vastly different results). Less straightforward is the impact of these different organizational and administrative processes for comparative work on similar phenomena of interest, such as racial classification in authoritarian states. Knowing how archives work(ed) means knowing what types of documents to look for and letting go of the fiction of finding "comparable" classification records required by what Ragin critiqued as variable-centered research (Ragin 1992). At the same time, knowing how archives operate as institutions or collaborators of the state within historically and nationally distinct information orders opens up opportunities for exploring alternative causal arguments and answering altogether different questions—in this case, about persecution, forced labor, and ethnic cleansing, and more generally about instances where knowledge and power intersect.

Conclusion

Anthropological and postcolonial engagements with the archive have challenged us to see state actors as "cultural agents of 'fact' production" and encouraged scholars to engage in ethnographic, rather than purely extractive, ways with the archive (Stoler 2002, 2010). This call is increasingly heeded in sociological methods, including the contributions of Kim, Lara-Millan, and Sargent, and of Wilson and Mayrl in this special issue. *Life on File* builds on these contributions as well as on parallel work in history, the anthropology of colonialism, and

---

[13] After the war, Knöpfler, like many of his fellow archivists without whom the Holocaust would not have been possible, successfully managed to exonerate himself from public culpability or any form of punishment.

library and information science, to create a sociological framework for analyzing the creation, preservation, and interpretation of archival documents. The framework connects documentary production to archivization and clarifies the multiple ways in which state and organizational power shape knowledge construction. It is my hope that the value of this can extend beyond comparative historical engagements with the archive—after all, "historicity is an ontological feature of human social life" (Brubaker 2003) that makes any archival record historical by principle.

The article's double aim was to connect documentary production to archivization and scholarly interpretation, and to ask how taking this relationship seriously can help our use of sociological methods and theorizing. On the most general level, I have argued that sociological archival work has to go above and beyond an immanent analysis of the textual source. The process of documentary creation and triage begins much earlier than when files hit the archive; if we only focus on the silences of the archive, we become capable of diagnosing omissions, but remain incapable of understanding the nature of the documents we encounter. As the methodological implications for inference and tracking change over time are developed throughout the main text, the conclusion hints at some consequences of "Life-on-File-thinking" for questions of casing, comparison, theorizing, and inference in archival-based qualitative sociological research.

Problematic archival research imposes preconceived notions on files that by their very nature cannot resist bad theorizing, nor attendant practices of thrusting reluctant records into ill-fitting and taken-for-granted units of analysis. Inscribing groupness into research design is an occupational hazard in sociology, where casing often goes hand in hand with generating a research puzzle post-hoc to justify and legitimate one's scholarly intervention (Brubaker 2003; Mears 2017). Rather than forcing an ample, messy, incomplete, and usually uncategorized set of archival records into impossible categories, we should heed Vaughan's pragmatic suggestion of availing ourselves of theories and concepts as "a heuristic tool to loosely organize the data," which allows for reasoning across complex cases later (Vaughan 2014, 65). Neither ethnography nor comparative historical sociology prescribe what comparison should look like—on the contrary, recent work in both fields is a testament to the method's disciplinary richness (Abramson and Gong 2020; Carpio 2019; Hanchard 2018; Pacewicz 2020; Starr 2019; Tavory and Timmermans 2009). Whether deploying abductive analysis, extended case method, process tracing, or other forms of the comparative method, Life on File has suggested multiple ways for producing informational value out of files that goes beyond that of the immediate textual or numerical particulars contained within. This is particularly pertinent where the study of action and classification unfolding over time and across national and linguistic boundaries is concerned.

Returning to Tilly's quest for reconstructing traces of real-life phenomena, we can readily acknowledge that like any other data source, archival records have limitations. But far from being found relics, they contain a treasure trove of information that exceeds the mere text.[14] Life on file is social: the files themselves take on different meaning depending on the site of preservation, and access options; their form and content by itself does not guarantee any particular type of interpretation. The relationship between files and the social world they purport to record is a complex one that can be rendered legible by unpacking the documentary creation process. If we think of authoritarian or colonial state records as biased fragments, and

---

[14] This poses a number of significant challenges to computational and quantitative sociological work that uses archival records that have to date not been sufficiently theorized or addressed.

the archives themselves as potentially unreliable informants, it becomes the task of the researcher to engage in acts of exposition, recovery, and rehabilitation that capitalize on, rather than see as detriment, the power relations contained within. Using archival records requires adequately stripping them down to their historical and archival-organizational scaffolding. Being transparent about contingencies in the production of the data we draw upon does not preclude theorizing, problem-driven causal argumentation, or thick description. On the contrary, recovering life from the archive demands "transcending the mere record" and vindicates a distinctively sociological approach to archival objects.

References

Abramson, Corey M., and Neil Gong, eds. 2020. *Beyond the case: The logics and practices of comparative ethnography*. Oxford: Oxford University Press.

Adams-Campbell, Melissa, Ashley Glassburn Falzetti, and Courtney Rivard. 2015. Introduction: Indigeneity and the work of settler archives. *Settler Colonial Studies* 5 (2): 109–116. https://doi.org/10.1080/2201473 X.2014.957256.

Adler, H.G. 1974. *Der verwaltete mensch: Studien zur deportation der Juden aus Deutschland*. Tübingen: Mohr.

Aronova, Elena, Christine von Oertzen, and David Sepkoski. 2017. Introduction: Historicizing big data. *Osiris* 32 (1): 1–17. https://doi.org/10.1086/693399.

Bayly, C.A. 1999. *Empire and information: Intelligence gathering and social communication in India, 1780–1870*. Cambridge: Cambridge University Press.

Bloch, Marc. 1953. *The historian's craft*. [1st American ed.] ed. New York: Knopf.

Born, Lester K. 1950. The archives and libraries of postwar Germany. *The American historical review* 56 (1): 34–57. https://doi.org/10.1086/ahr/56.1.34.

Briet, Suzanne. 1951. *Qu'est-ce que la documentation?* Paris: Éditions Documentaires Industrielles et Techniques.

Brubaker, Rogers. 2003. Beyond comparativism? Conference Manuscript.

Buckland, Michael. 1997. What is a "document"? *Journal of the American Society for Information Science* 48 (9): 804–809.

Cahill, Spencer E. 1998. Toward a sociology of the person. *Sociological Theory* 16 (2): 131–148.

Carpio, Genevieve. 2019. *Collisions at the crossroads: How place and mobility make race, collisions at the crossroads*. Oakland: University of California Press.

Christ, Michaela, and Maja Suderland, eds. 2014. *Soziologie und Nationalsozialismus - Positionen, Debatten, Perspektiven*. Berlin: Suhrkamp.

Cohen, Philip. 2019. Learn the right lessons from Naomi Wolf's book blunder: Expertise matters. But lane-policing is counterproductive. https://newrepublic.com/article/154019/learn-right-lessons-naomi-wolfs-book-blunder. Accessed September 1, 2019.

Conference on Captured German and Related Records, Robert Wolfe, and United States National Archives Records Service. 1974. Papers and Proceedings of the Conference on Captured German and Related Records, November 12–13, 1968. Athens, Ohio, 1975.

Daston, Lorraine, ed. 2017. *Science in the archives: Pasts, presents, futures*. Chicago: University of Chicago Press.

Demeter, Karl. 1969. Das Reichsarchiv Tatsachen und Personen. In *Mit 6 Bildtafeln*. Frankfurt am Main: Bernard und Graefe.

Derrida, Jacques. 1995. Archive fever: A Freudian impression. *Diacritics* 2: 9–63. https://doi.org/10.2307/465144.

Eckert, Astrid M. 2007. Managing their own past. German archivists between national socialism and democracy. *Archival Science* 7 (3): 223–244. https://doi.org/10.1007/s10502-007-9057-3.

Ernst, Wolfgang. 1999. Archival action: The archive as ROM and its political instrumentalization under national socialism. *History of the Human Sciences* 12 (2): 13–34. https://doi.org/10.1177/09526959922120225.

Fassin, Didier, and Estelle D'Halluin. 2005. The truth from the body: Medical certificates as ultimate evidence for asylum seekers. *American Anthropologist* 107 (4): 597–608. https://doi.org/10.1525/aa.2005.107.4.597.

Featherstone, Mike. 2006. Archive. *Theory, Culture & Society* 23 (2–3): 591–596. https://doi.org/10.1177/0263276406023002106.

Fehlauer, Heinz. 2010. NS-Unterlagen aus dem Berlin document center und die Debatte um ehemalige NSDAP-Mitgliedschaften. *Historical Social Research* 35 (3): 22–35.

Finn, Megan. 2018. *Documenting aftermath: Information infrastructures in the wake of disasters*.

Foucault, Michel. 1979. The life of infamous men. In *Michel Foucault: Power, truth, strategy*, eds. Michel Foucault, Meaghan Morris, and Paul Patton, 76–91. Sydney: Feral Publications.

Fritzsche, Peter. 2005. The archive. *History and Memory* 17 (1/2): 15–44,367. https://doi.org/10.2979/HIS.2005.17.1-2.13.

Fritzsche, Peter. 2006. The archive and the case of the German nation. In *Archive stories facts, fictions, and the writing of history*, ed. Antoinette M. Burton, 184–208. Durham: Duke University Press.

Fuentes, Marisa J. 2016. *Dispossessed lives: Enslaved women, violence, and the archive*: University of Pennsylvania Press, Inc.

Galison, Peter. 2004. Removing knowledge. *Critical Inquiry* 31 (1): 229–243. https://doi.org/10.1086/427309.

Garfinkel, Harold. 1984. *Studies in ethnomethodology*. Cambridge: Polity Press.

Gitlin, Todd, 1993. "I did not imagine that I lived in truth." *The New York Times,* December 28.

Goldthorpe, John H. 1991. The uses of history in sociology: Reflections on some recent tendencies. *British Journal of Sociology* 42 (2): 211–230. https://doi.org/10.2307/590368.

Gorichanaz, Tim. 2019. A first-person theory of documentation. *Journal of Documentation* 75 (1): 190–212. https://doi.org/10.1108/JD-07-2018-0110.

Hacking, Ian. 1995. The looping effects of human kinds. In *Causal Cognition: A Multidisciplinary Debate.* Clarendon Press/Oxford University Press.

Hanchard, Michael George. 2018. *The spectre of race: How discrimination haunts Western democracy*. Princeton: Princeton University Press.

Hartman, Saidiya V. 2007. *Lose your mother: A journey along the Atlantic slave route*. 1st ed. New York: Farrar, Straus and Giroux.

Hedstrom, Margaret. 2002. Archives, memory, and interfaces with the past. *Archival Science* 2 (1/2): 21–43.

Huener, Jonathan. 2014. Nazi Kirchenpolitik and polish Catholicism in the Reichsgau Wartheland, 1939–1941. *Central European History* 47 (1): 105–137. https://doi.org/10.1017/S0008938914000648.

Hull, Matthew S. 2012. *Government of paper: The materiality of bureaucracy in urban Pakistan*. Berkeley: University of California Press.

Jones, Taylor, Jessica Rose Kalbfeld, Ryan Hancock, and Robin Clark. 2019. Testifying while black: An experimental study of court reporter accuracy in transcription of African American English. *Language* 95 (2): e216–e252.

Kameo, Nahoko, and Jack Whalen. 2015. Organizing documents: Standard forms, person production and organizational action. *Qualitative Sociology* 38: 205–229. https://doi.org/10.1007/s11133-015-9302-7.

Kayser, Gerhard. 1939. Kirchenbuchfürsorge der Reichsstelle für Sippenforschung. *Archivalische Zeitschrift* 45: 141–163. https://doi.org/10.7788/az-1939-jg12.

Kittler, Friedrich A. 1987. *Grammophon, film, typewriter*. Berlin: Brinkmann und Bose.

Klingemann, Carsten. 1996. *Soziologie im Dritten Reich*. Baden-Baden: Nomos Verlagsgesellschaft.

Koselleck, Reinhart. 2004. *Futures past: On the semantics of historical time*. New York: Columbia University Press.

Latour, Bruno. 1986. Visualisation and cognition: Drawing things together. In *Knowledge and Society: Studies in the Sociology of Culture Past and Present*, ed. H. Kuklick, 1–40. Jai Press.

Lemov, Rebecca M. 2015. *Database of dreams: The lost quest to catalog humanity*. New Haven: Yale University Press.

Mahoney, James. 2004. Comparative-historical methodology. *Annual Review of Sociology* 30 (1): 81–101. https://doi.org/10.1146/annurev.soc.30.012703.110507.

Majer, Diemut. 2003. *"Non-Germans" under the Third Reich: The Nazi judicial and administrative system in Germany and occupied Eastern Europe with special regard to occupied Poland,* 1939-1945. Baltimore; London: Johns Hopkins University Press.

Manoff, Marlene. 2004. Theories of the archive from across the disciplines. *portal: Libraries & the Academy* 4 (1): 9–25 17p.

Mbembe, Achille. 2013. The power of the archive and its limits. In *Refiguring the archive*, eds. Carolyn Hamilton, Verne Harris, Jane Taylor, Michele Pickover, Graeme Reid, and Razia Saleh, 19–26. Dordrecht: Kluwer Academic Publishers.

Mears, Ashley. 2017. Puzzling in sociology: On doing and undoing theoretical puzzles. *Sociological Theory* 35 (2): 138–146. https://doi.org/10.1177/0735275117709775.

Milligan, John D. 1979. The treatment of an historical source. *History and Theory* 18 (2): 177–196.

Moore, Barrington. 1966. *Social origins of dictatorship and democracy: Lord and peasant in the making of the modern world*. Boston: Beacon Press.

Mora, G. Cristina. 2014. Cross-field effects and ethnic classification: The institutionalization of Hispanic panethnicity, 1965 to 1990. *American Sociological Review* 2: 183.

Morning, Ann Juanita. 2005. Multiracial classification on the United States census: Myth, reality, and future impact. *Revue Européenne des Migrations Internationales* 21 (2): 111–134. https://doi.org/10.4000/remi.2495.

Muta, Shohei, Japan Center for Asian Historical Records, and National Archives of Japan. 2007. Myth and reality about pre-World War II government records. *Japan-U.S.* Archives Seminar May *2007*: 111.

Ogborn, Miles. 2007. *Indian ink: Script and print in the making of the English East India company*. Chicago: University of Chicago Press.

Osborne, Thomas. 1999. The ordinariness of the archive. *History of the Human Sciences* 12 (2): 51–64.

Pacewicz, Josh. 2020. What can you do with a single case? How to think about ethnographic case selection like a historical sociologist. *Sociological Methods & Research*:0049124119901213. doi: https://doi.org/10.1177/0049124119901213.

Padgett, John F., and Paul D. McLean. 2011. Economic credit in renaissance Florence. *Journal of Modern History* 83 (1): 1–47.

Patterson, Orlando. 2018. American sociology's denial of slavery. *Trajectories - Newsletter of the ASA Comparative and Historical Sociology Section* 30: 17–23.

Petryna, Adriana. 2013. *Life exposed. Biological citizens after Chernobyl*. Princeton, NJ: Princeton University Press.

Pincus, Leslie. 2011. Revolution in the archives of memory: The founding of the national diet library in occupied Japan. In *Archives, documentation, and institutions of social memory: Essays from the Sawyer Seminar*, eds. Francis X. Blouin Jr and William G. Rosenberg, 382–392. Ann Arbor: University of Michigan Press.

Posner, Ernst. 1944. Public records under military occupation. *The American Historical Review* 49 (2): 213–227. https://doi.org/10.2307/1838607.

Ragin, Charles C. 1992. Introduction: Cases of "what is a case?". In *What is a case? Exploring the foundations of social inquiry*, eds. Charles C. Ragin and Howard Saul Becker, 1–17. Cambridge: Cambridge University Press.

Raman, Bhavani. 2012. *Document raj: Writing and scribes in early colonial South India*. Chicago: University of Chicago Press.

Reed, Isaac Ariail. 2010. Epistemology contextualized: Social-scientific knowledge in a postpositivist era. *Sociological Theory* 28 (1): 20–39.

Reichssippenamt. 1933. Anregungen zur Rasseforschung.- Private Mitteilungen, Berichte und Denkschriften. BArch R1509/1102.

Ricoeur, Paul. 2006 [1978]. Archives, documents, traces. In *The Archive*, ed. Charles Merewether, 66–69. Cambridge, Massachusetts: MIT Press.

Rousso, Henry. 1994. *The Vichy syndrome: History and memory in France since 1944*. Cambridge: Harvard University Press.

Salminen, Airi, Katri Kauppinen, and Merja Lehtovaara. 1997. Towards a methodology for document analysis. *Journal of the American Society for Information Science* 48 (7):644. https://doi.org/10.1002/(SICI)1097-4571(199707)48:73.0.CO2-V.

Sartori, Giovanni. 1970. Concept Misformation in comparative politics. *The American Political Science Review* 64 (4): 1033–1053.

Sengal, Paruh. 2019. Naomi Wolf's career of blunders continues in "Outrages." New York Times "Book of the Times"." https://web.archive.org/web/20190902011251/https://www.nytimes.com/2019/06/05/books/review-outrages-naomi-wolf.html . Accessed 1 September 2019.

Skarpelis, A.K.M. 2020. *Parenthetical logic: Towards a sociological theory of bracketing, absence and erasure*. Cambridge: Working Paper. Harvard University.

Skarpelis, A.K.M. 2019. *When whiteness fails: Racially supremacist thinking, classification and arbitration in National Socialist Germany, 1933 to 1945*. Cambridge: Working Paper. Harvard University.

Skocpol, Theda 1979. *States and social revolutions: A comparative analysis of France, Russia, and China* Cambridge: Cambridge University Press.

Smith, Dorothy E. 1974. The social construction of documentary reality. *Sociological Inquiry* 44 (4): 257–267.

Spencer, Quayshawn. 2014. A radical solution to the race problem. *Philosophy of Science* 81 (5): 1025–1038.

Starr, Paul. 2019. *Entrenchment: Wealth, power, and the constitution of democratic societies*. New Haven: Yale University Press.

Steinlight, Alexandra. 2017. The liberation of paper: Destruction, salvaging, and the remaking of the republican state. *French Historical Studies* 40 (2): 291–318. https://doi.org/10.1215/00161071-3761631.

Steup, Matthias. 2017. Epistemology. In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. https://plato.stanford.edu/archives/fall2017/entries/epistemology.

Stoler, Ann Laura. 2002. Colonial archives and the arts of governance. *Archival Science* 2 (1/2): 87–109.

Stoler, Ann Laura. 2010. *Along the archival grain: Epistemic anxieties and colonial common sense*. Princeton: Princeton University Press.

Strasser, Bruno J., and Paul N. Edwards. 2017. Big data is the answer … but what is the question? *Osiris* 32 (1): 328–345. https://doi.org/10.1086/694223.

Taeger, Angela. 2002. Analysis of records in historical research on criminal law. Criminal records on male homosexuality in Paris in the 18th century. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research: Qualitative Methods in Various Disciplines III: Criminology* 3 (1). doi: https://doi.org/10.17169/fqs-3.1.883.

Tavory, Iddo, and Stefan Timmermans. 2009. Two cases of ethnography: Grounded theory and the extended case method. *Ethnography* 10 (3): 243–263. https://doi.org/10.1177/1466138109339042.

Thomas, David, Simon Fowler, and Valerie Johnson. 2017. *The silence of the archive*. London: Facet.

Tilly, Charles. 2002. Event catalogs as theories. *Sociological Theory* 20 (2): 248–254. https://doi.org/10.1111/1467-9558.00161.

Trace, Ciaran B. 2002. What is recorded is never simply 'what happened': Record keeping in modern organizational culture. *Archival Science* 2 (1–2): 137–159.

Vaughan, Diane. 2014. Analogy, cases, and comparative social organization. In *Theorizing in social science the context of discovery*, ed. Richard Swedberg, 61–84. Stanford, California: Stanford University Press.

Vismann, Cornelia. 2008. *Files: Law and media technology*. Stanford: Stanford University Press.

Weber, Max. 2013 [1922]. Economy and society: An outline of interpretive sociology. Berkeley: University of California Press.

Weld, Kirsten. 2014. *Paper cadavers: The archives of dictatorship in Guatemala*: Paw prints.

Yeo, Geoffrey. 2011. Rising to the level of a record? Some thoughts on records and documents. *Records Management Journal* 21 (1): 8–27. https://doi.org/10.1108/09565691111125071.

A. K. M. Skarpelis is a postdoctoral fellow at the Weatherhead Center for International Affairs at Harvard University. Her first book manuscript, "Un/Wanted: Race, Sovereignty and the Governance of Expendable Populations," reframes the citizenship histories of Germany and Japan around the idea of expendable populations and shows how nationality principles are routinely circumvented when this serves geopolitical and other state interests. A cultural and comparative historical sociologist, her work on archives, National Socialist photography, race concepts, and computational methods has appeared in a number of sociological and historical venues and has received awards from the American Sociological Association, the Eastern Sociological Society, the Society for the Advancement of Socio-Economics and the École des Hautes Études en Sciences Sociales. Anna's second project, "Racial Vision: Failed Projects of Human Difference," brings together the fields of engineering and the history of science under a sociological umbrella and studies abandoned scientific projects whose purpose it was to ascertain somatic forms of human difference. Anna received her doctoral degree in Sociology from New York University.