



**Calhoun: The NPS Institutional Archive**  
**DSpace Repository**

---

Faculty and Researchers

Faculty and Researchers' Publications

---

2021

# The Provenance Problem: Research Methods and Ethics in the Age of WikiLeaks

Darnton, Christopher

Cambridge University Press

---

Darnton, Christopher. "The Provenance Problem: Research Methods and Ethics in the Age of WikiLeaks." *American Political Science Review* (2021): 1-16.  
<http://hdl.handle.net/10945/68519>

---

This publication is a work of the U.S. Government as defined in Title 17, United States Code, Section 101. Copyright protection is not available for this work in the United States.

*Downloaded from NPS Archive: Calhoun*



Calhoun is the Naval Postgraduate School's public access digital repository for research materials and institutional publications created by the NPS community. Calhoun is named for Professor of Mathematics Guy K. Calhoun, NPS's first appointed -- and published -- scholarly author.

**Dudley Knox Library / Naval Postgraduate School**  
**411 Dyer Road / 1 University Circle**  
**Monterey, California USA 93943**

<http://www.nps.edu/library>

# The Provenance Problem: Research Methods and Ethics in the Age of WikiLeaks

CHRISTOPHER DARNTON *Naval Postgraduate School, United States*

*How should political scientists navigate the ethical and methodological quandaries associated with analyzing leaked classified documents and other nonconsensually acquired sources? Massive unauthorized disclosures may excite qualitative scholars with policy revelations and quantitative researchers with big-data suitability, but they are fraught with dilemmas that the discipline has yet to resolve. This paper critiques underspecified research designs and opaque references in the proliferation of scholarship with leaked materials, as well as incomplete and inconsistent guidance from leading journals. It identifies provenance as the primary concept for improved standards and reviews other disciplines' approaches to this problem. It elaborates eight normative and evidentiary criteria for scholars by which to assess source legitimacy and four recommendations for balancing their trade-offs. Fundamentally, it contends that scholars need deeper reflection on source provenance and its consequences, more humility about whether to access new materials and what inferences to draw, and more transparency in citation and research strategies.*


National security leaks engender uncomfortable dilemmas for political leaders, citizens, and scholars. Barack Obama ran a “no drama” campaign in 2008 with tightly controlled communications and as US President prosecuted more leakers than all his predecessors combined (Dilanian 2019; Kornblut 2008). Nonetheless, in his final week in office, Obama commuted the prison sentence of Private First Class Chelsea Manning, whose leak of hundreds of thousands of classified documents had harmed US diplomatic relations, security, and intelligence, according to Obama’s secretaries of state and defense (Clinton 2014, 252, 368, 552–5; Gates 2014, 425–8; Savage 2017). The subsequent administration was further plagued by leaks and security breaches, often to Donald Trump’s ire, but frequently also at his instigation (Haberman and Rogers 2018; Rosenberg and Schmitt 2017). Although Trump repeatedly praised WikiLeaks while campaigning in 2016 (and earned the moniker “leaker in chief” from his staff), his raft of pardons after losing reelection four years later excluded Manning’s collaborator, WikiLeaks founder Julian Assange (Healy, Sanger, and Haberman 2016, 2018; Helderman, Dawsey, and Reinhard 2021; Parker and Sanger 2016). These cases challenge citizens to define justice and balance national security and the public interest, especially when leaked information includes not simply insider gossip but classified material. Scholars, too, debate the ethics of leaking and government secrecy (e.g., Cohen, Farrell, and Finne-more 2014; Feaver, Stanger, and Walzer 2018).

How should researchers approach the informational fallout from such leaks? The consequences and analytical challenges of treating unauthorized disclosures as social-

science evidence are extensive, complex, and underexamined (see, presciently, Drezner 2010). Political scientists who have addressed leaked documents’ research utility advocate freedom of inquiry, tout possible empirical revelations, and decry self-censorship (Gill and Spirling 2015; Michael 2015; O’Loughlin 2016) or emphasize human-subjects protections during research (Boustead and Herr 2020). In contrast, I argue that despite appearing to offer bountiful evidence, unauthorized sources are flawed and dangerous temptations, to which political science has too frequently and casually succumbed.

As with artworks and artifacts, how documents change hands after creation affects their meaning and value, with possessors and observers participating in the objects’ social history (Appadurai 1986, 34, 41–2, 56; cf. Higonnet 2013, 199–202). Sources’ provenance, their history of possession and transmission, generates methodological and ethical issues for researchers. These are especially acute when we recognize that specific works were, ultimately, stolen. Nonconsensual acquisition of documents complicates their analysis, raising questions of authenticity, legality, and selection bias. Moreover, engagement with those texts can reproduce real-world harms and condone security and privacy violations. As Baroness Ruth Deech argued in support of British legislation for restitution of cultural objects looted by the Nazis,

Art is an ethical issue. Displaying looted art, once it is known to be such, is not just an invasion of privacy and a demonstration that wrongdoers may indeed profit from their crimes; it is also putting on show something that the owners never meant to be seen in such circumstances. It has ceased to be an object of beauty and one that museums can be proud of or use for educational and aesthetic aims... . It taints the spectators who knowingly take advantage of the presence of the picture there and it speaks to them of loss and war, not creativity and insight. It is a well-known principle in physics that the act of observation changes the object observed and there is something of that principle in our viewing of looted art. (UK Parliament 2009; cf. Besterman 2014, 20–2)

Christopher Darnton , Associate Professor, Department of National Security Affairs, Naval Postgraduate School, United States, [cdarnto@nps.edu](mailto:cdarnto@nps.edu).

Received: October 21, 2020; revised: April 01, 2021; accepted: November 05, 2021.

Assange and his kindred are not the art-looting accomplices of Hitler and Göring, but disciplines such as art history have transformed their ethical practices regarding provenance in postwar and postcolonial contexts in ways that are instructive for social scientists. Given the proliferation of nonconsensually obtained evidence in security studies, particularly leaked classified documents, we should no longer elide the provenance problem but confront it to develop new standards.

This is an apt moment for reflection on documentary source legitimacy. First, the decade-long accumulation of published research since Manning's leaks is ripe for critical review. This work is pervasive and problematic: as I demonstrate, more than a hundred articles have employed leaked sources, in every leading journal, but only rarely with prominent disclosure and clear citations, let alone explicit discussion of evidentiary or ethical issues. Second, absent clear disciplinary standards, these publications steadily build norms that tacitly endorse further use of leaked sources and encourage future disclosures. Active deliberation of this trend is overdue: scholars concerned about leaked documents should contest their usage rather than just silently refraining from citing them. Third, the concrete problem of WikiLeaks in security studies should engage political science conversations on transparency and replication (Jacobs et al. 2021; Rinke and Wuttke 2021), and research ethics particularly regarding human subjects (APSA 2020; Boustead and Herr 2020; Kapiszewski and Wood 2021; Subotić 2021). How to handle primary-source evidence is increasingly recognized as a widespread and pressing concern.

In this paper, I first establish the problem's scope, outlining my journal sample and methodology and presenting descriptive data on publication patterns regarding leaked sources. Second, I critique existing guidance on source legitimacy from journals and professional associations. Third, I examine scholarly articles' apparent use of leaked sources, especially U.S. diplomatic cables, criticizing collective performance regarding citation practices, stated research designs, and ethics discussions. Fourth, I propose the concept of provenance as a foundation for improved professional standards, drawing insights from disciplines ranging from journalism to paleontology. Fifth, following John Gerring's (2001, esp. 22–31) "criterial framework," I operationalize eight provenance concerns, two primarily empirical (data richness, data reliability) and six predominantly ethical (legality, national security, public interest, policy relevance, human-subjects protections, and reflexivity). Sixth, I offer four suggestions to manage the inherent trade-offs between these criteria.

I conclude by urging greater humility, strategy, and transparency in political scientists' quest for more and better data. All sources have a cost. The provenance problem, as I see it, is that uncritical reliance on conveniently accessible but coercively or nonconsensually acquired documents casts a dark halo over the rest of our work. Research with such materials should be rare, prominently disclosed, carefully justified, and

extensively contextualized and cross-checked. Where scholars find sources and how we handle and label them affects what we can ethically and empirically claim and how others view our work. Scholars will make independent judgments in navigating these trade-offs, and further debate is inevitable. However, this paper clarifies the terms and stakes of those conversations. I float cautionary buoys over an array of submerged hazards, and I recommend only judicious departures from the main channel.

## LEAKS IN JOURNALS: METHODOLOGY AND PUBLICATION PATTERNS

Academic research that uses apparently leaked sources is widespread, persistent, and prominent. To track scholarly use of leaked sources since the Manning disclosures, I focused on the 20 journals listed in the 2011 William & Mary TRIP (Teaching, Research, and International Policy) survey (Maliniak, Peterson, and Tierney 2012; see Table 1). For scholars publishing WikiLeaks-based research in international relations and adjacent areas of comparative politics and American national security, this is a strong list of candidate outlets. It closely overlaps political-science journal rankings, including those measuring impact through weighted citations rather than reputation (Garand and Giles 2007, Tables 3–4), especially among top journals. I examined each journal's website for published guidelines about leaks, legality, and sources.

I built a dataset of articles mentioning either "Wikileaks" or "cable" (or both) published 2010–2020 in those journals. Full-text keyword searches using EBSCO, JSTOR, Project Muse, and ProQuest databases and publisher websites Sage, Wiley, Taylor & Francis, IngentaConnect, Lynne Rienner, Brill, Oxford University Press, and Cambridge University Press yielded 565 unique articles (see Darnton 2021). Overlapping searches produced varying results: not every portal catches keyword text within footnotes, references, hyperlinks, or larger words. I included shorter pieces (letters, reviews, editorial notes) but excluded early-view publications. I downloaded each article, searched its text for both terms, and examined all footnotes and references, to see whether leaked material was apparently cited. I did not access linked sources but relied on the text of notes and the sentences they supported. If articles mentioned diplomatic cables or leaked documents, and source material was unclear, or articles specifically mentioned online appendices, then I also consulted these.

From this review, I manually coded each article as 3 (*appears directly to cite, quote, or paraphrase leaked documents and/or material via WikiLeaks irrespective of origin*), 2 (*apparently refers to leaked material solely secondhand, via published articles rather than primary sources*), or 1 (*false positive: mentions the WikiLeaks organization without engaging leaked material or notes other cables, whether diplomatic, fiber-optic, or television*). I also read all 168 code-2 and code-3 articles in full. In 9% of articles referencing leaked material

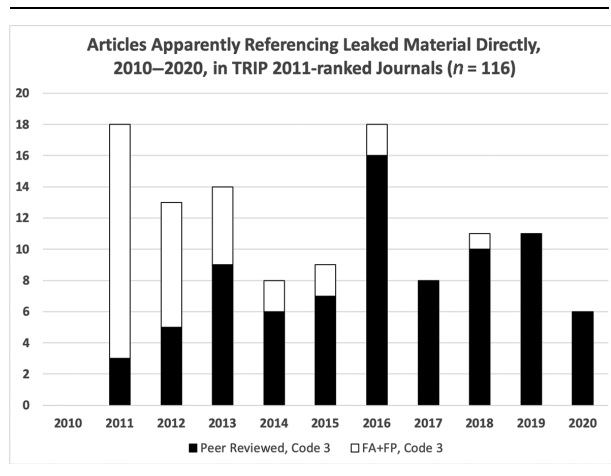
**TABLE 1. Journals Publishing Work with Leaked Material**

Title	Short Form	Code 3 Articles, 2010–2020	Code 2 Articles, 2010–2020	TRIP Rank, 2011	First Code 3 Article
Foreign Policy	FP	25	0	9	2011
International Affairs	IA	14	15	13	2011
International Security	IS	10	6	3	2012
Foreign Affairs	FA	10	0	4	2011
Security Studies	SS	9	4	14	2013
Review of International Studies	RIS	9	6	10	2013
International Studies Quarterly	ISQ	7	4	2	2012
Global Governance	GG	6	1	20	2014
International Studies Review	ISR	4	1	17	2013
European Journal of International Relations	EJIR	3	4	7	2012
International Organization	IO	3	0	1	2017
Review of International Political Economy	RIPE	3	1	15	2012
American Political Science Review	APSR	2	1	5	2016
Comparative Politics	CP	2	1	19	2017
International Relations	IR	2	0	18	2015
Journal of Peace Research	JPR	2	1	16	2017
World Politics	WP	2	2	6	2013
American Journal of Political Science	AJPS	1	1	12	2015
Journal of Conflict Resolution	JCR	1	2	8	2016
Millennium: Journal of International Studies	MIL	1	2	11	2011

(15/168), it was unclear either whether sources were primary documents or where they were obtained, so I also report my coding uncertainty. Among code-3 articles ( $n = 116$ ), I recorded whether leaked sources apparently included United States diplomatic cables. For those that did, in peer-reviewed journals ( $n = 64$ ), I coded further variables including number of cables cited, whether article body mentioned leaks or only references (or appendices) revealed this, whether articles mentioned documents' classification, what terms described WikiLeaks' role, and what information references included. The Appendix (see Darnton 2021) details all relevant information supporting these coding decisions.

Three findings stand out regarding journal publication patterns. First, an extensive body of work has employed leaked material: 116 articles cited it directly (code-3), and another 52 through secondary sources (code-2). This assuredly undercounts publications engaging leaked material including books, dissertations, and additional journals. It is impossible to measure self-censorship without the denominator, including scholars who could have cited leaked sources but declined. The TRIP survey (Maliniak, Peterson, and Tierney 2012, 44) provides a suggestive glimpse: 15% of scholars admitted using WikiLeaks for research (US respondents' figure was 10%; highest rates were France, 42%, and New Zealand, 31%). The scholarly community clearly has not scrupulously avoided WikiLeaks.

Second, by 2018, all 20 journals had published articles using leaked sources. Two of the most common outlets are *Foreign Policy* and *Foreign Affairs* (Table 1)—

**FIGURE 1. Articles Using Leaked Material**

editorially, rather than peer-reviewed—focused on current events and policy commentary, routinely featuring nonacademic authors, largely without footnotes. The most frequent peer-reviewed outlets—*International Affairs*, *International Security*, *Security Studies*, and *Review of International Studies*—comprise more code-3 articles (42) than the other 14 journals combined (39).

Third, the academic use of leaked material is persistent. The temporal pattern might suggest explosive interest following Manning's disclosures, gradually declining thereafter (total column height, Figure 1).

However, *Foreign Policy* and *Foreign Affairs* drive much of this (white columns). Among peer-reviewed

journals (*black columns*), direct engagement with leaks suggests the opposite: a gradual increase or steady annual production averaging nine articles, allowing a three-year research lag after the 2010 disclosures. The year 2016 saw the most peer-reviewed code-3 articles and 2019 the second-most. Additionally, rather than all journals publishing such work shortly after Manning's leaks, several only did so more recently (Table 1): eight published their first such articles during 2011–2012, five during 2013–2014, four during 2015–2016, and three in 2017.

## JOURNALS ON LEAKS: UNEVEN GUIDANCE ABOUT SOURCES

Research methods and ethics are matters of shared concern. Scholarly articles pass through peer review and editorial scrutiny, with a penumbra of responsibility regarding research ethics, legality, and scientific merit for journal editors, reviewers, publishers, and institutions. Thus, for context before evaluating how articles employed leaked sources, I review the limited and conflicting guidance available to their authors.

Academic associations and journals offer ambivalent stances, a decade after Manning's disclosures. The International Studies Association (n.d.; cf. Michael 2015, 176) professes willingness to publish research with still-classified documents but expresses "regret" that "some media are reporting otherwise," thereby recognizing that confusion is still rife; moreover, that ISA "does not have a policy rejecting" such work (due partly to the "cross-national, global complexity" of legal liability) is different than affirmatively defending such material's legitimacy. The *American Political Science Review* (2016) explains that it "will review papers that employ data whose legality is in question," yet passes responsibility onto authors, who "need to ascertain that they may legally use the data prior to the article appearing in print." Editorial interest in leaked documents' intellectual potential complicates this guidance: as the *APSR* editors noted (2017, iii), "One manifestation of 'big data' are (sic) massive data leaks, such as the Panama Papers or the US embassy cables"; an article employing Chinese government documents "illustrates how to use a massive data leak to answer social science questions," so editors "expect to see this research expanded to investigate other regimes."

Ethics are not reducible to legality, though. As the American Political Science Association (2020, 17–8) maintains regarding human-subjects research, scholars "should generally comply" with applicable laws and ethical "requirements may go beyond what the law ... may require"; conversely, researchers may violate laws given "reasoned justification ... based on ethical standards rather than convenience." APSA's (2020, 1) call for "openness and broader discussion" on research ethics is promising, and APSA and ISA deserve credit for posting even these limited guidelines. The other 15 TRIP-ranked peer-reviewed journals offer no public position regarding leaked sources, despite having

published such research. The *World Politics* (2018, 7) style sheet, though, explains how to format references to what it calls a "WikiLeaks cable." Ultimately, clearer positions on data legitimacy would help researchers resolve friction between laws and research ethics and slippage between treatment of human subjects and written documents.

## LEAKS IN ARTICLES

How have scholars proceeded in the absence of those standards? If leaks produce self-authenticating, excellent evidence, with negligible ethics concerns, then readers might expect robust defense of sources' research value, extensive documentary engagement, and detailed citations. Conversely, if scholars have concerns about sources' legitimacy or reliability, then readers might anticipate discussion of trade-offs and justifications. Either way, it should be straightforward—but is frequently challenging—to determine whether, why, and how scholars employed leaked material.

This criticism is not intended to name and shame individual authors, especially junior scholars, particularly since the explosion of online sources and the fog of ethics are so recent. Current guidance for authors is conflicting—and researchers, reviewers, and editors weighing these issues a decade ago had even less to work with. In this section, therefore, I highlight overall source usage trends rather than criticizing particular articles; the Appendix (Darnton 2021) provides information at the article level. (I also exclude *Foreign Policy* and *Foreign Affairs*, focusing on peer-reviewed journals.) Establishing patterns of practice regarding leaked documents and gaps between what articles say and what footnotes indicate demonstrates the urgency of further ethical and methodological discussion. Six issues are particularly salient.

## MISCELLANEOUS ORIGINS

Articles cited diverse leaked and hacked sources: corporate and governmental, US and foreign, military and civilian, domestic and diplomatic. Corporations include Stratfor, DynCorp, Blackwater, and multinationals doing business in Luxembourg. Documents appear from the governments of India and France, the United Nations High Commission on Refugees, and the International Security Assistance Force (Afghanistan). Several US agencies are included: US Trade Representative, US military video footage from Iraq, Joint Task Force-Guantánamo documents, the Iraq and Afghanistan "War Logs," Congressional Research Service reports, Secretary of State Hillary Clinton's emails, State Department incident reports from Iraq, Coalition Provisional Authority files, and Navy and CIA memoranda. Articles using leaks indirectly (code-2) suggest additional sources from China, Britain, Albania, Venezuela, the "Palestine Papers" and "tunileaks," alleged US intelligence files leaked by Edward Snowden, and John Podesta's emails.

Deciding what sources are fair game and which are off-limits for analysts is not a one-time problem from

the Manning disclosures. However, more than three-quarters of peer-reviewed code-3 articles (64/81) apparently employed at least one leaked US diplomatic cable; comparing these articles yields insights on how scholars addressed the dilemmas of leaked sources.

## FEW SOURCES

Of the 60 peer-reviewed articles referencing identifiable cables (disregarding two with specific cables only in appendices and two with cables not individually listed), half (30) cited a single cable, another eighth (8 articles) only two, and just a third (22, 37%) cited three or more (Figure 2). The top quintile (12/60) cited more cables than the rest combined. Published references may not reflect the extent of research conducted, though. Nor does citation quantity alone indicate analytical quality: specific projects or inferences might obtain great value from a single document (Bennett and Checkel 2015, 16–8).

However, readers need to know how and why authors approached a quarter-million cables, only to extract isolated citations. Further, given leaked documents' ethical concerns, scholars should consider whether they could perform substantially the same analysis without touching leaked sources at all. If this sparse evidence is unavailable elsewhere, and crucial for authors' analysis, that should be clearly indicated. Without clear methodological rationales, I find perplexing the frequency of uncritical reliance on solitary documents from a voluminous and problematic corpus.

## INCOMPLETE DISCLOSURES

In most cases (37/64 articles, 58%), readers would have no idea that leaked sources were used or WikiLeaks accessed, without checking notes, references, or appendices. Placing source and inference discussion in footnotes is often reasonable, especially to maintain case-study narratives. However, if authors engage sensitive or illegal material, prominent disclosure and discussion

in body text would help. Moreover, when articles mention using archival sources, government documents, or diplomatic cables but do not clarify that these included leaks, then stated research methods are misleading.

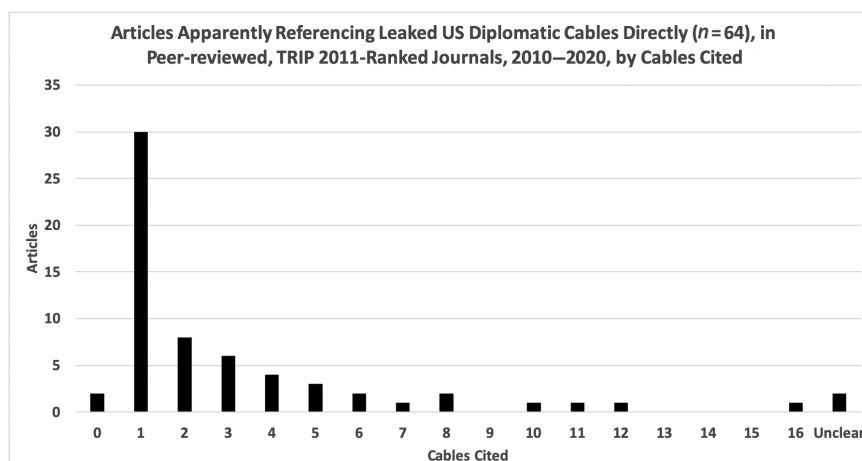
Similarly, only one-fifth of articles (13/64), even in footnotes or appendices, mentioned document classification (yet apparently cited classified material anyway); eighty percent cited similar documents either without recognizing, or acknowledging, that status. And, disregarding mentions of “WikiLeaks,” just one-third (21/64) noted their sources' leaked nature. Four articles, to their credit, acknowledged both leaked and classified aspects; thirty-four (53%) did neither. Neglect of document classification or leaks suggests incomplete understanding or analysis of source material, undermining empirical persuasiveness; conversely, disclosure of classified and unauthorized sources highlights the work's potential ethical precarity, requiring further authorial justification. (Intentional obscuring of source provenance would exacerbate both methodological and ethical problems.)

## OPAQUE REFERENCES

Collectively, citations fell short of transparency. Even when authors indicated leaked cables, hyperlinks to media outlets or organizations other than WikiLeaks often left citations unclear regarding whether authors engaged primary sources or secondary articles. Not all blame for truncated or misleading references (e.g., alphabetized bibliographies attributing cable authorship to WikiLeaks, not embassies) falls on authors: journal style sheets vary, and word count restrictions and parenthetical reference formats disadvantage documentary researchers (see Marc Trachtenberg, in Bütthe and Jacobs 2015, 14). But, problems extend beyond formatting.

Most articles (37/64, 58%) employed at least one cable where the reference did not identify the location, the US governmental origin, or the actual document.

**FIGURE 2. Articles Using Leaked Cables**



Specifically, nine articles (14%) discussed documents without citations, provided references without specifying source location, or just noted that texts are in the author's possession. If documents are not publicly available—perhaps obtained from personal contacts or through Freedom of Information Act (FOIA)—this should be disclosed. Cited documents without listed locations should be suspect, especially if these might remain classified or proprietary. Twenty-seven articles (42%; one overlaps with the previous nine) had at least one citation with a clear repository but unspecified government origin, including notes with unaccompanied hyperlinks and detailed WikiLeaks citations that omitted US authorship. Even if sources are linked and discoverable for replication, vague or implicit identification undermines persuasiveness.

Only three articles (/64, <5%), when citing apparently leaked cables, met three basic criteria: author (“Embassy X”), recipient (“State”), and document location (“WikiLeaks, [URL]”). Diplomatic cables are strategic communications, numbered in series, at specific classification levels, from embassies or consulates to the State Department and vice versa, written by and often addressed to individual diplomats, on specified topics. Good citations include this information and where authors found the source. Because our references do not sufficiently convey that communicative context, our analysis may fail to capture it as well.

## RARE RATIONALES

Why did scholars employ leaked material, and how did they address concerns about document legitimacy or reliability? Among 22 peer-reviewed articles citing three or more leaked cables, where WikiLeaks was evidently part of the research strategy, explicit methodological discussion of leaked documents' prospective empirical utility or limitations is sparse. Only one article extensively addresses document reliability: Elias (2018, 24–44) triangulates between leaked and public sources, considers what information might remain classified, and argues, “it is reasonable to expect the *WikiLeaks* records to be representative of” the “diplomatic processes” the article investigates. Predominantly, though, articles citing leaked cables apparently treated them as straightforward reflections of US policy or host country conditions.

Moreover, none of the 64 peer-reviewed articles employing leaked cables expressed ethical reservations, provided normative justifications, or mentioned potential illegality. Many scholars may not be contemplating sources' context and consequences or not disclosing this. Selection bias could be involved if researchers with ethical concerns simply avoid leaks. However, even authors without such qualms ought to defend their research strategies.

Some articles citing leaked documents other than cables provided clearer ethical position taking. One analyzes videos purportedly documenting human-rights violations by US forces in Iraq (Tidy 2017); leaks show “how dominant accounts of war can be complicated and contested” and how multiple perspectives are

“called upon as legitimate conduits for truths about war and are variously written into or out of accounts of war” (96–7). Further, Tidy (2017, 96–8, 102–5) critiques WikiLeaks' video editing and other source interventions, to resist taking official US stories or WikiLeaks counternarratives at face value. Another assembles data from leaked Joint Task Force-Guantánamo documents regarding detainees (Deutschmann 2016), declaring an “urgent need” for research on “prisoner treatment” and the consequences of prisoners' behavior for “how US authorities react.” A third uses US Trade Representative documents on the Anti-Counterfeiting Trade Agreement to analyze normative aspects of interest representation (Kuyper 2016), claiming that “interest in exclusion ... by governments and bureaucrats,” with information such as treaty drafts “deliberately obscured from public view ... violates standards of inclusive democracy” (318).

Notably, these rationales involve human rights and democracy, emphasizing governmental conduct rather than just documentary content. I am unconvinced that normative adherence to democratic deliberation and government transparency necessarily legitimizes scholarship with leaked classified materials or that the objective of improving human rights practices justifies any means of access to evidence. However, explicit authorial reflexivity facilitates more productive ethical debates than citations without such context.

## LEGITIMIZING EUPHEMISMS

Further complicating source assessment is what I see as legitimizing euphemisms, in 39% of articles (25/64). One-quarter (16/64) included positive verbs identifying WikiLeaks' connection to the cables (e.g., published, released) or adjectives about those documents' status other than classified and leaked (e.g., internal, private). These claims are accurate, particularly alongside admission of documents' unauthorized disclosure; besides, journals may influence citation phrasing. However, such diction suggests contestable claims about provenance and WikiLeaks' legitimacy. Distribution is not declassification, and availability is not authorization.

More problematic are references to “WikiLeaks cables,” which almost one-fifth of articles made (12/64), implying ownership or authorship, neither of which is true. This shorthand blurs the origin and disclosure of these sources and complicates scholars' ability to use them as evidence. The variation or omission of labels and what I read as hesitation or ambivalence about these documents' acquisition suggests that researchers are tiptoeing around the provenance problem rather than confronting it directly.

## PROVENANCE: AN INTERDISCIPLINARY FOUNDATION

Scholars need to reflect more carefully on the sources of our sources—rather than simply searching, citing, and moving on, we need to consider documents'

*provenance*. Here, insights from other disciplines should inform political science research methods. Despite extensive ethics guidance on interviews, experiments, surveys, and ethnography, political scientists working with documents are only beginning to chart this territory (Jacobs and Büthe 2021, esp. 190–1; Kapiszewski, MacLean, and Read 2015, 145–9, 172–87; Subotić 2021). And methods advice on archives and process tracing focuses on causal, empirical, and logistical concerns, with little attention to ethics (Bennett and Checkel 2015; Darnton 2018; Frisch et al. 2012). Even the transparency turn, which advocates precision on what sources are, why they were selected, and how they support empirical inferences (and which raises ethical considerations regarding scholarly opacity), has not emphasized the legitimacy of underlying data (Jacobs and Büthe 2021, supplement, II.1, 12–3; Moravcsik 2014, 665–8, 677–8). And leading journals offer mixed signals. Political science lacks clear bases for ruling any written sources as being out of bounds.

However, other fields emphasize that where scholars find a source matters greatly for how it should be understood—and whether researchers have rights of possession or access at all. In short, “Provenance is paramount” (Fraser 2018). Provenance is a major theme in archival management and archaeology and in collecting wine, art, books, and antiquities, where ownership history significantly influences market value (Gill and Chippindale 2007; Monks-Leeson 2011; Sweeney 2008). Opaque pathways between production and possession can render even authentic items worthless or fundamentally tainted. And the continued display of illegitimate objects casts a dubious pall over entire collections and institutions.

Anthropologists and museum curators reflecting on these problems increasingly advocate repatriation or restitution of culturally significant items, coercively or clandestinely obtained, to their original owners, heirs, or communities (Besterman 2014; Colwell 2015; La Follette 2017). Domains with greatest progress on normative agreement, legal provisions, and actual restitution along these lines are narrowly circumscribed: especially Nazi-looted art and cultural objects and Native American human remains and funerary items. Beyond these, masses of artifacts in private and public collections have gaps in provenance or uncomfortable legacies of looting or coercive exchange. Current museum practice requires that this history no longer be overlooked: provenance needs to be highlighted rather than finessed. Good-faith possession, diligently researched and openly maintained, can support objects’ retention, whereas nonconsensual acquisitions or opacity in investigating or disclosing these undermines claims to ownership or stewardship. Restitution debates over particular objects are contentious (consider the Parthenon Marbles [Doyle 2009; Rudenstine 2002]). Museums can reasonably claim public interest in display, preservation, and research, as well as realistic challenges in documenting transaction histories, but they face increased burdens of proof and diminished presumptions of legitimate possession, as against claims by previous owners or source communities.

Transparency is now necessary but not sufficient, as uncovering nonconsensual acquisitions has ethical consequences.

Historians maintain strong professional commitments to archival access and preservation (American Historical Association n.d, 2, 4), and appear less inclined than other disciplines to refrain from analyzing available sources, whatever their provenance. However, historians insist on deep contextual understanding and critical interpretation of sources, including sensitivity to archives as intentional collections with normative and power-laden dimensions. Selective creation, organization, destruction, and access of documents privileges some voices and themes and silences or obscures others—processes that are particularly salient in the records of colonialism and empire (see Burton 2005; Trouillot 1995). To read such archives against the grain requires systematic effort.

These reflections are not confined to the humanities or interpretivist wings of social science. Biologists and paleontologists increasingly weigh research revelations against the problematic origins of their specimens, from newly discovered blue tarantulas trafficked from Malaysia to hundred-million-year-old “blood amber” fossils from Myanmar conflict zones (Hunt 2020; Nuwer 2019). Fossil excavation in developing countries, younger scholars argue, should consult with local populations, involving participation, education, and repatriation, rather than colonial models of extraction and exploitation (Elbein 2021). The Society for Vertebrate Paleontology (2020, 2) has called on journals to “be mindful when handling manuscripts for publication that involve fossils from conflict zones” and to ban research with newly acquired Burmese amber, advising scholars “not to encourage a black-market for commercial trading,” which empowers violent actors. Moreover, the organization states (*ibid.*) that a core principle of paleontology is that only fossils “permanently accessioned and deposited in stable repositories within the public trust” are legitimate scholarly source material, whereas private collections “cannot be regarded as reliably available for study, cannot be considered part of reproducible science, and must not be introduced in scientific literature.”

Restraint comes, too, from the American Statistical Association (2018, 2–4): “The ethical statistician is candid about any known or suspected limitations, defects, or biases in the data,” and “In contemplating whether to participate in an analysis of data from a particular source, refuses to do so if participating in the analysis could reasonably be interpreted by individuals who provided information as sanctioning a violation of their rights.” Criminal justice, similarly, emphasizes the consequences of others’ access choices via the chain of custody: incriminating material (including digital files) from a suspect’s home might be thrown out of court as inadmissible if not obtained through lawful search or not properly handled thereafter; relatedly, material presented by prosecutors requires authentication against claims of manipulation, and using even one piece of fabricated incriminating evidence creates a presumption of innocence by implying the weakness



of the prosecution's case (Goodison, Davis, and Jackson 2015; Jones 2010).

Although journalism seems to welcome anonymous sources and blockbuster disclosures, media ethics actually identify numerous caveats and countervailing values, rather than providing blank checks for research. Professional journalists struggle to establish which truths are newsworthy and which leaks deserve publication. The documents they quote or publish are selected to support specific narratives, stories that reporters and editors have deliberately decided are in the public interest and that require exhaustive cross-checking. Media ethics experts argue that transparency must be “tempered by a commitment to responsible publishing and a concern for accuracy, verification, and minimizing harm” (Ward 2014, 54), along with other values such as context and support for community (Horner 2014, 194–200; McBride and Rosenstiel 2014). Moreover, editors consult government agencies for comment and context for impending stories and often agree to redact or withhold specific material (Foreman 2016, 86–7). On national security, journalists struggle to maintain independence without sacrificing access—and leaks are sometimes government sanctioned (Seib 2006).

These countervailing pressures affected newspapers' responses to disclosures brokered by WikiLeaks. *New York Times* Executive Editor Bill Keller expressed concerns (Keller 2011) about Julian Assange's motivations and selective withholding, coordinated with the US government about releases and stories, and sought to protect named US informants abroad from being endangered by unredacted documents. Alan Rusbridger (2018, 250, 312), editor-in-chief of the *Guardian*, recalled that Assange seemed happy to release the full tranche of documents from Manning and “let people around the world pick through the entrails of U.S. foreign policy,” whereas the *Guardian* wanted to be a more selective gatekeeper of public-interest newsworthiness; later, collaborating with the *New York Times* on Edward Snowden's leaks, Rusbridger pushed the *Times* to agree not to “use the archive as a bran tub to go fishing for stories unrelated to Snowden's primary focus.” Without equally careful ethical and inferential frameworks, scholars risk launching illegitimate or even illegal fishing expeditions.

### PROVENANCE: A CRITERIAL APPROACH

Several disciplines provide strong reasons to decline unauthorized sources, and they identify burdens of disclosure, deep interpretation, and public interest if scholars decide to proceed. To identify the major ethical and empirical issues at stake and how they conflict and to build toward disciplinary guidance, I follow John Gerring's (2001) “criterial” approach to research methods. This framework is particularly apt for its empathetic recognition of inherent trade-offs between scholarly values rather than a defined “rulebook” (22–3) while pushing political scientists to articulate their

strategies more explicitly and conscientiously. Ease of availability is weak justification for research design (29–30); scholars can disagree, and should make their case proactively, but they should expect that reviewers might prioritize other factors (26–7). This approach also comports with social-science ethics texts that advocate incorporating ongoing ethical practices rather than focusing narrowly on initial project approval or legal or Institutional Review Board (IRB) compliance (Fujii 2012; Israel 2015). The core issue, as the seminal Belmont Report (US Department of Health 1979) articulated regarding human-subjects protections, is how to balance potential harms and benefits throughout research.

In that spirit, I outline eight criteria for considering whether and how to analyze documentary sources of problematic provenance: data richness, data reliability, legality, national security, public interest, policy relevance, human-subjects protection, and reflexivity. The first two are primarily methodological, regarding the evidentiary value and inferential problems associated with leaked or hacked material. The latter six principally concern research ethics—the legitimacy, purposes, harms, and benefits of using nonconsensually obtained sources, and scholars' positions regarding documents' creators and brokers. Researchers may disagree over which attributes to maximize, but all approaches come with costs, and readers require persuasion. Dilemmas are irresolvable and scholars will contest fundamental issues such as whether cables on WikiLeaks' website are public-interest resources provided by whistleblowers or illegal materials that could harm national security and human subjects. Neither strict adherence to US laws, nor fixation on human-subjects protections, nor insistence on academic freedom, offers a straightforward path out of these quandaries.

### DATA RICHNESS

There is certainly a case for the evidentiary value of leaked documents. Focusing on originally classified material related to policy making reflects existing qualitative research methods guidance on process tracing, archives, and case studies, which emphasizes obtaining the best available evidence (Bennett and Checkel 2015, 18–9, 25–7; Darnton 2018, 18, 92–3, 110–6; George and Bennett 2005, 96–107; Trachtenberg 2006, 140–2, 153–7). Important channels for declassification and document release include the *Foreign Relations of the United States* series, mandatory declassification review by National Archives staff, and FOIA requests by researchers or institutions such as the National Security Archive (consult Brandon Rottinghaus in Frisch et al. 2012; Trachtenberg 2006, 252–5). Vast repositories of declassified sources now reside online on governmental and nongovernmental sites (see Connelly et al. 2020). WikiLeaks might appear to offer yet another collection of government documents, easily accessible and temptingly recent.

However, just on empirical grounds, even authentic and comprehensive tranches of leaked documents

present concerns. First, analysts must be careful not to read too much into individual sources, as classification implies neither accuracy nor authoritativeness and memorandum authors are not omniscient or objective. As retired diplomat Peter Galbraith (2011) explains, cables are crafted to command scarce attention at the State Department and perhaps the White House—salacious details, jokes, criticism of host-government officials, and narrative style are in, while truly sensitive matters are kept out. Is an exposé of foreign corruption factual, does it reflect embassy or US government positions, or is it rhetorically designed to focus attention on actionable problems? Moreover, documents' original classification may have talismanic appeal for researchers: the higher the better, with greater implied authority. These risks obtain with declassified documents too—but in archives, scholars can cross-reference and contextualize arrays of contemporaneous sources.

Second, those documents comprise better evidence for some questions than do others. Embassy cables, for instance, are not a clear window into either US policy making or political developments in host countries. If leaked documents comprise one type or series (outgoing cables, without internal memoranda or instructions), from one bureaucratic organization (State Department), scholars obtain a very partial picture of policy processes and perceptions. Thus, Robert Jervis (2015) critiques studies of torture and harsh interrogations based on CIA cables both for overreliance on a single source and because much of the debate took place orally and never made it into the files. Similarly, Bob Woodward argues, WikiLeaks “has been really overblown. Those documents are midlevel classification. They have virtually no standing in the White House, where decisions are made” (Glasser 2011). Importantly, if scholars rationalize using leaked documents because of low-level classification, then their sources do not reflect high-level decision making. Further, focusing on one classification level might bias the sample, as officials express different observations in different settings.

Last, given legal and ethical concerns with leaked documents, researchers analyzing such evidence should explain why alternative sources are insufficient. Scholars might argue that topics such as covert operations or ongoing trade negotiations have no methodological alternative to these materials. However, open sources can yield major insights even on sensitive subjects like targeted killings via drone strikes (Banka and Quinn 2018). And regarding host-country political and social conditions, diplomats' classified reports are not necessarily more accurate than, and frequently rely on, local media. Richness is relative: the empirical value of particular leaked materials for specific research questions needs to be argued, not just assumed.

## DATA RELIABILITY

Provenance concerns not just theft, but fraud. First, outright fabrication or manipulation of texts is worrisome with electronic leaks, though unlikely in official

archives (Trachtenberg 2006, 146–7). Verification is challenging because government officials are instructed not to confirm or deny leaked information. In the antiquities world, objects are often brokered by financially motivated, opaque entities, with rampant potential for forgery. Traffickers in looted goods are often willing to sell fraudulent ones; scrupulous collectors need consistently clean hands on both issues to preserve their reputations.

Second, even if individual documents are intact, selection effects are significant. If analysts do not know what was inaccessible to leakers or not released by them (or by intermediaries such as WikiLeaks or newspapers), this skews inferences from the publicly available sample (see Gill and Spirling 2015). For instance, when Daniel Ellsberg leaked the Pentagon Papers, he withheld volumes on US-North Vietnamese peace talks and kept back, for decades, other classified material he had illicitly photocopied at the same time (National Archives 2011; Savage 2021).

Leak motivations matter, and WikiLeaks is an especially suspect source. As Michael Walzer (2018, 58) argues, “In contrast to newspapers with long records of public service, WikiLeaks is the wrong kind of intermediary between a whistleblower and the American people,” with “narrowly partisan and personal aims.” One study suggested that less than 10% of the site's documents came from leakers or whistleblowers, almost 70% from hackers, and the remainder from open sources and FOIA requests—and that WikiLeaks withheld Trump-related documents while releasing Clinton material (Dorfman 2018). As unauthorized sources' authenticity and representativeness depends on the choices and goals of leakers, hackers, and intermediaries, researchers must consider not only empirical problems with the texts themselves but also the ethical baggage they carry. Ultimately, even reliable and useful data might not be worth employing.

## LEGALITY

Determining how to analyze leaked documents is secondary: the fundamental problem is whether to engage them at all. The US government's position on unauthorized disclosures is an important reference point: Manning was convicted (and Assange indicted) of violating the Espionage Act—disclosing classified information without authorization makes one not simply a leaker but a spy. It is unclear, even if reading leaked material does not constitute disclosure, whether publication of excerpts or information from those documents does: Espionage Act provisions, construed broadly, could encompass distribution (18 USC 37, §793, 798, <https://www.law.cornell.edu/uscode/text/18/part-I/chapter-37>). Consequences are more likely and more extensive—including loss of employment or security clearances—for government officials and contractors, including those in academic or research positions, who are explicitly reminded not to engage this material. One *Foreign Policy* commentator refers to post-WikiLeaks surveillance of federal employees as potential leakers or insider threats, by

their peers and supervisors, as “Orwellian” and an “inquisition” (Bamford 2016). This may be hyperbolic, but it is not hypothetical.

Formal liability may be lower for others. As private citizens, some scholars might see using publicly available material as minor, low-risk violations, akin to jaywalking or violating broadcast-television copyright notices. Society routinely distinguishes between johns and pimps, narcotics possession and distribution, and downloading music for personal use and operating file-sharing servers—although participating in such markets entails complicity with suppliers. And national laws vary: authors and journalists outside the United States might not worry about legal consequences from Manning’s leaked documents.

The ethics consideration, though, is not just whether research carries prosecutable consequences but whether scholars should knowingly exploit and endorse illegally obtained material or scrupulously avoid it. Whether documents are governmental, corporate, or personal, leaks and hacks generally mean dealing with stolen goods (see journalist Neil Sheehan’s recalled exchange, posthumously published, with Ellsberg about the Pentagon Papers; Scott 2021). This requires considering not just researchers’ home-country laws but also the context of documents’ authorship, ownership, and disclosure. A global discipline should advance global and reciprocal standards. APSA (2020, 17–8; likewise Boustead and Herr 2020, 508) guidelines permit research to cross legal lines, given sufficient justification, so legal concerns are not automatically paramount. However, they should be significant ethical considerations regarding leaked sources.

Hopefully no scholar would contemplate, as a research strategy, hacking into government systems to obtain classified documents. Why, then, are so many researchers apparently comfortable citing unauthorized material provided someone else stole it in the first place? The analogy to burglary and receipt of stolen property is uncomfortable but not farfetched. Many colleagues, during fieldwork, have experienced contacts or friends that helpfully but unofficially proffer collections of physical or electronic documents (see Lessing and Willis 2019). Before accessing leaked documents online, we might ask ourselves how confidently we would carry them through a departure terminal from their country of origin. Unprecedented and unauthorized access may be exciting, but scholars should tread carefully.

## NATIONAL SECURITY

Classified documents present special concerns, as leaks are not merely criminal but also harmful. U.S. presidential executive orders and Office of Management and Budget guidance unambiguously declare that leaked classified documents remain classified documents: unauthorized use and dissemination is illegal and harms national security (Lew 2010; National Archives 2009). Classification levels define the “damage to the national security” (i.e., “harm to the national defense or foreign relations”) that “reasonably could be

expected to result” from their disclosure, with Confidential signifying “damage,” Secret “serious damage,” and Top Secret “exceptionally grave damage” (National Archives 2009). Posting such material online neither constitutes, nor automatically produces, declassification (2009). Moreover, the US Government treats even some unclassified information as sensitive: Controlled Unclassified Information designations let executive agencies restrict dissemination, such as prohibiting release to foreign nationals, or identifying categories of sensitive content such as genetic, infrastructure, or nuclear information (National Archives 2019). Similarly, government employees watch for “classification by compilation”: juxtaposing specific pieces of information raises documents’ overall classification level; even some combinations of unclassified information must jointly be classified (Information Security Oversight Office 2018, 8).

There are reasons to doubt blanket assertions of national security prerogatives for controlling information access. Arguably, governments too often overclassify information and resist freedom-of-information requests (Aftergood 2011; Rudenstine 2016, 308–10; Shapiro and Siegel 2010), sometimes even reclassifying previously released material. Scholars might contend that certain documents are “only” Confidential or Secret; that materials are already available, with no additional harm from further use and that others will access it anyway; and that, years after particular disclosures, security risks have evaporated. Harm and risk might be small from specialized academic work with limited readership—but this is hardly guaranteed, after publication. And as Gregory Whitfield (2019 *passim*, at 534) astutely observes regarding experiments, the biomedical focus of human-subjects research guidance leads political scientists to underestimate scholarly risks, particularly “diffuse and group wrongs” including “violations of legitimate authority.” Even on the Pentagon Papers, which the Supreme Court ultimately ruled against preventing newspapers from printing, the Nixon Administration presented legitimate security concerns (Rudenstine 1998, 8–9, 327–9).

Moreover, scholars are unable to determine what constitutes national security harm (even security clearances do not convey that; it rests with an Original Classification Authority), especially because research necessarily involves compilation. As we cross several documents’ wires, insightful sparks emerge—which can produce unintended consequences. Just because data are already available, scholars should not assume that prospects for additional harm have disappeared. By default, protecting national security by forgoing unauthorized information might simply be the right choice. Researchers who disagree, rather than arguing that scholarship is harmless, might consider weighing public-interest benefits against legal or security externalities.

## PUBLIC INTEREST

Norms and laws sometimes conflict: limited leaks to protect human rights, correct abuses of power and secrecy, and serve the public interest may be justifiable.

One might proudly risk prosecution and imprisonment, and engage with flawed intermediaries like WikiLeaks, to illuminate egregious wrongs. This reflects common public conceptions of what whistleblowers do and foundations of investigative journalism. However, the US government distinguishes between whistleblowers, whose disclosures are protected to combat waste, fraud, and abuse (Shimabukuro and Whitaker 2012, 15–7), and leakers of classified material, who criminally harm national security. Debating which concept covers particular disclosures is reasonable, but we should not collapse the distinction altogether. If public benefits from particular projects drive evidentiary choices and these outweigh normative concerns about law, security, or privacy, then scholars should make this case openly.

Some might argue that uncovering government secrets is intrinsically a normative good, whereby anything that takes the national security state down a peg should be encouraged. Governments' insistence on secrecy for themselves sounds hypocritical if they routinely compromise their citizens' privacy and information security. Disclosures from this perspective are inevitable and desirable, and whatever emerges should become public domain, covered by academic freedom. In a 1980s hacker mantra, "information wants to be free" (Brand 1987, 204). If researchers' value beliefs categorically legitimize leaked material, they should disclose this vital point of reflexivity.

For many analysts, though, leaks' context and intent matter. Even if scholars perceive some leakers as whistleblowers, that does not necessarily legitimize sources like the Cablegate tranche that are massive rather than targeted and not designed to expose particular governmental abuses—especially if documents are then used for research that is further removed from uncovering wrongdoing or achieving concrete public benefits. This is different from the Pentagon Papers case (Ellsberg 2002, 289–95) and perhaps Manning's leak of the "war diaries." (A more acute debate concerns leaks by former National Security Agency contractor Edward Snowden, which apparently involved more highly classified sources, with declared public interest in revealing US domestic surveillance overreach but ultimately benefiting American adversaries including Russia and China.)

Likewise, not all research projects realistically contribute enough to the public good to outweigh ethical concerns about using leaked material. Some benefits from analyzing particular data accrue privately to researchers such as information, validation of valued arguments, and resulting professional publications. For the public, though, most research projects make incremental contributions to general knowledge. Generic claims of academic freedom do not sufficiently establish ethical foundations for research with leaked documents, which ought to clear high hurdles of real-world significance. Here, broadly positivist scholars might take cues from colleagues in critical security studies on emancipation, human rights and democracy promotion, or racial justice and gender equality (inter alia, Shepherd 2013). If research has a designed purpose and benefit, that may not justify all means and methods, but

it should weigh in those judgments, especially if the normative intent of the research and the motivation and scope of the leak are aligned.

## POLICY RELEVANCE

Engagement with leaks complicates the pursuit of policy relevance. International relations scholarship has long maintained ambivalent relations with policy making (Avey and Desch 2014; Nincic and Leggold 2000), and political science with American government (Smith 1997), and leaks may exacerbate these tensions. (Relevance is not just about policy, let alone US government policy, and may even oppose it [Sjoberg 2015], so I treat public interest separately.) One might argue that using the hottest, most recent materials enhances relevance (Elias 2018, 235). And as Stephen Walt (2011, 53) suggested, WikiLeaks may have "fostered greater transparency and made the marketplace of ideas somewhat more efficient" for enhancing policy debate. However, leaks can breed further information lockdowns rather than trust and sunlight. Scholars might have difficulty contributing to policy discussion after participating in, or endorsing, security breaches. It is hard to speak truth to power when shut out of certain information, but deliberation also suffers when officials with security clearances cannot listen or respond because of how researchers' information was acquired.

However, governments often encourage research with documents wrested from their adversaries. National Archives collections contain captured state records, including more than 70,000 rolls of microfilm reproducing documents from Nazi Germany and other Axis powers (National Archives 2016). Similarly, the Iraqi Ba'ath Party files now at the Hoover Institution and the Al-Qaeda and Iraqi files that were available at the now-shuttered Conflict Records Research Center (CRRC, 2005–2010, at the National Defense University at Ft. McNair), were also products of conquest rather than consent (ABC7 2008; Cox 2010; Gordon 2015). Scholars are privy to troves of Japanese and German official communications from World War II and Soviet records from the Cold War that passed through the hands of American and British intelligence agencies (Andrew and Mitrokhin 1999; Steil 2013).

In this time of declared great power competition, Washington might well be interested in analyses of Chinese regime security decision making and domestic propaganda operations based on leaked or hacked documents (King, Pan, and Peters 2017; Nathan 2019). The US Government might encourage or even sponsor such research—but from an ethics standpoint, that can hardly be the end of the conversation, as original governmental authors (or their successors) might see these seizures as illegitimate. Reciprocally, leaked US documents might enable policy-relevant research from the perspective of both adversaries and allies, for instance in European journals. Relevance involves trade-offs, and decisions to exploit or forgo certain leaked files build norms about future disclosures' legitimacy.

## HUMAN SUBJECTS PROTECTION

Even robust public-interest or policy-relevant assertions for document access confront human-subjects protection rationales for keeping certain information private. Leaked documents violate basic expectations of confidentiality and consent. Government officials know that much of their work ultimately joins the public record but communicate freely and effectively in classified or restricted channels knowing that decades will pass before mandatory declassification and that earlier FOIA requests would be carefully vetted, perhaps with redactions. (Even declassified archival documents containing personal information are often off-limits to researchers.) Diplomatic cables report on bilateral relationships, including frank conversations with host-government leaders and societal informants such as dissidents or members of vulnerable groups, all of whom anticipate confidentiality. Many US cables explicitly mark named interlocutors as “strictly protect” (a handling instruction separate from classification) because of harm that disclosure could cause them. Intelligence files have further issues if information about individuals was itself gathered non-consensually (see Subotić 2021). Documents taken from governments do not only affect the state or amorphous, collective national security; they can also damage real people.

Boustead and Herr (2020) suggest that authors handling documentary sources apply the Belmont principles (US Department of Health 1979) that undergird human-subjects research (HSR): respect for persons, beneficence, and justice. For instance, withholding personally identifiable information by anonymizing research participants guards against risks they might face after publication. This parallels many ethnographic researchers’ resistance to inflexible transparency mandates regarding sensitive interview transcripts and field notes (Büthe and Jacobs 2015). If researchers employ leaked sources, these recommendations are valuable for harm reduction. But as Subotić (2021, 343, 348–9) argues, archival researchers’ ethical obligations to other individuals’ “dignity, humanity, and voice” transcend the Belmont guidelines and should extend to research staff, deceased persons (especially victims of violence or those whose records reflect coercion), and surviving communities.

Focusing on HSR procedural safeguards is misleading. First, we may overestimate harm-reduction capacity. Anonymization is not foolproof: vast human-subjects datasets risk “reidentification” and loss of confidentiality as researchers connect the dots (Zimmer 2010). Anonymizing published quotations offers little protection if underlying documents are hyperlinked or discoverable. Second, HSR’s biomedical and experimental assumptions do not capture the range of ethical concerns with leaked documents, including institutional harms (see Whitfield 2019). Third, the Belmont “beneficence” principle suggests higher standards than merely to do no harm: even for minimal-risk projects, researchers should consider whether they actually provide public benefits,

especially to populations being studied. Fourth, the HSR lens might downplay the disciplinary community’s oversight responsibilities, burden-shift onto IRBs, and emphasize momentary compliance and approval instead of researchers’ ongoing ethical practice (see Fujii 2012; Israel 2015; Marzano 2012, 80–1).

Finally, prioritizing safeguards during research distracts from the underlying problem of source legitimacy. Scholars must explain why consensual methods and data were insufficient and defend the public value of the research against its potential harms (see Marzano 2012, 81–2, 87). HSR prioritizes protecting vulnerable populations, and governments are hardly weak, so power relations affect research legitimacy, and it matters whose information is being seized (82–3, 89). However, Belmont principles should cover the people entangled with governmental and corporate documents and apply, as scholarly values, across research domains. Analyzing sources taken against the will of their authors (and other participants and reported-on third parties), makes researchers complicit in violations after the fact. Even if scholars avoid new individual harms, reliance on unauthorized materials undermines the ethical principles on which human-subjects protections depend.

## REFLEXIVITY

The possibility of value-neutral social science is an eternal debate, but unauthorized sources make it especially acute. Borrowing from Robert Cox (1981, 128), sources and methods may not be “always *for* someone and *for* some purpose,” but decisions to analyze or forgo leaked documents certainly are. There is no easy way to avoid complicity: on Cablegate, scholars’ methodological choices tacitly either support US Government efforts to keep secrets or endorse other actors’ attempts to reveal these. Scholars might not want to associate themselves broadly with the actions and objectives of any government or of nonstate hackers, leakers, and brokers like WikiLeaks. However, citation choices are weighty: sources and analyses compromise some interests and benefit others. There is moral hazard too: the values we prioritize affect scholarly norms and incentivize individual decisions. Repeatedly avoiding or downplaying these choices, affirms the positions taken most loudly or often. Political scientists need to send stronger, clearer signals about whether we will provide a market for future leaks and how scrupulously we take the provenance of our data. As interpretivist scholars argue, how we approach our sources says something about who we are, and who we are affects our research (Schwartz-Shea and Yanow 2012, 38–40, 66–8, 95–104). Prominent, explicit reflexivity is an important criterion for establishing research legitimacy amid the challenges of leaked documents.

## RECOMMENDATIONS

Scholars will likely rank-order these conflicting values according to their own judgments. However, I offer

four suggested guidelines to weigh criterial trade-offs, avoid the “relativistic swamp” (Gerring 2001, 26–7), and chart reliable research courses amid the challenges of unauthorized, nonconsensual, or outright stolen evidence.

First, when in doubt, leave it out. Illegality or classification may not always bar investigative scholarship, but they are significant deterrents alongside complicity, reliability, and other concerns. Unless authors persuasively justify the necessity of tainted sources, accuracy of inferences from them, and public-interest value of the findings, they should probably shy away.

Second, access requires commitment. If researchers treat leaked material as an archive like any other, they are accountable for doing it justice, which invariably requires more than a citation or two. Even small claims, backed by primary sources, require scrutinizing those documents’ original communicative context and assessing biases in source repositories and in researchers’ sampling and interpretation of particular texts, as well as ethical concerns. With the proliferation of online sources, ease of access should not imply ease of analysis.

Third, do trust journalists, but don’t touch their sources. It would be unreasonable and impractical to avoid news articles based on leaks and later work citing these as “fruit of the poisonous tree” (inadmissible evidence from illegal searches). Because professional journalists and editors use ethical filters (including the balanced primacy of public interest) to present narratives, with excerpted evidence, scholars can trust and quote the paper of record. Meanwhile, scholarly restraint from rifling the “bran tub” of leaked documents curtails complicity with existing disclosures, removes the moral hazard from future ones by forswearing research access to unauthorized information except where narrated by reputable journalists, and reduces the burden of defending our work’s risks and benefits. Conversely, circumventing journalistic filters to access leaked documents requires shouldering responsibility to vet source legitimacy and reliability, weigh public interests, and mitigate individual and institutional harms.

Fourth, disclose, defend, and debate. Transparency should be a watchword not merely for empirical purposes of fact-checking and replication, or preventing error or fraud, but because individual choices about sources and methods have ethical dimensions and the accumulation of those decisions builds community norms. Responsibility lies not just with authors using leaked or hacked sources to justify their approach but also with other scholars to explain why they avoided those documents (e.g., Poor 2017). Journals, associations, advisors, and reviewers should do more to clarify scholarly principles and standards for publishable work.

## CONCLUSION

This paper examined the problem of nonconsensually obtained evidence in political science and international

relations, focusing on the case of leaked classified documents. When sources are essentially stolen—whether through insider leaks or external hacks—and disseminated without permission, scholars face both ethical and methodological dilemmas. Researchers considering such sources must confront documents’ uncertain integrity and legality, the challenges of analyzing and the consequences of publishing their contents, and their own professional values and project objectives. This paper seeks to stimulate disciplinary debate and individual reflection on these issues. Ultimately, I argue that the hazards from this research, from national-security harms and eroding human-subjects protections to scholarly complicity with rogue actors, generally outweigh the benefits and that exceptions and justifications need to be articulated much more explicitly and forcefully than is customary in existing work.

To establish the scope of the issue, I demonstrated that published scholarly use of leaked documents is widespread and opaque. All 20 TRIP-ranked journals have published such articles, but most articles using leaked diplomatic cables do not disclose this prominently and none explicitly addresses the legality of classified material. Reliance on leaked documents is rife, but open discussion of ethical and empirical pitfalls is rare.

Researchers need to do better but also deserve clearer guidance. Thus, I argued that incomplete and inconsistent guidelines from leading political science and international relations journals and associations provide an underdeveloped foundation for assessing source legitimacy. I outlined *provenance* as the primary concept for approaching this issue more systematically, and I reviewed how disciplines from journalism to statistics to paleontology address the origins of their evidence, highlighting several reasons to refrain from analyzing nonconsensual sources. Because unilateral rule making will not resolve fundamental ethical dilemmas, I presented eight criteria to help scholars sharpen provenance debates and assess whether and how to analyze leaked documents, plus four recommendations to balance intrinsic trade-offs among those criteria.

This approach suggests intellectual humility as a unifying principle. Scholars should avoid four forms of arrogance: entitlement to sources, straightforward inference from them, confidence in public value and minimal harm, and assumption that readers share our values and need no persuasion or will not notice or mind our methods. Readers exploring the museum gallery of our evidence and citations should be encouraged to ask why a particular piece is hanging on the wall, where it originated, and how it was obtained and to reassess our broader collection and analytical contribution in that light. Ethical dilemmas and methodological best practices evolve swiftly, and critiques of the field should stop short of chastising individual scholars or decreeing universal rules. However, I encourage more rigorous introspection and deliberation on disciplinary practices before the next leak—and there is always another one—puts us to the test.

## DATA AVAILABILITY STATEMENT

Research documentation and data that support the findings of this study are openly available at the American Political Science Review Dataverse: <https://doi.org/10.7910/DVN/SPLDTF>.

## ACKNOWLEDGMENTS

All claims in this paper are my own and do not reflect the position of the US Navy or any government entity. (Given the subject matter, though, reflexivity is important: this analysis was catalyzed by my personal experiences completing mandatory government trainings, serving on an IRB, holding a security clearance, and learning from new organizational cultures, after disciplinary socialization at other academic institutions.) Earlier versions were presented at APSA 2018, the NPS NSA Department Research Connections series, and Stanford's Center for International Security and Cooperation (CISAC); for their stimulating critiques and comments, particular thanks go to Thomas Christensen, Anne Clunan, Covell Meyskens, Dan Moran, Andrew Moravcsik, Chad Nelson, Rachel Sigman, James J. Wirtz, and Andrew Yeo. For research assistance, I thank Aaron Pennington.

## CONFLICT OF INTEREST

The author declares no ethical issues or conflicts of interest in this research.

## ETHICAL STANDARDS

The author affirms this research did not involve human subjects.

## REFERENCES

- ABC7 News. 2008. "Saddam Papers Come to Bay Area." June 18. <http://abc7news.com/archive/6212282/>.
- Aftergood, Stephen. 2011. "CIA Declassifies Documents from World War I." *Federation of American Scientists* (blog). April 20. [https://fas.org/blogs/secretcy/2011/04/cia\\_wwi/](https://fas.org/blogs/secretcy/2011/04/cia_wwi/).
- American Historical Association. n.d. "Statement on Standards of Professional Conduct, updated 2019." *Historians.org*. Accessed November 14, 2021. <https://www.historians.org/jobs-and-professional-development/statements-standards-and-guidelines-of-the-discipline/statement-on-standards-of-professional-conduct>.
- American Political Science Association (APSA). 2020. "Principles and Guidance for Human Subjects Research." APSA. April 20. [https://www.apsanet.org/Portals/54/diversity%20and%20inclusion%20prgms/Ethics/Final\\_Principles%20with%20Guidance%20with%20intro.pdf?ver=2020-04-20-211740-153](https://www.apsanet.org/Portals/54/diversity%20and%20inclusion%20prgms/Ethics/Final_Principles%20with%20Guidance%20with%20intro.pdf?ver=2020-04-20-211740-153).
- American Political Science Review. 2016. "APSR Frequently Asked Questions." *Apsanet.org*. September 14. <http://www.apsanet.org/Portals/54/files/Publications/Journal%20Files/Extra%20Journal%20Content/APSR-FAQ-FINAL-2016.pdf?ver=2016-09-14-121756-013>.
- American Statistical Association. 2018. "Ethical Guidelines for Statistical Practice." *Amstat.org*. April 14. <https://www.amstat.org/asa/files/pdfs/EthicalGuidelines.pdf>.
- Andrew, Christopher, and Vasili Mitrokhin. 1999. *The Sword and the Shield: The Mitrokhin Archive and the Secret History of the KGB*. New York: Basic Books.
- Appadurai, Arjun. 1986. "Introduction: Commodities and the Politics of Value." In *The Social Life of Things: Commodities in Cultural Perspective*, ed. Arjun Appadurai, 3–63. New York: Cambridge University Press.
- Avey, Paul, and Michael Desch. 2014. "What Do Policymakers Want From Us? Results of a Survey of Current and Former Senior National Security Decision Makers." *International Studies Quarterly* 58 (2): 227–46.
- Bamford, James. 2016. "The Espionage Economy." *Foreign Policy* 216: 70–2.
- Banka, Andris, and Adam Quinn. 2018. "Killing Norms Softly: US Targeted Killing, Quasi-Secrecy and the Assassination Ban." *Security Studies* 27 (4): 665–70.
- Bennett, Andrew, and Jeffrey Checkel, eds. 2015. *Process Tracing: From Metaphor to Analytic Tool*. Cambridge: Cambridge University Press.
- Besterman, Tristram. 2014. "Crossing the Line: Restitution and Cultural Equity." In *Museums and Restitution: New Practices, New Approaches*, eds. Louise Tythacott and Kostas Arvanitis, 19–36. Burlington, VT: Ashgate.
- Boustead, Anne, and Trey Herr. 2020. "Analyzing the Ethical Implications of Research Using Leaked Data." *PS: Political Science and Politics* 53 (3): 505–9.
- Brand, Stewart. 1987. *The Media Lab: Inventing the Future at MIT*. New York: Viking Penguin.
- Burton, Antoinette, ed. 2005. *Archive Stories: Facts, Fictions, and the Writing of History*. Durham, NC: Duke University Press.
- Büthe, Tim, and Alan Jacobs, eds. 2015. "Symposium: Transparency in Qualitative and Multi-Method Research." *Qualitative & Multi-Method Research* 13 (1): 2–8.
- Clinton, Hillary. 2014. *Hard Choices*. New York: Simon and Schuster.
- Cohen, Michael, Henry Farrell, and Martha Finnemore. 2014. "Hypocrisy Hype: Can Washington Still Walk and Talk Differently?" *Foreign Affairs* 93 (2): 161–5.
- Colwell, Chip. 2015. "Curating Secrets: Repatriation, Knowledge Flows, and Museum Power Structures." *Current Anthropology* 56 (Supplement 12): S263–75.
- Connelly, Matthew, Raymond Hicks, Robert Jervis, Arthur Spirling, and Clara Suong. 2020. "Diplomatic Documents Data for International Relations: The Freedom of Information Archive Database." *Conflict Management and Peace Science* 38 (6): 762–81.
- Cox, Douglas. 2010. "Archives & Records in Armed Conflict: International Law and the Current Debate over Iraqi Records and Archives." *Catholic University Law Review* 59 (4): 1001–56.
- Cox, Robert. 1981. "Social Forces, States and World Orders: Beyond International Relations Theory." *Millennium: Journal of International Studies* 10 (2): 126–55.
- Darnton, Christopher. 2018. "Archives and Inference: Documentary Evidence in Case Study Research and the Debate over U.S. Entry into World War II." *International Security* 42 (3): 84–126.
- Darnton, Christopher. 2021. "Replication Data for: The Provenance Problem: Research Methods and Ethics in the Age of WikiLeaks." Harvard Dataverse. Dataset. <https://doi.org/10.7910/DVN/SPLDTF>.
- Deutschmann, Emanuel. 2016. "Between Collaboration and Disobedience: The Behavior of the Guantánamo Detainees and its Consequences." *Journal of Conflict Resolution* 60 (3): 555–82.
- Dilanian, Ken. 2019. "Under Trump, More Leaks— and More Leak Investigations." NBC News. April 8. <https://www.nbcnews.com/politics/justice-department/under-trump-more-leaks-more-leak-investigations-n992121>.
- Dorfman, Zach. 2018. "Axios Codebook." *Axios*, July 24. <https://www.axios.com/newsletters/axios-codebook-96add155-824b-4609-9dc3-5ca319fa4439.html>.
- Doyle, Megan. 2009. "Ownership by Display: Adverse Possession to Determine Ownership of Cultural Property." *George Washington International Law Review* 41 (1): 269–97.
- Drezner, Daniel. 2010. "Why WikiLeaks is Bad for Scholars." *Chronicle of Higher Education*, December 5. <https://www.chronicle.com/article/Why-WikiLeaks-Is-Bad-for/125628>.

- Elias, Barbara. 2018. "The Big Problem of Small Allies: New Data and Theory on Defiant Local Counterinsurgency Partners in Afghanistan and Iraq." *Security Studies* 27 (2): 233–62.
- Elbein, Asher. 2021. "Decolonizing the Hunt for Dinosaurs and Other Fossils." *New York Times*, March 22. <https://www.nytimes.com/2021/03/22/science/dinosaurs-fossils-colonialism.html>.
- Ellsberg, Daniel. 2002. *Secrets: A Memoir of Vietnam and the Pentagon Papers*. New York: Viking.
- Feaver, Peter, Allison Stanger, and Michael Walzer. 2018. "The Secret Sharers: Leaking and Whistle-Blowing in the Trump Era." *Foreign Affairs* 97 (6): 199–206.
- Foreman, Gene. 2016. *The Ethical Journalist: Making Responsible Decisions in the Digital Age*, 2nd ed. Malden, MA: Wiley-Blackwell.
- Fraser, John. 2018. "The Elephant in the Room." *American Alliance of Museums* (blog). February 12. <https://www.aam-us.org/2018/02/12/the-elephant-in-the-room/>.
- Frisch, Scott, Douglas Harris, Sean Kelly, and David Parker, eds. 2012. *Doing Archival Research in Political Science*. Amherst, NY: Cambria Press.
- Fujii, Lee Ann. 2012. "Research Ethics 101: Dilemmas and Responsibilities." *PS: Political Science and Politics* 45 (4): 717–23.
- Galbraith, Peter. 2011. "How to Write a Cable." *Foreign Policy* 185: 102–3.
- Gates, Robert. 2014. *Duty: Memoirs of a Secretary at War*. New York: Knopf.
- Giles, Micheal, and James Garand. 2007. "Ranking Political Science Journals: Reputational and Citational Approaches." *PS: Political Science and Politics* 40 (4): 741–51.
- George, Alexander, and Andrew Bennett. 2005. *Case Studies and Theory Development in the Social Sciences*. Cambridge, MA: MIT Press.
- Gerring, John. 2001. *Social Science Methodology: A Criterial Framework*. New York: Cambridge University Press.
- Gill, David, and Christopher Chippindale. 2007. "Review Article: The Illicit Antiquities Scandal: What It Has Done to Classical Archaeology Collections." *American Journal of Archaeology* 111 (3): 571–4.
- Gill, Michael, and Arthur Spirling. 2015. "Estimating the Severity of the WikiLeaks U.S. Diplomatic Cables Disclosure." *Political Analysis* 23 (2): 299–305.
- Glasser, Susan. 2011. "Epiphanies: Bob Woodward." *Foreign Policy* 188: 25.
- Goodison, Sean, Robert Davis, and Brian Jackson. 2015. "Digital Evidence and the U.S. Criminal Justice System: Identifying Technology and Other Needs to More Effectively Acquire and Utilize Digital Evidence." Report, RAND Corporation. [https://www.rand.org/content/dam/rand/pubs/research\\_reports/RR800/RR890/RAND\\_RR890.pdf](https://www.rand.org/content/dam/rand/pubs/research_reports/RR800/RR890/RAND_RR890.pdf).
- Gordon, Michael. 2015. "Archive of Captured Enemy Documents Closes." *New York Times*, June 21. <https://www.nytimes.com/2015/06/22/world/middleeast/archive-of-captured-terrorist-qaeda-hussein-documents-shuts-down.html>.
- Haberman, Maggie, and Katie Rogers. 2018. "An Aggrieved Trump Wants Better Press, and He Blames Leaks for Not Getting It." *New York Times*, May 17. <https://www.nytimes.com/2018/05/17/us/politics/white-house-leaks.html>.
- Healy, Patrick, David Sanger, and Maggie Haberman. 2016. "Donald Trump Finds Improbable Ally in WikiLeaks." *New York Times*, October 12. <https://www.nytimes.com/2016/10/13/us/politics/wikileaks-hillary-clinton-emails.html>.
- Helderman, Rosalind, Josh Dawsey, and Beth Reinhard. 2021. "Trump Grants Clemency to 143 People in Late-Night Pardon Blast." *Washington Post*, January 20. [https://www.washingtonpost.com/politics/trump-pardons/2021/01/20/7653bd12-59a2-11eb-8bcf-3877871c819d\\_story.html](https://www.washingtonpost.com/politics/trump-pardons/2021/01/20/7653bd12-59a2-11eb-8bcf-3877871c819d_story.html).
- Higonnet, Anne. 2013. "Afterword: The Social Life of Provenance." In *Provenance: An Alternate History of Art*, eds. Gail Feigenbaum and Inge Reist, 195–209. Los Angeles, CA: Getty Research Institute.
- Horner, David. 2014. *Understanding Media Ethics*. Thousand Oaks, CA: SAGE.
- Hunt, Katie. 2020. "'Blood Amber' May Be A Portal Into Dinosaur Times, But The Fossils Are An Ethical Minefield For Palaeontologists." CNN, updated September 20, 2020. <https://www.cnn.com/2020/09/19/world/blood-amber-myanmar-fossils-sc/index.html>.
- Information Security Oversight Office. 2018. *Developing and Using Security Classification Guides*. National Archives. October. <https://www.archives.gov/files/isoo/training/scg-handbook.pdf>.
- International Studies Association. n.d. "Statement on the Use of Classified Materials in ISA Publications." Accessed November 14, 2021. <https://www.isanet.org/Publications/Classified-Materials>.
- Israel, Mark. 2015. *Research Ethics and Integrity for Social Scientists: Beyond Regulatory Compliance*, 2nd ed. Thousand Oaks, CA: SAGE.
- Jacobs, Alan, Tim Büthe, Ana Arjona, Leonardo R. Arriola, Eva Bellin, Andrew Bennett, Lisa Björkman, et al. 2021. "The Qualitative Transparency Deliberations: Insights and Implications." *Perspectives on Politics* 19 (1): 171–208.
- Jervis, Robert. 2015. "The Torture Blame Game: The Botched Senate Report on the CIA's Misdeeds." *Foreign Affairs* 94 (3): 120–7.
- Jones, Cynthia. 2010. "A Reason to Doubt: The Suppression of Evidence and the Inference of Innocence." *The Journal of Criminal Law and Criminology* 100 (2): 415–74.
- Kapiszewski, Diana, Lauren MacLean, and Benjamin Read. 2015. *Field Research in Political Science: Practices and Principles*. Cambridge: Cambridge University Press.
- Kapiszewski, Diana, and Elisabeth Jean Wood. 2021. "Ethics, Epistemology, and Openness in Research with Human Participants." *Perspectives on Politics* doi:10.1017/S1537592720004703.
- Keller, Bill. 2011. "Dealing with Assange and the WikiLeaks Secrets." *New York Times*, January 26. <https://archive.nytimes.com/www.nytimes.com/2011/01/30/magazine/30Wikileaks-t.html>.
- King, Gary, Jennifer Pan, and Margaret Peters. 2017. "How the Chinese Government Fabricates Social Media Posts for Strategic Distraction, Not Engaged Argument." *American Political Science Review* 111 (3): 484–501.
- Kornblut, Anne. 2008. "Obama Central: Peace, Harmony and Deep Secrecy." *Washington Post*, August 3. [https://www.washingtonpost.com/wp-dyn/content/article/2008/08/02/AR2008080201687.html?itid=ik\\_inline\\_manual\\_1](https://www.washingtonpost.com/wp-dyn/content/article/2008/08/02/AR2008080201687.html?itid=ik_inline_manual_1).
- Kuyper, Jonathan. 2016. "Systemic Representation: Democracy, Deliberation, and Nonelectoral Representatives." *American Political Science Review* 110 (2): 308–24.
- La Follette, Laetitia. 2017. "Looted Antiquities, Art Museums and Restitution in the United States since 1970." *Journal of Contemporary History* 52 (3): 669–87.
- Lessing, Benjamin, and Graham Denyer Willis. 2019. "Legitimacy in Criminal Governance: Managing a Drug Empire from behind Bars." *American Political Science Review* 113 (2): 584–606.
- Lew, Jacob. 2010. "WikiLeaks-Mishandling of Classified Information." *Obamawhitehouse.archives.gov*. November 28. <https://obamawhitehouse.archives.gov/sites/default/files/omb/memoranda/2011/m11-06.pdf>.
- Maliniak, Daniel, Susan Peterson, and Michael Tierney. 2012. "TRIP around the World: Teaching, Research, and Policy Views of International Relations Faculty in 20 Countries." Williamsburg, VA: The College of William and Mary. [https://www.wm.edu/offices/itpir/\\_documents/trip/trip\\_around\\_the\\_world\\_2011.pdf](https://www.wm.edu/offices/itpir/_documents/trip/trip_around_the_world_2011.pdf).
- Marzano, Marco. 2012. "Ethics and Social Conflict: A Framework for Research." In *Ethics in Social Research*, ed. Kevin Love, 73–90. Bingley, UK: Emerald Publishing.
- McBride, Kelly, and Tom Rosenstiel. 2014. "New Guiding Principles for a New Era of Journalism." In *The New Ethics of Journalism: Principles for the 21st Century*, eds. Kelly McBride and Tom Rosenstiel, 1–6. Los Angeles, CA: CQ Press.
- Michael, Gabriel. 2015. "Who's Afraid of WikiLeaks? Missed Opportunities in Political Science Research." *Review of Policy Research* 32 (2): 175–99.
- Monks-Leeson, Emily. 2011. "Archives on the Internet: Representing Contexts and Provenance from Repository to Website." *The American Archivist* 74 (1): 38–57.
- Moravcsik, Andrew. 2014. "Trust, but Verify: The Transparency Revolution and Qualitative International Relations." *Security Studies* 23 (4): 663–88.
- Nathan, Andrew. 2019. "The New Tiananmen Papers: Inside the Secret Meeting That Changed China." *Foreign Affairs* 98 (4): 80–91.



- National Archives. 2009. "The President Executive Order 13526, Classified National Security Information." December 29. <https://www.archives.gov/isoo/policy-documents/cnsi-EO.html>.
- National Archives. 2011. "Press Release: National Archives and Presidential Libraries Release Pentagon Papers." June 8. <https://www.archives.gov/press/press-releases/2011/nr11-138.html>.
- National Archives. 2016. "Collection of Foreign Records Seized." Last reviewed August 15, 2016. <https://www.archives.gov/research/captured-german-records/foreign-records-seized.html>.
- National Archives. 2019. "About Controlled Unclassified Information." Last reviewed August 1, 2019. <https://www.archives.gov/cui/about>.
- Nuwer, Rachel. 2019. "This Tarantula Became a Scientific Celebrity. Was It Poached from the Wild?" *New York Times*, April 1. <https://www.nytimes.com/2019/04/01/science/poaching-wildlife-scientists.html>.
- Nincic, Miroslav, and Joseph Lepgold, eds. 2000. *Being Useful: Policy Relevance and International Relations Theory*. Ann Arbor: University of Michigan Press.
- Parker, Ashley, and David Sanger. 2016. "Donald Trump Calls on Russia to Find Hillary Clinton's Missing Emails." *New York Times*, July 27. <https://www.nytimes.com/2016/07/28/us/politics/donald-trump-russia-clinton-emails.html>.
- O'Loughlin, John. 2016. "The Perils of Self-Censorship in Academic Research in a WikiLeaks World." *Journal of Global Security Studies* 1 (4): 337–45.
- Poor, Nathaniel. 2017. "The Ethics of Using Hacked Data: Patreon's Data Hack and Academic Data Standards." In *Internet Research Ethics for the Social Age: New Challenges, Cases, and Contexts*, eds. Michael Zimmer and Katharina Kinder-Kurlanda, 277–80. New York: Peter Lang.
- Rinke, Eike Mark, and Alexander Wuttke. 2021. "Open Minds, Open Methods: Transparency and Inclusion in Pursuit of Better Scholarship." *PS: Political Science and Politics* 54 (2): 281–4.
- Rosenberg, Matthew, and Eric Schmitt. 2017. "Trump Revealed Highly Classified Intelligence to Russia, in Break with Ally, Officials Say." *New York Times*, May 15. <https://www.nytimes.com/2017/05/15/us/politics/trump-russia-classified-information-isis.html>.
- Rudenshteyn, David. 1998. *The Day the Presses Stopped: A History of the Pentagon Papers Case*. Oakland: University of California Press.
- Rudenshteyn, David. 2002. "Lord Elgin and the Ottomans: The Question of Permission." *Cardozo Law Review* 23: 449–71.
- Rudenshteyn, David. 2016. *The Age of Deference: The Supreme Court, National Security, and the Constitutional Order*. Oxford: Oxford University Press.
- Rusbridger, Alan. 2018. *Breaking News: The Remaking of Journalism and Why It Matters Now*. New York: Farrar, Straus and Giroux.
- Savage, Charlie. 2017. "Chelsea Manning to Be Released Early as Obama Commutes Sentence." *New York Times*, January 17. <https://www.nytimes.com/2017/01/17/us/politics/obama-commutes-bulk-of-chelsea-mannings-sentence.html>.
- Savage, Charlie. 2021. "Risk of Nuclear War over Taiwan in 1958 Said to Be Greater than Publicly Known." *New York Times*, May 22. <https://www.nytimes.com/2021/05/22/us/politics/nuclear-war-risk-1958-us-china.html>.
- Schwartz-Shea, Peregrine, and Dvora Yanow. 2012. *Interpretive Research Design: Concepts and Processes*. New York: Routledge.
- Scott, Janny. 2021. "Now It Can Be Told: How Neil Sheehan Got the Pentagon Papers." *New York Times*, January 7; updated January 9. <https://www.nytimes.com/2021/01/07/us/pentagon-papers-neil-sheehan.html>.
- Seib, Philip. 2006. *Beyond the Front Lines: How the News Media Cover a World Shaped by War*. New York: Palgrave.
- Shimabukuro, Jon, and L. Paige Whitaker. 2012. "Whistleblower Protections under Federal Law: An Overview." Congressional Research Service. September 13. <https://fas.org/sgp/crs/misc/R42727.pdf>.
- Shapiro, Jacob, and David Siegel. 2010. "Is this Paper Dangerous? Balancing Secrecy and Openness in Counterterrorism." *Security Studies* 19 (1): 66–98.
- Shepherd, Laura, ed. 2013. *Critical Approaches to Security: An Introduction to Theories and Methods*. New York: Routledge.
- Sjoberg, Laura. 2015. "Locating Relevance in Security Studies." *Perspectives on Politics* 13 (2): 396–8.
- Smith, Rogers. 1997. "Still Blowing in the Wind: The American Quest for a Democratic, Scientific Political Science." *Daedalus* 126 (1): 253–87.
- Society of Vertebrate Paleontology. 2020. "Fossils from Conflict Zones and Reproducibility of Fossil-Based Scientific Data." Letter to Editors. <https://vertpaleo.org/wp-content/uploads/2021/01/SVP-Letter-to-Editors-FINAL.pdf>.
- Steil, Benn. 2013. "Red White: Why a Founding Father of Postwar Capitalism Spied for the Soviets." *Foreign Affairs* 92 (2): 115–29.
- Subotić, Jelena. 2021. "Ethics of Archival Research on Political Violence." *Journal of Peace Research* 58 (3): 342–54.
- Sweeney, Shelley. 2008. "The Ambiguous Origins of the Archival Principle of 'Provenance.'" *Libraries & the Cultural Record* 43 (2): 193–213.
- The Editors of American Political Science Review. 2017. "Notes from the Editors." *American Political Science Review* 111 (3): iii–ix.
- Tidy, Joanna. 2017. "Visual Regimes and the Politics of War Experience: Rewriting War 'from above' in WikiLeaks' Collateral Murder." *Review of International Studies* 43 (1): 95–111.
- Trachtenberg, Marc. 2006. *The Craft of International History: A Guide to Method*. Princeton, NJ: Princeton University Press.
- Trouillot, Michel-Rolph. 1995. *Silencing the Past: Power and the Production of History*. Boston: Beacon Press.
- UK Parliament. 2009. "Holocaust (Return of Cultural Objects) Bill" (second reading). Hansard, Lords, Vol. 712, July 10, 2009. [https://hansard.parliament.uk/Lords/2009-07-10/debates/09071034000391/Holocaust\(ReturnOfCulturalObjects\)Bill](https://hansard.parliament.uk/Lords/2009-07-10/debates/09071034000391/Holocaust(ReturnOfCulturalObjects)Bill).
- U.S. Department of Health, Education, and Welfare. 1979. "The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research." April 18. [https://www.hhs.gov/ohrp/sites/default/files/the-belmont-report-508c\\_FINAL.pdf](https://www.hhs.gov/ohrp/sites/default/files/the-belmont-report-508c_FINAL.pdf).
- Walt, Stephen. 2011. "Where Do Bad Ideas Come from and Why Don't They Go Away?" *Foreign Policy* 184: 48–53.
- Walzer, Michael. 2018. "Just and Unjust Leaks: When to Spill Secrets." *Foreign Affairs* 97 (2): 48–59.
- Ward, Stephen. 2014. "The Magical Concept of Transparency." In *Ethics for Digital Journalists: Emerging Best Practices*, eds. Lawrie Zion and David Craig, 45–58. New York: Routledge.
- Whitfield, Gregory. 2019. "TRENDS: Toward a Separate Ethics of Political Field Experiments." *Political Research Quarterly* 72 (3): 527–38.
- World Politics. 2018 "Style Sheet for World Politics." <https://www.cambridge.org/core/services/aop-file-manager/file/587f9e584dc12a014a4fb391/World-Politics-Style-Sheet.pdf>.
- Zimmer, Michael. 2010. "'But the Data Is Already Public': On the Ethics of Research in Facebook." *Ethics & Information Technology* 12 (4): 313–25.