Investigative Approaches to Information Technology Research

Daniel Carter, School of Journalism and Mass Communication, Texas State University

Amelia Acker, School of Information, University of Texas at Austin

Dan Sholler, Technology Management, University of California, Santa Barbara

Abstract

Recent events make clear that information technology companies often consciously design products that are societally harmful. They also make clear that these companies are unlikely to disclose their practices and often go to great lengths to keep them hidden. At the same time, however, these companies are responsible for wide-reaching, influential phenomena that should be of core interest to scholars of information technology. While scholars' methodologies and ethical conventions have tended to privilege the agency of research participants and focused on the prevention of harm, we argue that the current landscape calls for a reevaluation of these conventions. Specifically, we consider here what it would mean for researchers to take a more investigative approach by seeking to discover and reveal information that the subjects of the research would prefer to remain unknown. In reviewing literature on investigative journalism as a potential source of inspiration, we attempt to start a discussion about how our research can be more impactful, as well as the challenges that would accompany such a change.

Introduction

Shoshanna Zuboff's *The Age of Surveillance Capitalism* (2019) details a range of societally harmful data collection practices taking place in information technology industries. In writing this article we are motivated by the ongoing prevalence of these processes and others, ranging from threats to democracies and destabilization of knowledge institutions through disinformation and platform manipulation campaigns, to corporate fraud and small-scale grifts. But we are even more motivated by a quieter moment in Zuboff's narrative that highlights the distinct challenges facing researchers working, in the present, on topics related to information technology. Reflecting on Peter Drucker's extensive field work at GM and projecting his methods into the present organizational context at information technology firms, Zuboff notes:

Google is a notoriously secretive company, and one is hard-pressed to imagine a Drucker equivalent freely roaming the scene and scribbling in the hallways. Its executives carefully craft their messages of digital evangelism in books and blog posts, but its operations are not easily accessible to outside researchers or journalists. (p. 66)

Conventional approaches to scholarly research based on cooperation and consent are prohibited in these high-tech corporate environments, and academic researchers largely have no direct

access to the kinds of societally harmful processes that Zuboff describes. Indeed, many common practices in the technology industry allow companies to shape public knowledge in deceptive ways, harming academics', journalists' and others' abilities to understand or comment on unethical behavior. For example, Facebook has prohibited investigative and critical research of ad targeting by disabling the functionality of browser plugins and web scraping techniques that collect information from the platform, in favor of alternative data access regimes such as curated collections like the Facebook Ad Library. Bruns' (2019) recently evaluated social platforms and the implications of increasing API rollbacks and data grants (what he calls the 'APIcalypse') on independent and public interest scholarship. While corporate data philanthropy access schemes such as Facebook's partnership with Social Science One, the Ad Library, or the failed Twitter gift archive to the Library of Congress all appear to support access for researchers, severe limitations on access, terms of use, and search functionality are crippling the ability to investigate technology companies (Alaimo 2018; Mozilla 2019b; 2019a). Bruns argues that such corporate data philanthropy is "deeply flawed" for researchers while working as fig leaf to demonstrate corporate social responsibility that "generate[s] substantial positive publicity for the platforms" (p.11). Further, technology companies routinely release information to journalists "on background"—meaning, in these cases, that attribution of the information will be negotiated.

[This tactic] allows giants like Amazon and Tesla an opportunity to transmit their preferred message, free of risk, in the voice of a given publication. It leaves no trace of policy that might later be criticized—that could form part of the public record to be scrutinized by regulators, lawyers, or investors. If the company later reverses course or modifies its position, the egg is on the reporter's face, not the company's. (Merchant 2019)

However, while journalists and academics have limited direct access from which to counter these public relations tactics, investigative journalists—due to their unique approach and methods—have been instrumental in uncovering and publicizing phenomena that might otherwise remain hidden. For example, *ProPublica* recently published a series of articles looking at the tactics used by Intuit, the company behind the popular TurboTax tax filing software, to prevent people from discovering options to file their taxes for free and to instead route them to non-free alternatives. The articles focus on uncovering dark patterns, or techniques used in interface design and search engine optimization to route users away from a desired outcome or to otherwise confuse or hamper their actions in ways that are profitable. *ProPublica*'s articles received considerable attention, and the U.S. Senate subsequently modified a bill that would have helped to legitimize Intuit's harmful practices (Elliott 2019).

Dark patterns implemented by the developers of information technology are also of interest to academic researchers in fields such as human-computer interaction (e.g., C. M. Gray et al. 2018). However, the methods used by academic researchers who have studied dark patterns—as well as

their general attitude toward deceptive companies—differ considerably from those of investigative journalists. Where academic researchers, largely, describe and theorize concepts, journalists have been far more concerned with understanding the organizational decisions around dark patterns, the individuals responsible for their implementation and the specific ways they harm individuals. Both researchers and journalists hold values of accountability and the circulation of knowledge as motivations for their work, but each group has distinct methods and limits of inquiry, and each conforms to different pressures related to accountability and career advancement. A central part of the social and cognitive organization of a profession is the ability to shape events in the domain of its scrutiny into phenomena of importance, because these topics of interest make up the discourse that the profession is organized around as well as the way it trains its new members. Yet the unethical behaviors perpetrated by information technology firms are increasingly beyond the limits of scholars' methods as compared to our colleagues in journalism and civil society organizations.

We are concerned by this discrepancy of methods and attitudes and by what it means for the work of scholars who study information technology. We believe the processes that have developed around information technology warrant research approaches that account not just for their potential to do harm but also for the specific organizational decisions behind that harm and the possibility that those responsible will not be cooperative research participants. Reflecting on our own training and research practices, we began this project by questioning ourselves: how are we supposed to learn about processes that are hugely consequential—both as phenomena in themselves but also as potentially valuable grounds for theory development—when we lack the tools to tell stories that people don't want to be told? How should deeply rooted conventions related to research contributions and ethics be negotiated in order to access knowledge that is heavily guarded, hidden, and sometimes obfuscated due to specious, profit-driven motivations?

Responding to these questions, we suggest that researchers of information technology would benefit from exploring what it would mean to take a more investigative approach in their work—an approach that would counter current norms by seeking to reveal information that is purposefully hidden because it evidences harm, liability, or negligence. In this article, we first review current approaches that give some understanding of societally harmful processes but, we argue, fall short of giving researchers sufficient access to important phenomena. We then review literature describing the work of investigative journalists, highlighting features of their approaches and methods that we find provocative in the context of academic research. We conclude by discussing the challenges of adopting an investigative approach to academic research.

By arguing that researchers of information technology should consider taking an investigative approach, we do not intend to suggest that current approaches are naive or that they do not include verification techniques such as triangulation. Instead, we hope to broaden researchers'

repertoire of methods and to begin to explore research orientations that would draw on scholars' substantial domain expertise, theoretical perspectives and methods while providing new kinds of access to consequential phenomena and increasing the potential to produce scholarship that would be both theoretically fruitful as well as publicly engaging. As the "factories" of information technology increasingly harbor actors who set out to profit from societally harmful actions while closing their doors to academic researchers, pursuing ways to sneak into the factory or at least peep through its windows is an ethical response that warrants consideration.

Current approaches and limitations

Studies of Technical Objects

A collection of approaches—including some associated with platform studies, interface studies and software studies—focuses on revealing the harmful or unethical nature of information technologies by analyzing the formal features of objects. While few if any examples of these approaches would be described as entirely technologically determinative, they do shift away from the strong social constructivism that often characterizes sociological work on technology. Rather than looking, primarily, at social factors influencing the development and use of technologies, these approaches shift the focus to ask how features might, for example, reflect social conditions or structure use. As Bogost and Montfort (2009) argue in relation to platform studies, the brand of soft technological determinism they promote offers, in relation to approaches such as the social construction of technology, an awareness that the black box of technology is not empty—that it can in fact be opened—and that the specific features found within it can be revelatory. While the specific methods employed by researchers working from these broad approaches vary, including the close analysis of hardware (e.g., Dourish 2014; Montfort and Bogost 2009), code (e.g., Mackenzie 2005), interface (e.g., Andersen and Pold 2011; Bertelsen and Pold 2004) and platform (e.g., Gerlitz and Helmond 2013), we discuss them together here because they offer similar potentials to reveal and critique unethical patterns employed in information technology design. At the same time, they are broadly limited by their lack of access to the social processes and individual decisions behind technical features as well as the consequences these have for actual users.

Gray et al.'s (2018) study of dark patterns, for example, analyzes features such as the form, affordances and constraints of user interfaces. In their work, Gray et al. collect examples of dark patterns through online searches and then develop a classification system based on an iterative coding process. Their analysis is based on formal features such as the persistence of prompts or the obfuscation of important information. Similarly, in analyzing the Facebook Like button as a telling and visible feature of that social media platform, Gerlitz and Helmond (2013) argue that Facebook pursues what they term a "like economy" at the expense of the sociality the platform claims to encourage. While their analysis makes claims about Facebook's business practices and strategic rhetoric, their methodology is to extrapolate from what is accessible—largely, the

technical features of Facebook's like buttons and associated systems—rather than to pursue what we assume would not be made visible to them: the company's motivations and business strategies, as seen through direct empirical methods.

As a critical approach, attending to formal and technical features of objects is a direct way of addressing the unethical information technologies that are produced now. In relation to somewhat related approaches such as critical design (e.g., Bardzell and Bardzell 2013; Dunne 2005), critiquing and analyzing specific objects does provide a way to speak directly about harms that are actually being perpetrated. However, as Pasquale (2015) argues, opening the black box to reveal what is inside—even for the purpose of critique—can be an ineffective measure when it prompts increased complexity and obfuscation by businesses (p. 8). Further, such studies are most often unable to address the organizational processes behind unethical behavior, a limitation to both the development of theory and the viability of the work as a piece of public scholarship.

We are not suggesting that software studies or any of these other areas are not engaged with social processes or how the formal features they privilege came to be. Indeed, the work of researchers such as Dourish (2014) and MacKenzie (2006) is often strongly rooted in empirical stories that engage with historical or journalistic sources. Instead, we suggest that such approaches might be complemented by the methods and attitudes associated with investigation or that a focus on investigation might draw attention and provide a framework for discussing aspects of these studies that are otherwise difficult to place in the context of conventional methodologies.

Policy Studies and Proposals

One approach to culling the types of dark data practices discussed above is to implement governmental or industry-wide policies aimed at achieving fairness, justice, and other positive outcomes. Accordingly, scholars have conducted research that evaluates such policies or develops new policy strategies for ethical data collection and use. These types of scholarship tend to take two forms: (1) Studies and/or critiques of cases involving dark data practices; and (2) Construction of frameworks for developing and implementing data policies, such as those based on human rights frameworks (e.g., Human Rights Council, 2018). Both types are commonly found in technology law, information science, and political science outlets and often rely on journalistic investigations of unethical information technology design and uses.

Scholars of technology, policy, and law often adopt a version of the approach described above to identify ethical flaws in built systems and/or provide technical solutions. Others, however, take a case study approach in which they identify particular instances of unethical design and uses of information technology and relate them to the successes and shortcomings of governmental or organizational policies. For example, scholars have examined and critiqued the use of recidivism

prediction instruments and other data-centric tools in criminal justice (Barabas et al., 2018; Chouldecova, 2017; Kleinberg et al., 2018); technology-enabled consumer discrimination in the “gig” economy (Edelman and Luca, 2014); and information technology’s role in reinforcing and perpetuating poverty (Eubanks, 2018).

While case studies focus on solutions applicable to specific, local contexts, the production of frameworks attempts to provide macro-level solutions. For example, Mulligan and Griffin’s (2018) analysis of a search engine controversy, published in the *Georgetown Law Technology Review*, provides a foundational example of policy frameworking. The authors explore a case in which Google’s search engine returned seemingly-factual information from the query “did the holocaust happen.” The analysis attempts to delineate the roles and accountabilities of “imaginaries” who script search tools and describes the clash between engineering logics, social construction of historical truths, and stakeholder understandings of the First Amendment. The paper emphasizes the consequences of unethical information technology design—e.g., the spread of misinformation— and offers solutions informed by human rights law (particularly, the “right to truth” as an individual and collective right) that align with the realities of doing business. Other studies have similarly developed human rights-based analyses of unethical designs and uses of information technology, and sometimes explicitly build out existing human rights frameworks to include consideration of information technology. The rights issues and related technology policies scholars have examined include workplace discrimination (Rosenblat et al., 2017), human rights violations enabled by and examined through social media in Myanmar (Human Rights Council, 2018), and the ever-present topic of privacy.

Industry Partnerships

Some researchers may argue that access is already provided by platforms through their developers’ SDKs and APIs, services that are even, in some cases, specifically provided for academic research. For example, the Yelp Open Dataset or Google’s Cloud Public Datasets are both used for educational and research purposes. Yet these curated datasets are often partial, incomplete or stripped of context. Many social media researchers have found that the tools to access these data are “black boxes” and the developers’ terms of service for most platform APIs limit data transparency, prohibit archival access and cripple attempts at reproducibility—further, in the wake of public scandals such as Cambridge Analytica’s extensive use of Facebook data, companies have sharply curtailed access (Bruns 2019). Similarly, Freelon (2018) warns that researchers can no longer rely exclusively for access to social media data and should instead pursue less sanctioned methods such as web scraping.

One way to avoid the limitations of API access is for entrepreneurial academics to negotiate access to platform data through special research grants and competitive internal competitions such as the WhatsApp Research Awards or the Snap Research Fellowships. A related arrangement, proposed by King and Persily (2019) would see researchers apply for access to

industry data through a committee of scholars, with accepted projects funded through charitable foundations. Social Science One, an organization formed by King and Persily, currently implements this plan to allow scholars to work with data from Facebook.

However, the circulation of King and Persily's proposal and the announcement of the first round of Facebook data access granted to researchers by Social Science One coincided with Mark Zuckerberg's testimony before two Senate committees in a hearing on Data Privacy and Protection, leading some to speculate that the announcement would primarily serve as a public relations tool for the beleaguered company. While grantees have been announced, the project's website still claims that Facebook audits what researchers do with the data and that the data management is managed by Facebook ("Our Facebook Partnership" n.d.).

We now see also new industry-academic models from federal research agencies as well, such as the NSF Program on Fairness in Artificial Intelligence in Collaboration with Amazon (FAI). The program solicitation is the first of its kind to feature a corporation's logo alongside the two NSF directorates sponsoring the call for proposals ("NSF Program on Fairness in Artificial Intelligence in Collaboration with Amazon (FAI) (Nsf19571)" n.d.).

While private-public access models such as these do provide researchers with access to quality data and remove restrictions on publishing that might limit the publication opportunities of embedded researchers, the selection of scholars who review proposals and the organizations that fund projects, as well as the methods of data collection and the organization of data by companies still limits researchers' access, especially to data that would be potentially harmful to the company. For these reasons, private-public access models might still be more accurately described as corporate sponsored research.

The Investigative Approach

As a resource for learning about investigative approaches, we focus here on investigative journalism. This is certainly not the only profession that might offer ideas and guidance for researchers; we might, for example, have alternatively looked to the work of insurance or government auditors, forensic investigators or detectives. However, we find the work of investigative journalists useful for several reasons. In producing published, narrative accounts, journalists' products likely have much in common with the output of academic researchers, and, recently, journalists have conducted investigative work on information technology processes similar to those that interest researchers. Further, as we discuss below, we find the literature on investigative journalism provocative—and at this stage, provocation is, we argue, what is needed. The scholars we hope to address here exist within distinct epistemological communities and will need to debate and work out what an investigative approach looks like for them. There is no reason to think that the attitudes and methods of investigative journalists will translate

seamlessly; at the same time, we've found reading this literature a good way to initiate conversations among ourselves.

As Zelizer (2004) notes, academic researchers have often downplayed the work of journalists, with even those scholars focused on it largely limiting themselves to a narrow slice of the profession. Indeed, the majority of work on journalism falls within the realm of communication or media studies, with scholars looking more to the reception of journalists' products than to the work and knowledge practices of journalists themselves. However, Dahlgren (1992) does point to a strand of journalism research focused on the production of news from a sociological perspective, often critiquing, for example, the reliance on powerful institutional sources (e.g., Schudson 1989). While the influence of this strand of research persists, more recently, in examinations of the work of data journalists (e.g., Lewis and Westlund 2015; Parasie 2015), the practices of investigative journalists have received strikingly little attention despite the large role their output has played in the public consciousness (e.g., Bob Woodward and Carl Bernstein's work on the Watergate case and Glenn Greenwald, Laura Poitras, and Ewen MacAskill's work based on NSA documents leaked by Edward Snowden).

In the context of sociological work on investigative journalism, James Ettema and Theodore Glasser's articles and books from the 1980s and 1990s represent the bulk of what is available. Fortunately, this work is deeply interested in investigative journalism as a form of inquiry and engages with discussions of epistemology and method in a way that aligns well with our interest in investigation as a paradigm for academic research. In pursuing a "sociology of epistemology," Ettema and Glasser (1984) inquire into how journalists produce and justify knowledge claims, a project that involves drawing distinctions between the epistemological practices of investigative journalists and those engaged in daily reporting. We find these epistemological distinctions useful for thinking about researchers' practices and ways that they might pursue more investigative approaches to studying information technology processes. For these reasons, in the following sections we draw primarily on Ettema and Glasser's body of work.

Moral Judgement

Investigative journalism is inherently moral, as it focuses on exposing injuries and injustices, especially those perpetrated by individuals and institutions in power. While journalists uncover and expose specific injustices, they describe their work to Ettema and Glasser in the context of broader patterns and systemic problems. As one journalist told Ettema and Glasser:

"I look for something that is more than just a single individual who may be hurting a small group of people. [...] I look to whether or not it's an area where something needs to be done. Is there a need for regulation? Is there a need for legislation? Is there a need for stepped-up enforcement by some governmental agency that's not doing its job?"

However, Ettema and Glasser note that the obvious moral component of investigative journalism does sit in tension with the conventional view of journalists (held by the public as well as practitioners) as objective, outside observers. While journalists intend to point out injustices, they require that the judgement of what is unjust come from an explicit, external source. Established standards such as laws, ethical codes, professional standards and expert opinions serve this purpose, allowing for knowledge with obvious moral impact to be produced in a way that preserves the professional standards and public expectations of journalists.

Distrust of statements from those in power

Ettema and Glasser (1984) distinguish between daily reporters and investigative journalists primarily on the basis of their orientations to verification. Daily reporters, Ettema and Glasser argue, judge knowledge claims based on their credibility, merely passing on information from sources rather than attempting to verify the truth of the information. Consequently, daily reporters do not question information put out through bureaucratic processes such as press conferences, assuming the credibility of these sources at face value. Fishman (1980) thus argues that daily reporters “participate in upholding a normative order of authorized knowers in society” (p. 96). Rather than challenging power, everyday reporters tend to affirm it.

In contrast, Ettema and Glasser argue that investigative journalists are often prompted in their work by precisely incredible sources. Anonymous tips and hunches, for example, call for verification, but investigative journalists also operate, even when dealing with bureaucratic statements and documents, under the assumption that, while some forms of evidence carry more weight than others, no knowledge claims are, as daily journalists assume, self-evident. An investigative journalist interviewed by Ettema and Glasser (1998), for example, characterizes the difference between the work of daily reporters and his own as such: “Typical reporting gets the charge and gets the response and stops there. But what we tried to do with the racism series—what investigative reporting tries to do—is to go back and say, ‘Well, that response doesn’t make sense.’ Or, ‘That response is a lie’” (p. 10).

To some extent investigative journalists’ refusal to accept statements from people in positions of authority mirrors social scientists’ practices of triangulation and seeking out contradictory cases. However, as becomes clear below, in our discussions of interviewing techniques and institutional standards related to research ethics, journalists’ distrust of statements from those in power has distinct consequences that enable them to discuss societally harmful processes in ways that academic researchers often cannot.

Release of information against the will of those to whom it pertains

The moral orientation outlined above assumes that the subjects of investigation do not willingly reveal their harmful practices. Indeed, investigative journalism is premised on the assumption that the knowledge produced is either intentionally concealed or at least overlooked.

As Ettema and Glasser (1998) argue, publicity—defined as “the transmission of public information [as well as] the construction of the place and the occasion for the conduct of public affairs”—is a core value of investigative journalism, with the added stipulation that the way investigative journalists enact this value is often contrary to the wishes of those in power (p. 190). Here Ettema and Glasser cite the work of sociologist Alvin Gouldner (1976) in highlighting the difference between publicity as a democratic principle and the term’s use in the field of public relations. By challenging the statements of those in power and revealing information that others would choose to remain hidden, investigative journalists work against public relation professionals’ “fusion of media accounts with manager accounts of social reality” (qtd. in Ettema and Glasser, p. 190).

As we discuss below, the release of information against peoples’ desires and the consequent damage to reputations or organizations sits in stark contrast to the institutional norms and expectations of academic research.

Reliance on distinct methods

In pursuing the investigative approach outlined above, journalists rely on a range of methods that overlap with those of social scientists but also diverge in notable ways. Here, we focus on two methods that we find especially provocative in relation to the work practices and community standards of academic researchers.

Adversarial interviews

Adversarial interviews differ from those conventionally conducted as part of social science research in that the interviewer pushes back against or questions the assertions of the interviewee. Adversarial interviews are most often discussed in relation to broadcast news (e.g., Clayman and Heritage 2002), a context in which politicians and public figures can be called to justify their actions. The concept of adversarial or confrontational interviewing also exists in police work (e.g., Baldwin 1993) and forensic auditing (e.g., D. Gray 2008), where questioners work to prove the occurrence of actions that are purposefully hidden. In the context of investigative journalism, adversarial interviews (sometimes also referred to as confrontational interviews) conducted with the target of an investigation provide journalists with a chance to test hypotheses and to obtain contradictory evidence or explanatory statements. Less related to establishing facts but still an important aspect of these interviews, from the perspective of meeting community knowledge standards, is the goal of allowing the subjects of an investigation to present their own narratives.

A handbook published by the nonprofit Investigative Reporters and Editors organization (Houston and Investigative Reporters & Editors, Inc. 2008) suggests conducting adversarial interviews near the end of an investigation, once documents have been collected that can be used

to determine if the subject of investigation is telling the truth. Journalists interviewed by Ettema and Glasser (1998) are split on the order of document collection and adversarial interviews, with some advocating confronting the subject of investigation as early as possible in order to explore possibilities, while others recommend conducting such interviews as late as possible, in order to be able to present the interviewee with confirming documents or to avoid warning them of the investigation too early.

Formers and leaks

Because investigative journalism focuses on revealing wrongdoing, their most important sources of information are often former employees or others who have deep knowledge of an organization's or individual's actions but are not in a position to be harmed by the revelation of those actions. Indeed, the motivation to leak a story or documents to a journalist or to participate in an interview can stem from negative feelings or a desire to seek retribution.

While investigative journalists treat leaks and interviews with former employees with incredulity, the validity of such potentially revealing sources for academic researchers is much harder to establish. A strongly motivated interviewee, for example—or, more troublingly, one who approaches the researchers volunteering information that has not been previously sought—raises red flags in a context in which objectivity or disinterest are often if not assumed then at least strived for. In case studies focusing on firms, for example, the objectivity of former employees can be questioned if they left on contentious terms (Vissak 2010). At the same time, these former employees might be the only sources for information on the firm's unethical behavior.

In an example of published research that does rely on information provided by former employees, John (2019) attempts to understand the processes and rationales behind Facebook's claims related to the number of connections made on the platform between people on opposing sides of major conflicts and ultimately documents a series of failed attempts to contact the company for more information. Here, the researcher is only able to learn about these numbers by contacting a former Facebook employee, who is able to speculate on the "organizational and bureaucratic" processes that might provide an explanation (p. 12).

While still relatively rarely discussed, leaks' relevance to academic research is often more associated with datasets than interview data. WikiLeaks' release of the Pentagon Papers in 2010, for example, gave researchers an important source for estimating the death toll in Iraq—however, the credibility of such leaked data is uncertain at best within academic contexts. Gary King, director of the Institute for Quantitative Social Science at Harvard, for example, has noted that students who have sought IRB permission to use such data have been told "they weren't allowed to touch it" (qtd. in Bohannon 2010).

Challenges

Methods

Investigative work seeks to establish truthful accounts of specific situations. While investigative journalists might hope that their work raises awareness of broad societal patterns, their visible output is the narration of a single story. Auditors and detectives have even less interest in anything beyond the question at hand. Was the fire started purposefully? Was the cause of death natural? Were funds transferred and to whom?

But researchers often attempt to generalize from specific cases to more abstract theories, which calls into question the methodologies on which they might draw. While research of information technology does result in individual case studies and ethnographic accounts, all but the most descriptive of studies aim to produce theories or insights with broader implications. This goal holds, as well, for methodologies that have been suggested recently and that propose alternatives to inductive studies that generalize from multiple observations (e.g., McFarland, Lewis, and Goldberg 2016; Tavory and Timmermans 2014).

With such methodologies in mind, we suggest that the movement from specific investigations to theories presents less of a challenge than the new data collection techniques that investigative approaches require. The problem is not the relationship between data and theory but between, perhaps, methods and methodologies, with important questions raised about the appropriateness of a given method within an existing methodological framework or community of practice.

For methods such as interviewing, which are common in a range of academic areas, the challenge will be understanding how an investigative attitude changes the method and the extent to which that change then aligns or does not with community norms. While academic researchers conventionally conduct interviews, for example, most would find the concept of an adversarial interview to be quite foreign. A rare discussion of “confrontational” interviews in the context of social science research, for example, notes that in contrast to the more conventional interview goals of consensus formation or empathy, confrontation highlights the elements of conflict and power present in the interview situation (Kvale 2007, p. 75). Still, even this conception of confrontational interviewing diverges from the realist underpinnings of investigative journalism. Kvale describes confrontation as aiding in the construction of understanding, but this methodological approach becomes nearly nonsensical in the context of interviewing people who do not wish to have information revealed about themselves or activities of which they are a part.

While we assume that such interviews would find some useful role in researchers’ work, we are also unsure of what that role would be, the specific methodological choices around it and the way that it would be articulated. For example, conducting confrontational interviews early in the research process might, as some journalists suggest, widen the possible field of inquiry. On the

other hand, waiting to conduct confrontational interviews until the end of an investigation might play a more confirmatory role or satisfy ethical expectations of a community. The methods of receiving leaks and seeking out formers would require similar negotiation.

A related challenge is the lack of training available to graduate students in investigative methods. Indeed, in our experience, the methods we discuss here—and, more generally, the entire idea of an investigative approach—runs orthogonal to how we were trained to conduct research. Most glaringly, within the areas with which we are familiar (largely the social sciences, but individual authors have taken graduate methods courses, as well, in design, business and the humanities), instruction in methods places a warranted emphasis on a conception of ethics and pragmatics that gives research participants a great deal of agency to determine how they interact with researchers and what information they provide. If students are taught to uncover societally harmful behavior, they will need guidance in negotiating the agency of participants as well as in new forensic methods that will likely vary by domain.

Institutional

The methods described above are almost impossible to imagine in the context of an IRB proposal due to methodological and ethical assumptions. Following on the origins of IRB in medical research, as well as the institutional history of applying these standards to some kinds of social science work, an assumption of inductive methodologies dominates the process. The clearest signal of this assumption is found in the conventional definition of research as producing generalizable knowledge. The logic of investigation, however, rests firmly on proving the specific case without showing its generalizability. Even if investigative journalists do intend for their narratives to address broad societal patterns, their standards for making this connection are far different than those of academic review. While the problem of scholars' work not aligning with IRB assumptions is certainly not unique to investigative approaches and scholars in different communities and institutions have various processes for managing this difficulty (from heuristics, tricks and magic words used in applications to simply not applying for IRB approval), if researchers pursue investigative approaches, these questions are likely to remain relevant.

A perhaps more fundamental challenge related to IRB and, more broadly, to institutional standards of research is the assumption that research does not harm those who participate in it. Applying for IRB approval, for example, entails outlining the potential harms of research and how they will be communicated to research participants. However, the moral assumptions of investigative work to some extent guarantee harm—revealing information that someone wants to hide is, after all, a form of harm, and the potential results of investigations (criminal convictions or damage to someone's reputation, for example) are unlikely to be risks that the subjects of investigation would sign on for. Indeed, projects in which everyone involved has given consent are probably by definition not investigative.

Obtaining consent from leakers or former employees presents another challenge related to IRB. Because investigative journalists often rely on such sources to both prompt an investigation and also to provide what academic researchers would consider data, it is unclear how informed consent might be obtained. Many investigative journalists, for example, maintain systems such as encrypted messaging through which they can obtain information while ensuring the integrity of the communication. While such practices represent attempts to act in an ethical manner that protects informants' privacy and attempts to prevent potential harms, they do not align with the institutional conception of research projects as conceptualized and approved before data collection begins.

While the simple answer to these challenges is that investigative projects conducted by scholars are likely a poor fit for IRB review, the more complicated answer is that institutional standards and values are wrapped up in scholars' career advancement in ways that will likely require community effort to (continue to) define alternative research approaches as valuable contributions.

Moral

A final challenge to investigative approaches is negotiating their moral nature. In order to justify the kind of harm entailed by revealing information against a subject's wishes, a moral argument is needed. For journalists, the community standards for producing this argument require that behavior violate an explicit, external rule, such as a law or code of business ethics. Alternately, as Glasser and Ettema (1989) note, journalists can rely on an understanding of their community's moral standards, stopping themselves from declaring actions to be immoral or unethical but assuming that readers will see them from this shared perspective; however, this method of moral justification risks reinforcing consensus values.

These standards are unlikely to satisfy communities of academic researchers, for whom such external standards would likely require additional articulation or theoretical grounding. Indeed, the justification for academic research is so firmly tied to the goal of not harming research participants that it is difficult to imagine how communities of scholars might evaluate arguments justifying harm in order to expose unethical practices. At the same time, practices such as anonymizing participant data or purposefully obscuring the context of a study seem, in an investigative context, as unethical as revealing them might be in other cases.

The explicit articulations that investigative approaches to academic researchers are likely to require have the potential to lead investigative work in new directions: the goal of academics pursuing investigation, after all, is not to recreate the work of investigative journalists but to draw on it as inspiration for enlivening research, to obtain access to new empirical phenomena and to produce scholarship that has interesting public potential.

Conclusion

Research conventions in the social sciences were developed under assumptions that do not hold in relation to contemporary information technology, a context where companies that produce the most influential and impactful phenomena work to keep their practices hidden from government oversight, academic researchers and an increasingly skeptical public. It is not ethical to allow the individuals and companies who perpetuate harm to control how researchers access their processes and perspectives. Broadly, we have suggested that, contra practically everything we have been taught and continue to teach our students, these data should not be obtained through consent but through invasive methods that go against the wishes of those being studied. If we do not at least explore this possibility for information technology research, then we risk losing access to phenomena that are potentially crucial to our field. We also risk losing an opportunity to perform a public service by turning our expertise to exposing harm and injustice.

In order to fully understand contemporary information technology, it is insufficient to only analyze technical objects without also digging into the processes that made them. It is inappropriate to partner with technology companies, especially when the partnership is not recognized as such and research questions are implicitly controlled. And policy suggestions already rely on information that has been obtained through investigation.

Exploring investigation as a research approach will require looking to new models, such as the investigative journalists discussed here. It will require methodological articulation, ethical debate, methods training and institutional negotiation. We do not know what the research outcomes of such an approach might look like, but we think this is a conversation worth having.

References

- Alaimo, Kara. 2018. "Twitter's Misguided Barriers for Researchers." *Bloomberg*, October 16, 2018. <https://www.bloomberg.com/opinion/articles/2018-10-16/twitter-s-barriers-for-academic-researchers-are-misguided>.
- Andersen, Christian Ulrik, and Søren Bro Pold. 2011. *Interface Criticism: Aesthetics Beyond the Buttons*. ISD LLC.
- Baldwin, John. 1993. "POLICE INTERVIEW TECHNIQUE Establishing Truth or Proof?" *The British Journal of Criminology* 33 (3): 325–52. <https://doi.org/10.1093/oxfordjournals.bjc.a048329>.
- Bardzell, Jeffrey, and Shaowen Bardzell. 2013. "What Is 'Critical' about Critical Design?" In *Proceedings of the 2013 ACM Annual Conference on Human Factors in Computing Systems*, 3297–3306. CHI '13. New York, NY, USA: ACM. <https://doi.org/10.1145/2466416.2466451>.
- Bertelsen, Olav W., and Søren Pold. 2004. "Criticism as an Approach to Interface Aesthetics." In *Proceedings of the Third Nordic Conference on Human-Computer Interaction*, 23–32. NordiCHI '04. New York, NY, USA: ACM. <https://doi.org/10.1145/1028014.1028018>.
- Bogost, Ian, and Nick Montfort. 2009. "Platform Studies: Frequently Questioned Answers." *Digital Arts and Culture* 2009, December. <http://escholarship.org/uc/item/01r0k9br>.
- Bohannon, John. 2010. "Leaked Documents Provide Bonanza for Researchers." *Science* 330 (6004): 575–575. <https://doi.org/10.1126/science.330.6004.575>.
- Bruns, Axel. 2019. "After the 'APIcalypse': Social Media Platforms and Their Fight against Critical Scholarly Research." *Information, Communication & Society* 0 (0): 1–23. <https://doi.org/10.1080/1369118X.2019.1637447>.
- Clayman, Steven, and John Heritage. 2002. *The News Interview: Journalists and Public Figures on the Air*. Cambridge University Press.
- Dahlgren, Peter. 1992. "Introduction." In *Journalism and Popular Culture*, edited by Peter Dahlgren and Colin Sparks. SAGE.
- Dourish, Paul. 2014. "No SQL: The Shifting Materialities of Database Technology." 2014. <http://computationalculture.net/article/no-sql-the-shifting-materialities-of-database-technology>.
- Dunne, Anthony. 2005. *Hertzian Tales: Electronic Products, Aesthetic Experience, and Critical Design*. Cambridge: MIT Press.
- Elliott, Justin. 2019. "Congress Scraps Provision to Restrict IRS From Competing With TurboTax." *ProPublica*, June 5, 2019. <https://www.propublica.org/article/congress-scraps-provision-to-restrict-irs-from-competing-with-turbotax>.
- Ettema, James, and Theodore Glasser. 1998. *Custodians of Conscience*. New York: Columbia University Press.
- Ettema, James S., and Theodore L. Glasser. 1984. *On the Epistemology of Investigative Journalism*. <https://eric.ed.gov/?id=ED247585>.
- Fishman, Mark. 1980. *Manufacturing the News*. University of Texas Press.
- Freelon, Deen. 2018. "Computational Research in the Post-API Age." *Political Communication* 35 (4): 665–68. <https://doi.org/10.1080/10584609.2018.1477506>.
- Gerlitz, Carolin, and Anne Helmond. 2013. "The like Economy: Social Buttons and the Data-Intensive Web." *New Media & Society* 15 (8): 1348–65. <https://doi.org/10.1177/1461444812472322>.

- Glasser, Theodore L., and James S. Ettema. 1989. "Investigative Journalism and the Moral Order." *Critical Studies in Mass Communication* 6 (1): 1–20.
- Gouldner, Alvin Ward. 1976. *The Dialectic of Ideology and Technology: The Origins, Grammar, and Future of Ideology*. Seabury Press.
- Gray, Colin M., Yubo Kou, Bryan Battles, Joseph Hoggatt, and Austin L. Toombs. 2018. "The Dark (Patterns) Side of UX Design." In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 534:1–534:14. CHI '18. New York, NY, USA: ACM. <https://doi.org/10.1145/3173574.3174108>.
- Gray, Dahli. 2008. "Forensic Accounting and Auditing: Compared and Contrasted to Traditional Accounting and Auditing." *American Journal of Business Education* 1 (2): 115–26.
- Houston, Brant, and Investigative Reporters & Editors, Inc. 2008. *Investigative Reporter's Handbook: A Guide to Documents, Databases, and Techniques*. Fifth edition. Boston: Bedford/St. Martin's.
- John, Nicholas A. 2019. "Social Media Bullshit: What We Don't Know About Facebook.Com/Peace and Why We Should Care." *Social Media + Society* 5 (1): 2056305119829863. <https://doi.org/10.1177/2056305119829863>.
- King, Gary, and Nathaniel Persily. 2019. "A New Model for Industry-Academic Partnerships." *PS: Political Science and Politics*.
- Lewis, Seth C., and Oscar Westlund. 2015. "Big Data and Journalism." *Digital Journalism* 3 (3): 447–66. <https://doi.org/10.1080/21670811.2014.976418>.
- Mackenzie, Adrian. 2005. "The Performativity of Code: Software and Cultures of Circulation." *Theory, Culture & Society* 22 (1): 71–92. <https://doi.org/10.1177/0263276405048436>.
- . 2006. *Cutting Code: Software and Sociality*. Peter Lang.
- McFarland, Daniel A., Kevin Lewis, and Amir Goldberg. 2016. "Sociology in the Era of Big Data: The Ascent of Forensic Social Science." *The American Sociologist* 47 (1): 12–35. <https://doi.org/10.1007/s12108-015-9291-8>.
- Merchant, Brian. 2019. "Tech Journalism's 'on Background' Scourge." *Columbia Journalism Review*, July 17, 2019. <https://www.cjr.org/opinion/tech-journalism-on-background.php>.
- Montfort, Nick, and Ian Bogost. 2009. *Racing the Beam: The Atari Video Computer System*. Platform Studies. Cambridge, Mass: MIT Press.
- Mozilla. 2019a. "Open Letter: Facebook, Do Your Part Against Disinformation." *The Mozilla Blog* (blog). February 11, 2019. <https://blog.mozilla.org/blog/2019/02/11/open-letter-facebook-do-your-part-against-disinformation>.
- . 2019b. "Facebook and Google: This Is What an Effective Ad Archive API Looks Like." *The Mozilla Blog* (blog). March 27, 2019. <https://blog.mozilla.org/blog/2019/03/27/facebook-and-google-this-is-what-an-effective-ad-archive-api-looks-like>.
- "NSF Program on Fairness in Artificial Intelligence in Collaboration with Amazon (FAI) (Nsf19571)." n.d. Accessed July 24, 2019. <https://www.nsf.gov/pubs/2019/nsf19571/nsf19571.htm>.
- "Our Facebook Partnership." n.d. Accessed July 24, 2019. <https://socialscience.one/our-facebook-partnership>.
- Parasie, Sylvain. 2015. "Data-Driven Revelation?" *Digital Journalism* 3 (3): 364–80. <https://doi.org/10.1080/21670811.2014.976408>.
- Pasquale, Frank. 2015. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Cambridge, Massachusetts ; London, England: Harvard University Press.

- Schudson, Michael. 1989. "The Sociology of News Production." *Media, Culture & Society* 11 (3): 263–82. <https://doi.org/10.1177/016344389011003002>.
- Tavory, Iddo, and Stefan Timmermans. 2014. *Abductive Analysis: Theorizing Qualitative Research*. Chicago: University Of Chicago Press.
- Zelizer, Barbie. 2004. *Taking Journalism Seriously: News and the Academy*. SAGE.
- Zuboff, Shoshana. 2019. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. 1 edition. New York: PublicAffairs.