Research Article

Toward Data Justice: Understanding Police Shooting Data Systems and Narratives

—RACHEL ATHERTON

***Abstract*—Background:** The Washington Post*'s Fatal Force fatal police shooting database was first created in 2015 to fill a gap in official data collection on police use of force.* **Literature review:** *Critical data studies scholarship suggests that data system design is rhetorical and communicates cultural values, not just numerical data. Narrative research methods, which focus on thick, rich contextual data, could help address the rhetorical and cultural dimensions of data system design.* **Research questions:** *1. How do data collection, aggregation, and curation practices influence data stories about crime and violence? 2. How does Fatal Force, a data system about police use of force that originated outside law enforcement, prioritize and organize information? 3. How does designing data systems with explicit, highly specific goals and aims (like the inclusion criteria and purpose of Fatal Force) influence the system as a whole?* **Methodology:** *Using "unblackboxing," a combination of narrative and critical data studies methods, I analyzed the Fatal Force database and its accompanying data stories. I compare this database with its institutional counterpart in the FBI's fledgling Use of Force database.* **Results:** *Fatal Force is constructed with particular attention to questions that users may have about police brutality, police shootings, and race. Fatal Force's data stories use narrative commonplaces like sociality and temporality to humanize the issue of police use of force and communicate greater nuance. The FBI's Use of Force database shows an orientation toward police perceptions of use-of-force incidents and a lack of attention to national conversations about police brutality.* **Conclusions:** *Data systems show clear perspectives on the issues that they describe, which influence how users encounter the data system, how useful the system can be for various users, and how inclusive or just the data system is.*

Index Terms*—Data system design, narrative, rhetorics of data, unblackboxing.*

The field of technical communication has enjoyed a growing body of research on social justice in the last several years. Scholars have advocated for a social justice turn (see [1], among others), and others have opened up avenues by which technical communicators can do this social justice work in our research and pedagogy [2]–[5]. My work aims to continue these lines of inquiry by addressing unjust law enforcement data practices that disproportionately target Black, Indigenous, and People of Color (BIPOC) [6] and do not address experiences of marginalized people. Citizens have, in many cases, worked to replace or augment these institutional data with richer, more contextual data that focus on social justice, such as *The Washington Post*'s Fatal Force database documenting shootings perpetrated by on-duty police, and I analyze these citizen data initiatives to make interventions in law enforcement data systems.

Big data condenses complex human behavior down into easily digestible insights, useful for understanding trends, making predictions, and taking action to solve problems. Data-driven and evidence-based approaches dominate many industries, both for their usefulness in creating profit and for their apparent lack of bias. But as scholars studying big data and algorithms have shown [7]–[9], decentering humans from the decision making process does not eliminate systemic inequities such as racism, misogyny, or homophobia; in fact, these inequities may become further pronounced. Big data allow for incredible efficiency, but they lack the human and cultural exigencies necessary to understand and address social problems ethically and justly.

Data systems may be largely composed of quantitative information, but they are still texts with rhetorical purpose and communicative effects. Regardless of medium, texts contain and communicate narrative elements that shape the way audiences understand them [10]. The lens and practice of narrative inquiry can help us study data

Manuscript received March 15, 2021; revised October 15, 2021; accepted October 16, 2021. Date of publication January 26, 2022; date of current version March 17, 2022.
The author is with the Department of English, Purdue University, West Lafayette, IN 47907 USA (email: ratherto@purdue.edu).

IEEE 10.1109/TPC.2021.3138008

Practitioner Takeaway

- Unblackboxing, a methodology that looks at narrative and rhetoric in data systems, can help show designers what messages data systems are sending and how those systems impact users.
- Fatal Force, a database describing fatal police shootings in America, differs significantly in orientation and construction from the FBI's own database about police use of force.
- The perspectives from which a data system is constructed strongly influence what information is considered valuable about that issue and what questions users can ask about the issue.

systems, their implicit and explicit messages, and their impacts. Technical communication, with its grounding in rhetorical theory and its focus on the experience of users (not just designers), is well-suited for interventions in data system design. Using narrative principles to understand user experience has proven successful in prior work (i.e., [11]). Combining narrative inquiry with critical data study methods in analyzing data systems can help humanize these systems, shedding light on the invisible work and tacit knowledge embedded within to answer questions about whose knowledge counts and why.

The article reports on research that I conducted on *The Washington Post*'s Fatal Force fatal police shooting database, which was first created in 2015 to fill a gap in official data collection on police use of force. I focus specifically on crime, violence, and policing in the US as these are systems that have become shaped and defined by data; that focus on data has produced a disproportionate and negative impact on marginalized and vulnerable populations. Technical communication scholars have an ethical obligation to investigate and improve technical systems that disenfranchise vulnerable populations [12]. I examine data collection, maintenance, and publication practices in Fatal Force to promote what Taylor calls data justice [13]. Taylor's definition hinges on three principles: "visibility, digital (dis)engagement, and countering data-driven discrimination," but also emphasizes the importance of a

> methodological engagement with the political economy of data, to determine not only *what*, but *who* is important and *how* they relate to the desired outcomes. [p. 8, emphasis in original]

What is crucial about data justice in Taylor's framework is that both the necessity of data for human growth and thriving, and the risks and harms of data, especially for already marginalized people, are held in balance—Taylor acknowledges that data justice must be a middle way [13].

In what follows, I outline how data are used to make decisions about crime and policing and how those decisions impact multiply marginalized populations. In this article, I present two phases of a larger project. I couple a narrative inquiry with a new method that I develop called unblackboxing. While the term is not new—Dorpenyo used it in his study of biometric election technology in Ghana [14]—it has yet to be taken up as a method for critically studying data systems. Phase 1 analyzes what datasets and documents are collected in a case, how and by whom they are collected, what the collection instruments are, as well as what maintenance and storage procedures influence the end data output to a user. Phase 2 investigates how the institution or community in question stories their data, especially if the story has changed over time via news, texts, reports, or any other documents that talk about the data. Unblackboxing Fatal Force reveals implications for rhetorical data system design, specifically in how data elements, public interfaces, and stories work together to achieve specific goals.

Literature Review

Data in Crime, Policing, and Violence: Law enforcement agencies (LEAs) regularly use data to make decisions, both large and small. Every traffic stop and engagement with law enforcement produces datasets that are aggregated into national databases. In the process, much of the richer, thicker data present in the original reports get lost because the national database collects only basic information. Researchers, policymakers, LEAs, and others use these data to make statements about trends in crime over time. Consequences unfold from there—incomplete data do not give an accurate picture of crime trends; so data narratives about how to police, where to police, and whom to police are necessarily incomplete.

Police reports and other data about crime in an area (arrests or call-outs, for instance) are used by algorithmic audiences too. In recent years, Benbouzid writes, police forces across the US have turned to predictive software such as Hunchlab or PredPol to help map trends in their jurisdictions and manage resources [15]. These programs initially allowed officers to test theories and intuitions against actual data about crime in their locales; essentially, they were developed as tools to help officers do their work better and target their investigations and patrol efforts. Later, though, these tools turned from intuition to prediction and stopped requiring interpretation by police officers—in other words, officers would not have to determine whether the machine's output was reasonable or accurate. Benjamin critiques this practice as a tool of the "New Jim Code," a technological way of segregating and discriminating against people of color and particularly Black people [9], and empirical studies back up this critique.

As Lum and Isaac found, the widespread predictive policing algorithm PredPol perpetuated narratives about illegal drug use occurring mostly in low-income neighborhoods of color, despite the fact that these crimes occur at similar rates across all neighborhoods [6]. In other words, the algorithm identified poor people and people of color as more likely to be criminals not because they were engaged in more illegal drug use than wealthier people or white people (the kind of trend big data is useful for identifying) but because police had made more arrests in poor neighborhoods of color in the past (local context that the algorithm cannot account for).

Transparency and neutrality are common reasons cited for using predictive policing and other data-driven approaches to law enforcement, but scholars who study big data and policing often disagree that these approaches are not transparent or neutral at all. Brucato, for instance, wrote that

> The turn to big data for transparency is intended to avoid any need for interpretation, and thus sets databases up to obscure the interpretive work done by data scientists, data visualizers, reporters, and the recipients of their presentations. [16, p. 4]

Brucato observes that data in the case of crime and police action are given agency and independence, which is propped up as a speaker in its own right to defer responsibility away from humans in charge. This application of data in policing tells a narrative of impartiality via technology, but antenarratives from activists and marginalized communities suggest otherwise.

Crime data's first manifestation might be the police report. Leslie Seawright writes in her study of the police report writing process that an individual police report describes one incident, potentially where multiple crimes or no crimes occurred, and gives both quantitative and qualitative data—including narrative—about the incident [17]. Yu and Monas note similar processes in their study of police report writing, adding that despite the complexity of the task of writing these reports, police officers themselves are often not actually well-equipped to work in the genre [18]. The police officer writes the report, sometimes with contributions from peers or from superior officers, and the report becomes evidence—another kind of data—in each connected crime.

The report is used by several groups such as officers, lawyers, judges, and even the people mentioned in the report (victim or perpetrator) as any charges make their way through the justice system, all according to their contexts. Importantly, Seawright reports that often officers can use police reports, whether intentionally or not, to hide information and avoid admitting their own mistakes in each encounter—each officer writes their own reports, and so the narrative is always from the officer's perspective only. When the officer later uses the report as a mnemonic device for testifying in court, whatever is in the police report becomes fact because of the length of time between different benchmarks in criminal proceedings [17].

Any error introduced in the police report stage is then reinforced and magnified as the incident is entered into data for the LEA itself, and these data are then sent up the chain to eventually reach national crime data. These national crime databases contain information from most LEAs across the country but not all, and not all agencies report all their data for various reasons. In 2019, the FBI reported that only 44% of US LEAs participating in the Uniform Crime Report (UCR) program were submitting statistics to the FBI using the National Incident-Based Reporting System (NIBRS), which is currently the most detailed, richest data entry method [19]. Other agencies still used the Summary Reporting System, a thinner system with less detail.

The UCR program's data reach is estimated to cover just over 90% of the population in small cities and rural areas and about 98% in metropolitan areas [20]—but not all agencies report to the UCR. Reporting is voluntary such that, for whatever reason, reporting agencies may not send all their data. A 2016 Bureau of Justice Statistics report concluded that "a comprehensive and current roster of all LEAs in the US is currently unavailable" [21, p. 7]. These incomplete crime databases, nonetheless, are the most complete source of information that the government possesses to create national crime reports that are used by policymakers, researchers, and LEAs to make decisions. In this capacity, these incomplete data become reality.

Crime, violence, and victimization are also defined by data in many ways. Evidence-based policymaking requires data on crime rates, types of crime, locations of crime, and characteristics of both victims and offenders (among other factors). Essentially, crime problems that get addressed are crime problems that are supported by data. However, these data are created by people working in a system rife with systematic oppression. For a crime to be recorded and reported at all, LEAs must first recognize it as such; counts for citizens killed by police, for instance, do not match across LEA data, Centers for Disease Control (CDC) data, and citizen data [22]. Similarly, Indigenous activists and researchers have found that national data on missing and murdered Indigenous women do not match with Indigenous communities' own knowledge of this crisis [23]. Conflicting data and their concomitant narratives reveal fundamental differences in the way different publics and counterpublics view and construct the world.

Law enforcement is not the only stakeholder working to create data about crime and violence, however. Private citizens, community organizations, and activist groups recognize that government data are incomplete and have begun working to augment or in some cases, outright supplant or replace government data. These efforts are particularly interesting in that they are often created using some crowd-sourced information from social media or news reports, whereas government data are exclusively sourced from law enforcement officers' own reports. In other words, these community data systems are more inclusive of individual peoples' stories.

Community data systems also tend to be more specialized than government data, and often a function of the questions each is created to answer. Because government data are incomplete, community data systems are created to fill gaps—these may be gaps in reporting, such as the Murder Accountability Project, which attempts to record all murders in the US and describe a clearer picture of murder in the country than the one in the FBI's national data, or they may be gaps in what is considered relevant or useful data, such as *The Washington Post*'s Fatal Force project that tracks all fatal shootings by on-duty police officers in the US. In each case, the data system is not created as a measure to answer **every** question about crime data but specific questions about specific types of crime and violence.

Defining Blackboxing: What is blackboxing in the first place? "Blackboxing," from Latour, refers to "the way scientific and technical work is made invisible by its own success," where a system or process's work is not the object of focus but only the inputs and outputs are [24, p. 304]. As with the proverbial "black box" in mathematics, an input goes into the box, something happens inside the box, and an output emerges, but what process caused the changes inside the box is unknown. The idea of blackboxing or a black box is often used in talk of tech development, referring to hidden, usually proprietary processes and algorithms that cannot be seen by the user, such as Google's search algorithm. We can guess at what work the algorithm is doing by testing different inputs, recording and analyzing the outputs, and comparing the differences, as Noble did in *Algorithms of Oppression* [8]. But despite making inroads, the algorithm is still blackboxed—we cannot read it or analyze it, just the inputs and outputs. At best, our efforts are guesswork.

We need to think of systems here not only as the computer database but also the whole ecology of the data, from collection to reporting and use. If blackboxing is obfuscation of processes and work, then we need to acknowledge that the work includes a lot of human action and decision making. Data often get blackboxed at more than one stage in the data life cycle—depending, of course, on who is collecting them and how they are being presented. If a raw dataset is not included with a report on the data, that is one form of blackboxing—the report forwards an interpretation without also offering the complete basis for that interpretation. Further back, data are often blackboxed in scrubbing processes that "cook" the data to make them easier to use and also remove outliers that could lead to other interpretations.

Even further back, data can be blackboxed in collection, especially if norms, values, and assumptions of study designers, database designers, or data collectors go uninterrogated.

At each stage I enumerate above, one or more people made value judgments about what material was worthy of being data and what material was not. Data collection, scrubbing, analysis, reporting, and use are all necessary functions of data work; I am not suggesting that to unblackbox a data system means dismantling the whole thing or never valuing data. Rather, I am suggesting that to unblackbox a data system means regarding these necessary functions as products of social, cultural, and rhetorical forces embodied in and enacted through data workers—as active choices with many possible alternatives. To unblackbox a data system means analyzing this rhetorical work and this web of connections at every stage so that the values and stories that are baked into a data system are apparent and can be evaluated, rather than remaining hidden and tacitly accepted.

Other scholars have used the term "unblackboxing" before. In technical communication, Isidore Dorpenyo used the term in his doctoral dissertation on Ghana's biometric election technology [14]. Specifically, he defined unblackboxing as accounting for rhetorical conditions, firmly emplacing technology within and tracing out the network of connections to various actors and forces. For Dorpenyo, unblackboxing as a concept works with and is embedded in rhetorical concerns. In my use of unblackboxing, I devote special attention to invisible rhetorical work within the data system. Because data are often treated as an objectively true measure of the world, I argue that it is especially important to understand the constituent narratives and rhetorical moves that are at work in a data system but are "made invisible by [the system's] success" [24, p. 304].

Critical Data Studies, Narrative, and Technical Communication: Dalton and Thatcher named critical data studies in 2014; their work as geographers emphasized not just space but also place and context; so they saw a need for critical examination of data collection, analysis, and use practices, especially with the growth of data science and the proliferation of big data. When "big data" as a term was normalized, they argued, it would "[recede] from conscious consideration." In their initial sketching of the field, they listed six principles for critical data studies as follows.

- [Situate] big data in time and space.
- Technology is never as neutral as it appears.
- Big data does not determine social forms.
- Data is never raw.
- Big is not everything.
- "Counter-Data" can allow for the possibility of subversive and liberatory data use [25].

These principles have been picked up by scholars in a number of disciplines, and similar critiques of big data have emerged in other fields. As boyd and Crawford observed, one of the reasons big data has power in discourse is that it is taken as automatically valuable and real, without critical questioning of whether a dataset's prediction fits with reality for all or even whether big data is the appropriate response to the problem at hand [26]. Gitelman, similarly, wrote that data are always "baked," not raw, because they are constructed and interpreted—they do not have meaning without the work of humans to make them tell a story [27]. Benjamin picks up on the importance of story for data communication as well, arguing that not only do data tell a story but also that big data carries the risk of dominating and excluding other stories [9]. Other works such as Noble's *Algorithms of Oppression* may not call themselves critical data studies as such, but they do embrace some of that discipline's frameworks and motivations—namely, cracking open the black box of data and algorithms in the interests of showing that technology's effects are not equal [8]. Ultimately, critical data studies is a useful gathering place for scholars in multiple disciplines bringing multiple disciplinary approaches to the problem of uncritical big data collection, analysis, and distribution.

Scholars in technical and professional communication (TPC), rhetoric and composition, and data science are currently grappling with big data as it informs social change. In composition, we teach students to write with data or to consider algorithmic audiences [28], [29]. In rhetoric, scholars are examining how personal data influence corporate messaging, construct us as subjects, or offer new opportunities for argument [30], [31]. And in technical communication, data offer new methods; data visualization invites questions about user experience and technical communication in a variety of public, educational, and corporate settings; and big data projects invite a mixture of critique and intervention [32]–[36]. Graham et al. [33] conducted a pilot mixed-methods study using big data for genre analysis, not unlike some corpus research but emphasizing the importance of retaining rhetorical analysis in addition to the digital data analysis.

Similarly, the common thread among most of the recent data visualization work in TPC is the necessity of rhetorical and technical communication expertise to make sure that visualizations are legible, accessible, useful, and ethical for diverse audiences.

- I argued in a recent study of US CDC visualizations that even the most objective-seeming data tables were making statements about what COVID-19 information was valuable and worth viewing together; officials had made rhetorical choices about what data to display and how to display them and were passing their choices on to the public without acknowledging those choices [32].
- In the realm of pedagogy, Wolfe wrote similarly that students need to be taught how to make ethical and rhetorical choices about which data to visualize and in what ways [36].
- In the context of scientific big data at the Los Alamos National Laboratory, Overmyer described scientists' work designing interfaces to help experts glean insights from big data visualizations; crucial in this process was human expertise and attention because experts need to monitor and understand the data quickly to avoid costly lab mistakes [35].

This wide reach of data-related topics translates to other disciplines as well, where data research is thriving in fields like sociology, history, science and technology studies, information studies, and political science [7], [8], [37]–[39]. Common to most of these texts is an acknowledgement of or emphasis on the rhetorical, communicative, and interpretative aspects of data—in all these areas, scholars agree that data are not neutral, and although they are certainly useful, they are also just as subject to rhetoric and narrative as anything else.

Narrative inquiry remains a key methodology in humanities. As Clandinin writes, narrative inquiry is about studying experience; specifically, "narrative inquirers understand experience as a narratively composed phenomenon" [40, p. 16]. Part of that includes honoring individual experience—one person's story—as real, legitimate, and true, in so much as anything can be true, but narrative inquiry is concerned not only with the highly particular. Rather, this methodology is

> an exploration of the social, cultural, familial, linguistic, and institutional narratives within which individuals' experiences were, and are, constituted, shaped, expressed, and enacted. [40, p.18]

In other words, narrative inquiry attempts to marry individual experience with cultural experience, small with big, and instance with context. This focus lends itself particularly well to pairing with big data because big data typically has many instances but little context (i.e., it is big but thin), and narrative inquiry may have few instances but a great deal of context (i.e., it can be small but thick).

Narrative inquiry's focus on personal experience does not mean that it cannot be used to understand connections across artifacts, texts, and stories, sometimes with disparate sources. Its three commonplaces (temporality, sociality, and place) structure narrative analysis such that researchers can gain comparative insights that speak to an ecology of being rather than only an individual.

- Temporality is not just the linear passage of time as a fact but also considers time to be an ordering mechanism for a person's history and life (which may or may not be linear). Temporality as a narrative commonplace is "embodied in the person" [40, p. 40] and tied also to relationships and places.
- Sociality, the second commonplace, describes the "milieu" where narratives unfold, specifically the "cultural, social, institutional, familial, and linguistic" forces surrounding and interacting with a person as well as the relationships between these narratives [40]. Narrative is personal, and so too is the sociality commonplace, in that the external forces described above interact with and shape personal experience, and vice versa.
- The third commonplace of narrative inquiry, place, speaks to other kinds of narrative (personal, cultural, and institutional), which also speak back to place, mutually shaping and reinforcing one another. Grounding a thing (a data system, an institution) or a person in place ensures that the narrative has a concrete setting with specific, contextual questions and considerations rather than existing in a nebulous abstraction (where it is easier to accept narrative whole cloth and without critiques) [40].

Narrative's staying power has largely been in its ability to draw out tacit knowledge and provide rich qualitative data about place and lived experience that most other methods cannot accomplish. Narrative inquiry attends to place, context, and culture in unique ways and is inherently participatory, in that participants get to tell their own stories. Narrative is a key part of unblackboxing because data that are easily blackboxed are usually data that are thin. They have lost context, been taken out of their place and time, and been disconnected from the material thing or physical event that they represent. Unblackboxing's goal is to rebuild data's web of connections to place, people, culture, story, and systems. Narrative inquiry is an especially useful tool for this aim because of its emphasis on making tacit knowledge explicit—blackboxing makes so much of a system tacit, behind the scenes, accepted and taken for granted, that we miss how the system functions. We ignore the tacit knowledge—the stories—that formed the data system in the first place.

The three commonplaces of narrative inquiry also lend themselves to mixing with large datasets; temporality, sociality, and place all add context to big data that may include elements of these commonplaces but lack richness. For instance, a dataset gathered from global positioning system (GPS) tracking in Verizon cell phones in an area may include millions of data points recording phone users' movements across town, tracking who went where at what time, but that dataset likely would not include thicker components such as relationships between individuals, the importance of certain places, or the reason why a person took one route in the morning but another in the evening. In other words, narrative inquiry's specific commonplaces provide the why and the how that big data tend to lack.

The narrative inquiry's focus on the particular and the individual also makes it especially well-suited to research that seeks to amplify the experiences of the invisible, ignored, and marginalized. Scholars like Jones have called for and begun to use more narrative research methods in technical communication as part of a social justice approach to the work of the field [41]–[43]. Specifically, Jones builds on Clandinin's narrative commonplaces (described above) to show how they can be used as tools for social justice and coalitional learning; she argues that place, temporality, and sociality are "well suited to dwelling in the spaces between" because they allow researchers to "[acknowledge] complex, pluralistic, and contextual realities" that multiply marginalized groups tend to experience [43, p. 520]. Especially crucial for using this narrative perspective to analyze data systems is Jones' point that researchers must "linger with the ambiguous, malleable, and movable, rather than the static and stayed"—because blackboxing crystallizes a data system into a singular object that fades out of sight, unblackboxing needs to bring the "ambiguous, malleable, and movable" back to the fore [43, p. 520].

Research Questions

In this study, I am guided by the following research questions:

RQ1. How do data collection, aggregation, and curation practices influence data stories about crime and violence?

RQ2. How does Fatal Force, a data system about police use of force that originated outside law enforcement, prioritize and organize information?

RQ3. How does designing data systems with explicit, highly specific goals and aims (like the inclusion criteria and purpose of Fatal Force) influence the system as a whole?

Methodology

In what follows, I outline the two phases of my methodology and the way that I analyzed the data. Together, the two phases combine critical data study methods with narrative analysis using Clandinin's commonplaces [40] to unblackbox *The Washington Post*'s Fatal Force police shooting database.

Phase 1 Data Systems This stage involves analyzing the data themselves: information, storage system, collection and maintenance practices, etc. Different cases will have different materials in this category based on the databases used in the case and the level of documentation and updating taking place. Analyzing data in this stage means viewing datasets like texts, and using the same methods one might use for grounded theory research. Because of the focus on categories at this stage, institutional critique [44] is useful here in determining what to look for, though because every data system is different, researchers need to leave room for uniqueness and for encountering a system on its own terms [45]. A great deal of the hidden rhetorical work in designing a data system has to do with the categories and options for data collection and entry that are baked into the system, along with the stakeholders and values that underlie those choices of information inclusion and exclusion. Identifying these categories, stakeholders, and values lets researchers more clearly analyze the rhetoric of the system itself.

TABLE I
HEURISTIC QUESTIONS AND EXAMPLE FINDINGS FROM PHASE 1 OF UNBLACKBOXING FATAL FORCE

Phase 1 Questions	Example Findings (Fatal Force)
What counts as data in this system?	Each row in the database represents a person fatally shot by on-duty police; data elements focus mainly on the victim and circumstances of the shooting such as weapons involved, body cameras, and mental health.
How are categories related to one another?	Each data element focuses on either the victim's characteristics, the circumstances of the shooting, or the location.
Where are the data stored (with whom, and/or physically on which machines or in which spaces?)	With the *Washington Post*; on GitHub
How are the data cleaned?	Entries are not removed based on lack of information; they may be updated as more information becomes available
How are the data retrieved for analysis?	Through .csv or Excel file; data are stored in "raw" form on GitHub and available in an interactive interface on the *Washington Post*'s website
Who collects data?	Team of *Washington Post* reporters collects data via police sources, news media nationwide, original reporting, social media, tips (so additional data collectors may include police, other reporters, and citizen stakeholders).

Phase 1 answers the questions listed in Table I, which can be adapted or specified for a specific case; the table also lists major findings for each question from Fatal Force to exemplify what this phase might look like.

Each question in this heuristic gets at one or more of the three goals of this phase: categories, stakeholders, and values. Individual data systems may require adding, removing, or modifying questions in the heuristic.

Phase 2 Data Documents This stage is one level removed from the data system itself. Objects of interest here are training materials for system users, reports about the data, and any other documentation that discusses how the data are collected or can be used. The goal here is to understand how the data system is framed narratively by designers and users, a process that can happen at all stages in the data's life cycle and that may change based on who is doing the framing or storying.

Phase 2 analysis builds on results from Phase 1 to look for narrative arcs, using Clandinin's narrative commonplaces of temporality, sociality, and place, which, as Jones argued, can be used as tools for social justice and coalitional learning because of their ability to reveal history and connections. Clandinin's commonplaces are all about finding common axes of connection between personal, "cultural, social, institutional, familial, and linguistic" influences on experience [43, p. 40]—researchers can use them as a heuristic for reading data stories across documents [40]. In addition to lending themselves well to social justice work, these commonplaces are flexible and adaptable enough to work for reading more than just typical narratives with characters and plot. Given that unblackboxing needs to flex to work for various kinds of data systems, the kinds of narratives that researchers might encounter vary widely from traditional to totally implicit and nontraditional. In other words, Clandinin's commonplaces are more likely to still serve as a useful heuristic for even the most technical, "objective" data stories.

Phase 2 answers the questions listed in Table II, which can be adapted or specified for a particular case. Each question in this list gets at one or more of the three goals of this phase: connections, perspectives, and actors.

Each system is different enough that one set of guiding questions will not work similarly enough across systems. In Phase 1, a standard heuristic works better because all data systems need to have some common elements by virtue of how databases are organized and constituted—for instance, every data system sorts information into categories and defines those categories in specific ways. Stories and narrative, on the other hand, are dependent on the case such that a premade heuristic will not be specific enough. Rather, coding with the narrative commonplaces, using open coding, and encountering the system on its own terms must inform the questions instead.

TABLE II
HEURISTIC QUESTIONS AND EXAMPLE FINDINGS FROM PHASE 2 OF UNBLACKBOXING FATAL FORCE

Phase 2 Questions	Example Findings (Fatal Force)
What connections between individuals or groups are evident or implied in the narrative? What types of connections are present? Which group sets the terms of the connection?	Often connects shooting victims with loved ones; some narratives draw connections between individuals and demographic groups (i.e., young men, Black men, etc). Journalists typically set terms.
How does place impact social connections or temporality of the narrative?	Place is rarely dominant in any narratives; exceptions usually happen when describing a scene that readers are meant to imagine.
What are the events described in the narrative? Are they historical or conjecture about the future, or both?	Events tend to be historical, though often described cinematically rather than at a distance.
Who or what acts in the narrative?	Shooting victims, their families and loved ones, police officers or agencies.
What language or metaphors are used to convey the events?	Count language, relative time between events, mix of objective “police report” type reporting and emotional “cinematic” description.

All these also build on Phase 1, and because each case will produce different findings there, the Phase 2 work changes accordingly. The complexity in Phase 2 creates challenges, to be sure, but the tensions that may exist between various documents that tell the stories of a data system can be just as revealing and instructive as a coherent narrative across all the documents that a researcher analyzes in this phase. Table II presents rather unified example results from Fatal Force because the stories around this database are generally similar; this may not be the case for every system, and rather than attempting to reconcile differences across narratives, unpacking the roots of those differences (perhaps going back to Phase 1’s categories, stakeholders, and values) could be productive.

Methods In Phase 1, I analyzed *The Washington Post*’s Fatal Force database itself, available via their GitHub as a downloadable .csv file [46], [47]. Also included in Phase 1 was the methodology document linked on the database, which defined how the *Post* created the database and updates it year over year [22]. Lastly, Phase 1 included the interactive web database as well as the archived versions of the database from each year of the project, which tells the stories behind the data but also presents it to the public in an engaging and easily used manner (and therefore can be counted as “the data” for the purposes of this method because many users would likely experience this web database rather than the downloadable .csv files) [46], [48]–[50].

To determine which data stories I would analyze for Phase 2, I first included reporting that was linked on the database itself because these stories are most closely linked with the data for users. I then expanded the document set to include some of the articles from the *Post*’s Police Brutality section. Finally, I searched the exact phrase “fatal force” on the *Post* archive as a whole and included any article whose title indicated it focused on data about police brutality or police shootings. For both the Police Brutality section articles and the articles retrieved via search, I included only articles that focused on broader claims about police shootings and their circumstances. I chose not to include pieces that focused on a single specific incident (such as one shooting, or reports on court hearings or criminal charges for officers involved in one shooting) because those pieces were less related to the Fatal Force database itself. They usually did not engage strongly with the database and either referenced it in one or two lines or included the database in a link in the middle of the article like an advertisement or a “read more” feature rather than as a central component of the story. Similarly, I removed articles that focused on police shootings but made only a brief mention of the Fatal Force database and rather made it a basis for more sustained engagement. Therefore, all the documents analyzed for Phase 2 are either stories specifically about the Fatal Force database or articles reporting on insights derived from Fatal Force data.

Fatal Force’s original exigence was to track cases like Michael Brown’s 2014 killing because, at that time, no LEA was tracking those incidents and releasing data to the public. In 2017, however, the FBI did respond to growing social pressures about police use of force and began its own collection efforts. To shed further light on Fatal Force, I also look at what is currently available from this institutional attempt to fill the same gap that Fatal Force aims to fill. This analysis was limited because, to date, only about 40% of police officers in the US report use-of-force data, meaning that only participation data has been released nationwide and only data for some states [51]. We know what data the FBI are collecting, and a pilot study report has been released, but until participation rises to 60% and 80% nationwide, actual use-of-force data will be held back, presumably to prevent inaccurate conclusions from being drawn from nonrepresentative data [52], [53].

Results

Fatal Force, Phase 1 The dataset has specific criteria for inclusion that are based on the killing of Michael Brown in Ferguson, MO, USA, in 2014. Each person included in the data system was a civilian who was shot and killed by a police officer in the line of duty. Any killing by a police officer who was not on duty or by a police officer who was on duty but did not shoot the victim or the killing of a person who was in custody is not recorded in the database. For instance, although George Floyd's death initiated a worldwide wave of protests in 2020, his killing is not recorded in Fatal Force because he was suffocated and not shot. The reason for choosing this specific incident—Michael Browns murder—as the basis of the Fatal Force database is that this killing was a catalyst for the Black Lives Matter movement as it is today [22].

This system can be construed as person-based or as incident-based; because of the subject of the data, each item in the database comprises one person killed by police. In other words, each person killed is one incident. The name of the person killed is one of the first categories, second only to the ID number, signaling its relative importance—the system privileges information about both the victim and the circumstances of the killing. Table III shows the categories in this dataset, in order.

TABLE III
Categories in the Post Fatal Force Dataset [47]

Data Categories
ID number
Name
Date of death
Manner of death
Armed status
Age
Gender
Race
City
State
Signs of mental illness
Threat level
Fleeing status
Body camera
Coordinates (expressed in two variables, latitude and longitude)
Accuracy of geocoding

All categories are focused on the victim of the fatal police shooting in question, with the exception of the body camera and the coordinates, which describe the existence of footage of the killing and the placement of the killing on a mapped data visualization, respectively. Order of categories matters in data systems; the order dictates which points are listed first and which points are seen as being related to which other points. Importance can also sometimes be inferred from primacy since the most crucial information tends to be placed first for sorting and searching purposes. Name and date taking prime placement indicate their importance to the incident-based nature of the database; manner of death and armed status are also near the front of the list, suggesting that they are key pieces of information in accordance with common claims of misuse of force in discussions of police brutality in the US. In this way, these categories appear to be directly influenced by the motivations for creating the database and by the medium much of the data come from (news reports).

The relationship between categories in the dataset itself seems to be different than in the interactive database online, where information is organized in a new order to create a narrative similar to one that readers might find in the headline and lede of a news article (shown in Fig. 1) [46]. In the interactive database on the Fatal Force website, each row in the dataset is displayed on a card which lists the victim's name, age, race, and gender before indicating whether the victim was armed (and with what), the manner of death and date, and where the victim was killed (in a house, in a parking lot, etc., and then city and state). Whether the victim displayed signs of mental illness, whether a body cam recording exists, and whether the victim was

Fig. 1. Example of Fatal Force interface entry.

fleeing (and how) are all indicated below the headline of the card in brief bullets. Finally, the card lists one or more sources for information about the killing. This interactive database serves two functions: first, it allows users to filter and search through the data using any of the variables; second, it collects the individual bits of information into a clear narrative with a consistent format [46].

Data are stored with *The Washington Post* and also publicly updated on GitHub, a popular platform for sharing digital materials like web code, apps, code snippets in various languages, and datasets. GitHub also hosts a description of each variable in the dataset along with its potential options [47]. For instance, the "gender" category can be filled by options "male," "female," and "unknown"—the description also specifies that the category includes information about the gender the victim identified with, if that is not the same as information in official reports (such as with transgender victims), but notably there is no option for a nonbinary or even "other" gender marker.

Data that are updated as new reports are published, but no data are deleted for missing information; the categories with no information are simply left blank. There are several entries in the database with no name, for instance, either because an investigation is still ongoing or because the person has yet to be identified to the media. Data are pulled from news reports, law enforcement reports, and social media reports, as well as from independent databases that record information about police killings in the US. The database's methodology also indicates that journalists working on the Fatal Force project conducted their own reporting in several cases; so data may be coming from or otherwise corroborated by witnesses, relatives, or law enforcement, but the sources of information for individual entries are not listed [22].

Fatal Force, Phase 2 The Fatal Force project extends to a larger reporting agenda at the *Post*, where journalists are working to investigate and bring to light the various stories around police brutality. Narratives included in this case take, in general, one of two forms—explicitly story-driven narratives focusing on individuals affected by fatal police shootings such as family members, or implicit narratives that work on an abstract level without specific characters. Most articles use both for differing effects; story-driven narratives tend to illustrate points and bring human emotion or illuminate particular ways in which a blanket policy fails individual people. Abstract narratives, on the other hand, are often used to create a more authoritative voice, especially when making claims that critique powerful institutions.

Most of the longer investigative pieces are in large part composed of explicit story-driven narratives, relying for their poignancy on detailed accounts of people being fatally shot. Temporality and sociality tend to be the dominant narrative elements. In reporting 16-year-old Robert Dentmond's story, authors emphasize Dentmond, his age and familial circumstances, and his relationship with his girlfriend. They say that Dentmond had a generally poor social support system, emphasize the importance of his girlfriend, and then rely on that setup to explain Dentmond's actions on the night that he was killed [54]. Dentmond even briefly gets his own voice as authors report a line from a Facebook message he sent before he called the police on himself, expressing that he wanted the police to shoot him. Temporality takes over as the authors signal the date, the time of Dentmond's call to police, the time between the call and the police's arrival, Dentmond's advancing on them with a fake weapon, and police opening fire. These times are expressed in exact terms (March 20, 10:07 P.M., 24 min). Less exactly, authors use temporality as an implicit factor—these instances are usually used more for emotional impact. In one line, authors write, "After 24 minutes, Dentmond leveled …," and in the very next sentence, they follow that exact time marker with an implicit one: "Nine officers opened fire and didn't stop *until the gun fell from his hand*" (emphasis mine) [54]. This mix of explicit and implicit time markers works on two levels. Explicit markers function like details in a stereotypical police report or court documents, and implicit markers function more cinematically, inviting the reader to empathize and embed themselves in the scene, imagining how long officers fired, how many bullets were fired, and how many times Dentmond was hit.

Implicit narrative in the Fatal Force data stories is usually couched between explicit narratives—in this example, several explicit narratives are covered in a sentence or two, to show a pattern across unique incidents, and each explicit narrative gets marked by place and sociality, though temporality is mostly elided. Implicit narrative in one exemplar piece on shootings of unarmed Black men makes some key moves that are echoed in other examples in the data system [55]. The first is counting—authors explicitly name and emphasize the number of the dead. This continues as authors begin to describe the frequency of incidents (every nine days, three shootings in two weeks). Counting and frequency, both used to communicate "a lot, too much," are followed by a twist—these incidents are just "a small fraction" of the total—and then a resolution: Proportionally, unarmed Black men are seven times more likely to be shot and killed by police than unarmed white men [55]. This implicit narrative does not seem like a story compared to what follows—the story of Christian Taylor, a Black, unarmed college football player shot by police—because Taylor's story has characters, a plot. But implicit narratives still tell stories that can be constructed; this example's story reads something like

> unarmed Black men are shot and killed by police with alarming frequency, and though it looks like other groups have it worse 'by the numbers,' the truth about how much more danger Black men face than other races is even worse than you thought. [55]

The key difference in commonplaces between implicit and explicit narratives in these data stories is that implicit narratives rarely include elements of place—they focus specifically on sociality and, to a slightly lesser degree, on temporality. Explicit narratives typically feature place heavily, even if only to ground the story—and although they also include sociality and temporality like implicit narratives, the ways that these commonplaces manifest tend to diverge. Implicit narratives use sociality in more global or large-scale ways, describing groups rather than individuals, while explicit narratives use it to communicate information about individual people, their relatives, and their relationship with the world and people around them. These rhetorical moves match up with the dichotomy of purpose between these two narrative types (implicit functions to critique powerful institutions; explicit functions to humanize and facilitate empathy).

In terms of narrative commonplaces, the Fatal Force database mostly focuses on place and sociality. Categories here describe a person—the victim—and how they were being perceived by those around them (i.e., whether someone suffered from mental health issues, whether someone was armed and how, what someone's race and gender were, etc.). Place is less key in the data themselves, mostly present in mapping the location of the death. Fatal Force data stories, however, add temporality back in—not just the date of the death, which is already in the database, but the series of events and their temporal relationship to one another. Stories can convey how fast something occurred, or how slow, as in the case of Dentmond's story described earlier, which specifically noted that Dentmond hesitated for 24 minutes between officers telling him to drop his (replica) weapon and his advancing on them with it [54].

The database flattens the story of each line, each shooting, into one moment—as the cards in the web-based interactive database all read, a person of a given race and gender, armed or unarmed, was shot in a place on a date. Data stories give these events their time back. Data stories also typically include a great deal of more information about the people who were shot and often even include the words of those who loved them. This is a clear move to embed shooting victims in the world and in social networks, both to humanize these data and to build empathy.

The Fatal Force database is very simple, but searchability, importance of names, and filterability are key factors that work to humanize the data and reinforce that each entry in the database is the story of a life lost. The presence of links out to news stories, as well, concretizes each card in the interactive web database into an event. Their chronological order serves this same function, too, playing into temporality. Given the exigence of the database, with Michael Brown's murder in Ferguson, MO, USA, as one of the catalysts for the Movement for Black Lives, it is perhaps surprising that the database itself does not focus heavily on Black Lives Matter, rather keeping the focus on police killings more generally [22]. However, most of the data stories surrounding the database and comprising the data system add information that makes up for that lack by putting the tallied killings in the context of the country's demographic breakdown.

FBI Use-of-Force The FBI began its National Use-of-Force Data Collection with a pilot study in 2017, two years after the start of the Fatal Force database. The Use-of-Force database was not used in earnest, however, until 2019 [53]. The FBI's Use-of-Force data, like Fatal Force, are collected in response to growing social awareness of police use and misuse of force; unlike Fatal Force, however, the FBI's data do not name any of the high-profile killings or protests in the last decade nor are Black Lives Matter or other organizations focusing on police brutality mentioned at all. Rather, a summary of the FBI's 2017 pilot study states that the Use-of-Force collection was begun "at the request of major law enforcement organizations" to "promote more open and informed conversations" about police use of force, with no further information about exigence or purpose. This database collects the information shown in Table IV, according to Use-of-Force FAQ.

Currently, no actual results have been reported—the FBI has set participation benchmarks that must be reached before various levels of data will be made available to the public, and, currently, just one of those benchmarks has been met (just over 40% of officers are represented; so national participation data have been released—essentially communicating how many agencies have submitted incident reports, with zero reports meaning no use of force occurred or that the agency is not currently participating). Participation is voluntary for LEAs. Currently, only about 5000 agencies out of about 18,000 are participating in Use-of-Force data collection [51].

The FBI's Use-of-Force database asks for more detail about a given incident than Fatal Force does—likely a result of access since the police have more access to information about an incident than journalists do. However, in the pilot study report, authors note that victim injuries should be entered "from the officer's perspective," rather than from medical reports [53]. Similarly, whether the victim was armed and whether the victim appeared to be suffering from mental health issues or other impairments are both based on the officer's perception at the scene. Because no data have been released from this project yet, it is unclear whether information can be added after the fact or to represent specific categories found in Fatal Force, such as whether the victim had a replica weapon. It is likely that this "officer's perspective" requirement is related to either victim privacy or expediency for law enforcement, but all the same, the information entered following this guideline is not necessarily an accurate reflection of a victim's injuries.

The FBI database also includes far more information regarding the officers involved—this database is linked to other FBI data collection efforts that include that data; so including officer data makes sense, but it also betrays the database's focus. Fatal Force is focused almost entirely on victims. This focus is reinforced again by one of the repeated disclaimers on the FBI's database: The Use-of-Force database "does not assess or report whether officers followed their department's policy or acted lawfully" but merely records that a use of force occurred [52].

The difference between the National Use-of-Force Data Collection and Fatal Force is that Fatal Force was created to address a specific issue and takes its cues from one murder—Michael Brown's—so its inclusion criteria and design are therefore very specific but can answer questions that FBI data are not meant for. The FBI's use-of-force database is made for specifically working on a large-scale, abstract level—thinner data, even though there are more categories that could potentially add more nuance—to answer various questions about what types of force are used by police, where, and under what circumstances. The participation benchmarks for data release, at least for now, serve to increase that abstraction; users can see how many participating agencies reported that use of force either occurred or did not, but no more detail than that. As Jones argues, abstraction rather than particularity often works to minimize or make invisible the experiences of marginalized people [41]; in this case, however, the abstraction created by these benchmarks also serves to protect and reinforce privilege for law enforcement.

TABLE IV
INFORMATION COLLECTED IN FBI USE OF FORCE DATABASE [52]

Incident Information	Subject Information	Officer Information
Date and time	Age, sex, race, ethnicity, height, and weight	Age, sex, race, ethnicity, height, and weight
Total number of officers who applied force	Injury/death of subject	Years of service in law enforcement
Number of officers from reporting agency who applied force	Type of force used	Was the officer a full-time employee?
Location	Did the subject direct a threat to the officer or another person?	Was the officer on duty?
Location type (street, business, home, etc.)	Did the subject resist?	Did the officer discharge a firearm?
Did the officer(s) approach the subjects?	Types of resistance or weapon involvement (threats, active aggression, firearms, etc.)	Was the officer injured?
Was it an ambush incident?	Did the subject have a known or apparent impairment, such as mental health condition or being under the influence of drugs or alcohol?	If so, what was the officer's injury type?
Was a supervisor or senior officer consulted during the incident?	Was the subject believed to have a weapon?	
Reason for initial contact (routine patrol, traffic stop, etc.)		
If the initial contact was due to unlawful activity, what was the most serious offense the individual was suspected of?		
If applicable, the reporting agency will include the National Incident-Based Reporting System record or local incident number of the report detailing criminal incident information on the subject and/or assault or homicide of a law enforcement officer.		
If the incident involved multiple agencies, the reporting agency should provide case numbers for the other agencies' incident reports.		

The perspectives and goals of both police use-of-force data systems diverge in that one is highly specific about its stance and goals, while the other acknowledges very little context in regard to purpose. Fatal Force, as an entire system including database and data stories, was designed to address social inequities as they manifest in policing and police violence, from the perspective of citizens, community members, and allies. The FBI's use-of-force database, which has not yet released much information to the public at all, was designed "for the purpose of compiling national statistics" [53, p. 2] or "to promote more open and informed conversations regarding law enforcement use of force in the United States" [53, p. 1], from the perspective of law enforcement institutions and police officers [53].

It seems clear that the Fatal Force system is better geared toward users encountering the system and looking for answers. The Fatal Force database is augmented with qualitative data in the form of data visualization, journalism, and investigative reporting that goes along with the database—these data stories draw from the trends that the database can show, and adding context that acknowledges and centers social justice issues. For example, Fatal Force clearly shows that more White

Americans are killed by police than Black Americans, just by the numbers—but Fatal Force also puts that data in context of US demographics to show that Black Americans are disproportionately harmed by police shootings, debunking anti-#BlackLivesMatter arguments. The FBI's use-of-force database is meant to answer similar questions and focuses on the same issue, though with a wider inclusion (all use of force, and not just fatal police shootings with on duty officers), but unlike Fatal Force, the officer's perspective is the one that matters—not the victim's.

CONCLUSION

I have argued previously that data systems themselves are rhetorical; they make arguments, even when they are intended by their designers to be "objective" or "impartial" or to simply report information without taking a stance or making a claim [32]. These arguments are implicit in a data system from the beginning because the system must name categories of data that it will collect and list possible options for that data (i.e., the "armed" category in Fatal Force has four possible options, which allow for more nuance than a "yes/no" answer but less than a free text entry). In so doing, the system's designers place value on certain kinds of information and argue that the situation being described by the data in that system is accurately and appropriately captured by those categories and those entry options. Designing a data system means predicting what users will want to ask of the data and working to meet those needs. And as technical and professional communicators invested in enacting social justice, we need to be in the room asking whose questions do these data answer? Whose reality is captured by these data? And does this system uphold Taylor's principles of data justice: "visibility, digital (dis)engagement, and countering data-driven discrimination" [13, p. 83]?

This work opens pathways for future scholarship as well. First, there is a need for further research on the rhetorical dimensions of law enforcement's technical communication work. Some scholarship on technical communication and policing exists but not a great deal, and as I have shown in this article, the invisible rhetorical work going on in this area is a crucial social justice concern. Second, scholars should examine many more examples of data systems in crime and policing. Fatal Force is an interesting and instructive case, but it is only one case—there is still a great deal more to learn. Finally, although my current work with unblackboxing and data system design deals with crime and policing, this is far from the only social justice issue that is described and dealt with using data. The unblackboxing methodology would benefit a great deal from testing on data systems of different types and with different audiences and purposes.

REFERENCES

[1] N. N. Jones, K. R. Moore, and R. Walton, "Disrupting the past to disrupt the future: An antenarrative of technical communication," *Tech. Commun. Quart.*, vol. 25, no. 4, pp. 211–229, Oct. 2016, doi: 10.1080/10572252.2016.1224655.
[2] G. Y. Agboka, "Participatory localization: A social justice approach to navigating unenfranchised/disenfranchised cultural sites," *Tech. Commun. Quart.*, vol. 22, no. 1, pp. 28–49, Jan. 2013, doi: 10.1080/10572252.2013.730966.
[3] E. B. Carlson, "Embracing a metic lens for community-based participatory research in technical communication," *Tech. Commun. Quart.*, vol. 29, no. 4, pp. 392–410, Oct. 2020, doi: 10.1080/10572252.2020.1789745.
[4] N. N. Jones, "The technical communicator as advocate: Integrating a social justice approach in technical communication," *J. Tech. Writing Commun.*, vol. 46, no. 3, pp. 342–361, Apr. 2016, doi: 10.1177/0047281616639472.
[5] C. Shelton, "Shifting out of neutral: Centering difference, bias, and social justice in a business writing course," *Tech. Commun. Quart.*, vol. 29, no. 1, pp. 18–32, Jan. 2020, doi: 10.1080/10572252.2019.1640287.
[6] K. Lum and W. Isaac, "To predict and serve?," *Significance*, vol. 13, no. 5, pp. 14–19, 2016, doi: 10.1111/j.1740-9713.2016.00960.x.
[7] V. Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York, NY, USA: St. Martin's Press, 2017.
[8] S. U. Noble, *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York, NY, USA: New York Univ. Press, 2018.
[9] R. Benjamin, *Race After Technology: Abolitionist Tools For the New Jim Code*, 1st ed. Medford, MA, USA: Polity, 2019.
[10] M. Bal, "The point of narratology," *Poetics Today*, vol. 11, no. 4, pp. 727–753, 1990, doi: 10.2307/1773075.

[11] C. Spinuzzi, *Tracing Genres Through Organizations: A Sociocultural Approach to Information Design*. Cambridge, MA, USA: MIT Press, 2003.
[12] R. Walton, K. Moore, and N. Jones, *Technical Communication After the Social Justice Turn: Building Coalitions for Action*. London, UK: Routledge, 2019.
[13] L. Taylor, "What is data justice: The case for connecting digital rights and freedoms globally," *Big Data Soc.*, vol. 4, no. 2, pp. 1–14, 2017, doi: 10.1177/2053951717736335.
[14] I. Dorpenyo, "Unblackboxing' technology through the rhetoric of technical communication: Biometric technology and Ghana's election," Ph.D. dissertation, Dept. English, Michigan Technol. Univ., Houghton, MI, USA, 2016. [Online]. Available: https://digitalcommons.mtu.edu/etdr/125
[15] B. Benbouzid, "To predict and to manage. Predictive policing in the United States," *Big Data Soc.*, vol. 6, no. 1, 2019, Art. no. 205395171986170, doi: 10.1177/2053951719861703.
[16] B. Brucato, "Big data and the new transparency: Measuring and representing police killings," *Big Data Soc.*, vol. 4, no. 1, 2017, Art. no. 205395171769633, doi: 10.1177/2053951717696332.
[17] L. Seawright, *Genre of Power: Police Report Writers and Readers in the Justice System*. Urbana, IL, USA: NCTE, 2017.
[18] H. Yu and N. Monas, "Recreating the scene: An investigation of police report writing," *J. Tech. Writing Commun.*, vol. 50, no. 1, pp. 35–55, Jan. 2020, doi: 10.1177/0047281618812441.
[19] Federal Bureau of Investigation, "NIBRS," 2019. Accessed: Dec. 13, 2019. [Online]. Available: https://www.fbi.gov/services/cjis/ucr/nibrs
[20] Federal Bureau of Investigation, "Summary of the uniform crime reporting (UCR) program," US Department of Justice, 2014. Accessed: Jul. 28, 2020. [Online]. Available: https://ucr.fbi.gov/crime-in-the-u.s/2013/crime-in-the-u.s.-2013/about-ucr/aboutucrmain_final.pdf
[21] D. Banks, J. Hendrix, M. Hickman, and T. Kyckelhahn, *National Sources of Law Enforcement Employment Data*. New York, NY, USA: US Department of Justice—Bureau of Justice Statistics, 2016.
[22] J. Tate *et al.*, "How the *Washington Post* is examining police shootings in the United States," in *The Washington Post*. Tampa, FL, USA: Fred Ryan, 2016.
[23] A. Lucchesi and A. Echo-Hawk, *Missing and Murdered Indigenous Women & Girls*. Seattle, WA, USA: Urban Indian Health Institute, 2019. Accessed: Feb. 05, 2020. [Online]. Available: https://www.uihi.org/resources/missing-and-murdered-indigenous-women-girls/
[24] B. Latour, *Pandora's Hope: Essays on the Reality of Science Studies*. Cambridge, MA, USA: Harvard Univ. Press, 1999.
[25] C. M. Dalton and J. Thatcher, "What does a critical data studies look like, and why do we care? Seven points for a critical approach to 'big data'," *Soc. Space - Environ. Plan. D*, May 28, 2014. Accessed: Feb. 25, 2020. [Online]. Available: https://web.archive.org/web/20140528052630/http://societyandspace.com/material/commentaries/craig-dalton-and-jim-thatcher-what-does-a-critical-data-studies-look-like-and-why-do-we-care-seven-points-for-a-critical-approach-to-big-data/
[26] d. boyd and K. Crawford, "Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon," *Inf. Commun. Soc.*, vol. 15, no. 5, pp. 662–679, Jun. 2012, doi: 10.1080/1369118X.2012.678878.
[27] L. Gitelman, *Raw Data is an Oxymoron*. Cambridge, MA, USA: The MIT Press, 2013.
[28] A. Beveridge, "Writing through big data: New challenges and possibilities for data-driven arguments," *Comp. Forum*, vol. 37, no. Fall, 2017, Art. no. 2017, Accessed: Jan. 16, 2020. [Online]. Available: https://compositionforum.com/issue/37/big-data.php
[29] J. R. Gallagher, "Writing for algorithmic audiences," *Comput. Comp.*, vol. 45, pp. 25–35, Sep. 2017, doi: 10.1016/j.compcom.2017.06.002.
[30] C. Lanius and G. S. Hubbell, "The new data: Argumentation amid, on, with, and in data," in *Theorizing Digital Rhetoric*, A. Hess and A. L. Davisson, Eds. London, UK: Taylor & Francis Group, 2017, pp. 126–139.
[31] J. Reyman, "The rhetorical agency of algorithms," in *Theorizing Digital Rhetoric*, A. Hess and A. L. Davisson, Eds. London, UK: Taylor & Francis Group, 2017, pp. 112–125.
[32] R. Atherton, "'Missing/unspecified': Demographic data visualization during the COVID-19 pandemic," *J. Bus. Tech. Commun.*, vol. 35, no. 1, pp. 80–87, Jan. 2021, doi: 10.1177/1050651920957982.
[33] S. S. Graham, S.-Y. Kim, D. M. DeVasto, and W. Keith, "Statistical genre analysis: Toward big data methodologies in technical communication," *Tech. Commun. Quart.*, vol. 24, no. 1, pp. 70–104, Jan. 2015, doi: 10.1080/10572252.2015.975955.
[34] L. Melonçon and E. Warner, "Data visualizations: A literature review and opportunities for technical and professional communication," in *Proc. IEEE Int. Prof. Commun.*, 2017, pp. 1–9.
[35] T. C. Overmyer, "UX methods in the data lab: Arguing for validity," in *Proc. 37th ACM Int. Conf. Design Commun.*, 2019, pp. 1–6, doi: 10.1145/3328020.3353954.
[36] J. Wolfe, "Teaching students to focus on the data in data visualization," *J. Bus. Tech. Commun.*, vol. 29, no. 3, pp. 344–359, Jul. 2015, doi: 10.1177/1050651915573944.
[37] T. Bucher, *If…Then: Algorithmic Power and Politics*. Oxford, NY, USA: Oxford Univ. Press, 2018.
[38] C. Koopman, *How We Became Our Data: A Genealogy of the Informational Person*. Chicago, IL, USA: Univ. Chicago Press, 2019.
[39] S. E. Merry, *The Seductions of Quantification*. Chicago, IL, USA: Univ. Chicago Press, 2016.
[40] D. J. Clandinin, *Engaging in Narrative Inquiry*. New York, NY, USA: Routledge, 2016.
[41] N. N. Jones, "Narrative inquiry in human-centered design: Examining silence and voice to promote social justice in design scenarios," *J. Tech. Writing Commun.*, vol. 46, no. 4, pp. 471–492, Oct. 2016, doi: 10.1177/0047281616653489.

[42] N. N. Jones, "Rhetorical narratives of black entrepreneurs: The business of race, agency, and cultural empowerment," *J. Bus. Tech. Commun.*, vol. 31, no. 3, pp. 319–349, Jul. 2017, doi: 10.1177/1050651917695540.
[43] N. N. Jones, "Coalitional learning in the contact zones: Inclusion and narrative inquiry in technical communication and composition studies," *Coll. English*, vol. 82, no. 5, pp. 515–526, May 2020.
[44] G. C. Bowker and S. L. Star, *Sorting Things Out: Classification and Its Consequences*. New York, NY, USA: MIT Press, 1999.
[45] P. Sullivan, "Beckon, encounter, experience: The danger of control and the promise of encounters in the study of user experience," in *Rhetoric and Experience Architecture*, L. Potts and M. Salvo, Eds. Anderson, SC, USA: Parlor Press, 2018, pp. 17–40.
[46] "Fatal Force: Police shootings database," in *The Washington Post*. Tampa, FL, USA: Fred Ryan, Accessed: Oct. 15, 2020. [Online]. Available: https://www.washingtonpost.com/graphics/investigations/police-shootings-database/
[47] *The Washington Post*. (2021). Washington Post/data-police-shootings. San Francisco, CA, USA. [Online]. Available: https://github.com/washingtonpost/data-police-shootings
[48] "Police shootings 2016 database," in *The Washington Post*. Tampa, FL, USA: Fred Ryan, Accessed: Mar. 10, 2021. [Online]. Available: https://www.washingtonpost.com/graphics/national/police-shootings-2016/
[49] "Police shootings 2017 database," in *The Washington Post*. Tampa, FL, USA: Fred Ryan, Accessed: Mar. 10, 2021. [Online]. Available:https://www.washingtonpost.com/graphics/national/police-shootings-2017/
[50] "Fatal Force: 2019 police shootings database," in *The Washington Post*. Tampa, FL, USA: Fred Ryan, Accessed: Mar. 10, 2021. [Online]. Available: https://www.washingtonpost.com/graphics/2019/national/police-shootings-2019/
[51] Federal Bureau of Investigation, "Use of force," *Crime Data Explorer*, Accessed: Mar. 14, 2021. [Online]. Available: https://crime-data-explorer.fr.cloud.gov/officers/national/united-states/uof
[52] Federal Bureau of Investigation, "National use-of-force data collection," Accessed: Mar. 15, 2021. [Online]. Available: https://www.fbi.gov/services/cjis/ucr/use-of-force
[53] Criminal Justice Information Services Division, "National use-of-force data collection pilot study summary," *FBI File*, Dec. 2018, Accessed: Mar. 15, 2021. [Online]. Available: https://www.fbi.gov/file-repository/ucr/national-use-of-force-data-collection-pilot-study-121018.pdf/view
[54] J. Sullivan, J. Jenkins, J. Tate, S. Courtney, and J. Houston, "In two years, police killed 86 people brandishing guns that look real—But aren't," in *The Washington Post*. Tampa, FL, USA: Fred Ryan, 2016.
[55] "The 24 unarmed black men who have died in 2015," *The Washington Post*, Tampa, FL, USA: Fred Ryan, Accessed: Mar. 10, 2021. [Online]. Available: http://www.washingtonpost.com/sf/national/2015/08/08/black-and-unarmed/

Rachel Atherton received the B.A. degree in English and the B.A. degree in German from Southern Illinois University, Carbondale, IL, USA, in 2015, and the master's degree in Rhetoric and Composition in 2017 from Purdue University, West Lafayette, IN, USA, where she is currently working toward the doctoral degree with the English Department. Her research interests include rhetorics of data and data system design. Her work has been published in the *Journal of Business and Technical Communication*, *Computers and Composition*, and *Technical Communication*.