ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/312187007

To Post or Not to Post: Online Discussion of Gun Permit Mapping and the Development of Ethical Standards in Data Journalism

Article *in* Journalism & Mass Communication Quarterly · January 2017

DOI: 10.1177/1077699016684796

CITATIONS
9

READS
1,165

3 authors, including:



Mohammad Yousuf
University of New Mexico
26 PUBLICATIONS **157** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Data Journalism View project

News Engagement View project

All content following this page was uploaded by Mohammad Yousuf on 12 June 2018.

The user has requested enhancement of the downloaded file.

Check for updates

Journalism & Mass Communication Quarterly
2017, Vol. 94(1) 168–188
© 2017 AEJMC
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/1077699016684796
http://journals.sagepub.com/home/jmq
SAGE

To Post or Not to Post: Online Discussion of Gun Permit Mapping and the Development of Ethical Standards in Data Journalism

David Craig[1], Stan Ketterer[2], and Mohammad Yousuf[1]

Abstract

This study investigates the ethical dimensions of data journalism, an area of growing public importance in journalistic practice, by examining journalists' discussion related to controversies over handling of gun permit data. Three discussion threads on the listserv of the National Institute for Computer-Assisted Reporting were analyzed. Frames were freedom versus responsibility, privacy and verification, consequences, and alternatives. The findings highlight the benefits of pooling the practical wisdom of participants in an evolving practice. They also suggest recommendations for evaluating the ethics of data journalism.

Keywords

data journalism, ethics, practice, technology

Data analysis and visualization is an increasingly important part of 21st-century journalism. Although the terms and definitions differ and encompass a range of approaches (Coddington, 2015), the common goal of what professionals now commonly call "data journalism" is, as Aron Pilhofer of the *New York Times* put it, "providing information and analysis to help inform us all about important issues of the day" (Gray, Chambers, & Bounegru, 2012). Internet technology allows journalists to post data to media websites, making it easier for the public to access the information and to assess it. Data journalists can also connect that information to Google Maps, making it even more accessible.

[1]The University of Oklahoma, Norman, USA
[2]Oklahoma State University–Stillwater, USA

Corresponding Author:
David Craig, Gaylord College of Journalism and Mass Communication, The University of Oklahoma, 395 West Lindsey Street, Room 3000, Norman, OK 73019, USA.
Email: dcraig@ou.edu

These changes in journalism are occurring during an unprecedented increase in the volume of data collected by private industry and government, including personal data, and an equally unprecedented expansion of computing power and storage necessary to process and analyze data in new ways (Carter & Lee, 2016; Mayer-Schonberger & Cukier, 2013; Podesta, Pritzke, Moniz, Holdren, & Zients, 2014). As a White House review of big data (Podesta et al., 2014) conducted in the wake of the National Security Agency (NSA) scandal (Liptak & Peters, 2013) pointed out, big data are a two-edged sword: "Like other transformative factors of production, big data generates value differently for individuals, organizations and society. While many applications are unequivocally beneficial, some of its uses impact privacy and other core values of fairness, equality, and autonomy" (p. 48).

Less than 2 weeks after the shooting deaths of 20 children and six staff members at Sandy Hook Elementary School in Newtown, Connecticut (Jonsson, 2012), a journalist at the *Journal News* in White Plains, New York, wrote a story about the possible expansion of publicly accessible gun permit data that was accompanied by three online maps with the locations of gun permit holders (Worley, 2012). The clickable maps of a two-county area in the New York suburbs also included the names and addresses of the gun permit holders. The detailed maps with personal information prompted a public outcry both locally and nationally, mainly involving privacy and safety concerns. They also prompted discussion among journalists, including participants in NICAR-L, the listserv of the National Institute for Computer-Assisted Reporting.

The purpose of this article is to explore the ethical dimensions of data journalism by examining these journalists' discussion of this controversy and related issues that preceded it. The study's importance goes beyond this case because the questions it raises about the ethics of data journalism are ongoing ones relevant to journalists' work in analyzing and visualizing public data worldwide. From a scholarly standpoint, it is important because it contributes to the still-small amount of literature examining the ethics of data journalism. The findings also point to the potential value of online interaction among professionals for the evolution of standards of excellence in journalism, based on philosopher Alasdair MacIntyre's (2007) concept of a practice.

Literature Review

This literature review will examine scholarly and professional discussion of the meaning and potential of data journalism, studies related to the ethical issues and challenges of data journalism, and literature on excellence in journalism and the evolution of excellence.

Meaning and Potential of Data Journalism

Data journalism does not have a uniform definition and has been used in overlapping ways with the older terms *computer-assisted reporting* (Garrison, 1998; Houston, 1999) and *computational journalism* (e.g., Cohen, Hamilton, & Turner, 2011; Daniel & Flew, 2010). Anderson (2011) defined computer-assisted reporting as "forms of

algorithmic, social scientific and mathematical forms of newswork" (Abstract). Hamilton and Turner (2009) described computational journalism as "the combination of algorithms, data, and knowledge from the social sciences to supplement the accountability function of journalism" (p. 2). The social science element echoes the work of Philip Meyer, who pioneered "precision journalism" using scientific methods to analyze data for journalistic purposes (Meyer, 2002).

The term *data journalism* became commonly used among journalists in place of computer-assisted reporting following a January 2011 discussion on NICAR-L. Listers agreed the term *computer-assisted reporting* no longer fully described the practice. However, Coddington (2015) argues,

> The three practices, then, are distinct quantitatively oriented journalistic forms: CAR is rooted in social science methods and the deliberate style and public-affairs orientation of investigative journalism, data journalism is characterized by its participatory openness and cross-field hybridity, and computational journalism is focused on the application of the processes of abstraction and automation to information. (p. 337)

Although the case in this article involved discussion by a group identified with computer-assisted reporting, it also involved public participation because of the online presentation of data for exploration through interactive mapping.

The fundamental difference between data and traditional journalism is the analysis involves digitized data. Journalists traditionally conducted investigations using paper records and analyzed the resulting data using mathematical techniques. With the advent of personal computers and the increased availability of government databases, the analysis is now often conducted using statistical programs that require specialized knowledge to interpret. As Aucoin (2005) pointed out regarding Meyer's pioneering investigative work, "But it was the computer analysis that set their study apart from previous journalistic efforts and set a new standard for investigative journalists" (p. 196). Computer-assisted data analysis often yields a more precise idea for a story that is still reported and written using traditional journalistic techniques. The same digitization allows the public to access the data for online exploration through easy-to-use interfaces and data visualizations, such as clickable maps. Recognizing the importance of such developments, journalism schools are paying increasing attention to the teaching of data journalism (Krueger, 2014).

Daniel and Flew (2010) mentioned four potential contributions of data journalism. It can:

1. Increase the depth of quality original investigative reporting;
2. Differentiate a digital masthead from competitors;
3. Accelerate the process of journalism from news source to delivery; and
4. Provide a factual basis for analysis, which may minimise the risk of incorrect reporting and counter the influence of public relations. (p. 3).

Houston (2010) focused on visualization of "large government data sets in order to detect favoritism, incompetence, and corruption" (p. 51). Flew, Spurgeon, Daniel, and

Swift (2011) stated using large databases in journalism requires collaboration between journalists and computer scientists. They implied data journalism could contribute to a story's accuracy, originality, and reliability.

Gray et al. (2012) asked some proponents and practitioners of data journalism about its importance. The issues include filtering data by presenting the important and relevant information, presenting information so readers can verify it and replicate the methodology, monitoring trends, providing independent interpretations of official information, and scrutinizing and holding the powerful accountable. Data journalism also may help investigative journalism survive when support for such journalism is dwindling (Anderson, 2011).

Taken together, this literature on the meaning and potential of data journalism—though it does not discuss ethics directly—implies a role for data journalists in maximizing the ethical potential of journalism to provide truthful, in-depth reporting.

Ethical Issues and Challenges of Data Journalism

Although little scholarly work has systematically addressed the ethical issues and challenges in data journalism, ethical concerns have emerged in a few academic discussions and in other sources. These concerns cluster around three overlapping issues: truthfulness, harm, and privacy.

"It's not a question of whether your data is dirty, but how dirty" is a common expression among data journalists (Vallance-Jones & McKie, 2009) that points to problems of truthfulness and harm. It indicates a recognition that all data have flaws that must be corrected or at least understood if they cannot be corrected. Data journalists must screen the data to find these problems (Garrison, 1998) and clean them to eliminate them (Vallance-Jones & McKie, 2009). If not, the subsequent analysis will be flawed. Messner and Garrison (2007) reviewed the scholarly and trade literature regarding uses and problems with databases in reporting. Problems included data import problems, incomplete data, missing data, inaccurate data, lack of verification, flawed documentation, inconsistent or inaccurate coding, misspellings, inappropriate uses of data, and faulty data analysis. Other problems with data include duplicate records, outdated data, and lack of precautions to protect unnamed sources of data. For important stories, the data journalist's methodology and problems with data are usually reported in a "nerd box" so the public can assess them. Because data journalists are essentially creating news through their analysis, these practices for dealing with data problems must be followed or the story or online data can mislead the public. Bradshaw (2015) and Cairo (2016) noted the importance of providing context in presentation of data, including visualizations, to avoid misrepresentation.

Thus, problems of truthfulness can lead to harm. Daniel, Flew, and Spurgeon (2010) found computational journalism directly involved ethical issues such as double falsehoods, misrepresentations of results, and presentation of incomplete data that might harm people. As Meyer (2002) noted, "Numbers are like fire. They can be used for good or for ill. When misused, they can create illusions of certitude and importance that render us irrational" (p. 19). When numbers are attached to a problem, it might

make the problem look more important than it originally is (Boyle, 2001). As boyd and Crawford (2012) said of big data, "what it quantifies does not necessarily have a closer claim on objective truth" (p. 667). Lewis and Westlund (2015) noted problems with public data conflict with many data journalists' desires to share data sets with the public.

Moreover, privacy concerns have increased with the explosion of data gathering and analysis capabilities in recent years. These carry accompanying concerns about harm. Mayer-Schonberger and Cukier (2013) point out government regulation has not kept up with private industry's ability to aggregate and process information for its own ends, including detailed information about private citizens.

A White House report in the wake of the NSA scandal stated,

> In addition to creating tremendous social good, big data in the hands of government and the private sector can cause many kinds of harms. These harms range from tangible and material harms, such as financial loss, to less tangible harms, such as intrusion into private life and reputational damage. (Podesta et al., 2014, p. 51)

Fairfield and Shtein (2014) explored the ethical problems of using data science as reporting and newsgathering tools. They argued the nature of big data challenges the traditional way of framing ethical principles: "Big data technologies increase costs of compliance with traditional ethical values, while steeply lowering costs of invasion of privacy" (p. 50). Daniel et al. (2010) suggested private data collected from data sets should be aggregated and de-identified to stop data misuse. Culver (2013) argued that the *Journal News* should have aggregated the gun permit data to minimize harm.

These discussions of data journalism ethics, as well as big data more broadly, raise a range of issues highlighting the importance of continued analysis of ethics in this area.

Excellence in Journalism

Data journalism also fits within the scholarly study of the meaning of excellence and its development in the practice of journalism. Scholars have examined both specific criteria of quality or excellence and the evolution of excellence more broadly.

A review of the literature suggests some quality criteria such as accuracy and comprehensiveness have remained constant over time, although the list varied widely in different studies (e.g., Merrill, 1968; Newspaper Advertising Bureau, cited in Burgoon, Burgoon, & Atkin, 1982; Westin, 2003).

Bogart (2004) did a meta-analysis on studies examining excellence in journalism and presented a list of attributes. They include "integrity, fairness, balance, accuracy, comprehensiveness, diligence in discovery, authority, breadth of coverage, variety of content, reflection of the entire home community, vivid writing, attractive makeup, packaging or appearance, and easy navigability" (p. 40).

With emergence of the Internet, some scholars have examined the best practices in online journalism. Based on a survey of 723 online media editors, Gladney, Shapiro, and Castaldo (2007) concluded the traditional standards remained similar on the web.

The highest ranked quality criteria for online journalism are credibility, utility, immediacy, relevance, ease of use, fact–opinion separation, clear paths, simplicity, and exclusivity. Hayes, Singer, and Ceppos (2007) found three fundamental values that are important for online journalism—authenticity, accountability, and autonomy.

Some scholars suggest accuracy, authenticity, balance, independence from faction (e.g., political party), and objectivity are the core conventions of journalism that define the system of values and beliefs of journalists—also known as the journalism paradigm (Berkowitz & Eko, 2007; Hindman & Thomas, 2013; Kovach & Rosenstiel, 2007; Reese, 1990; Thomas & Hindman, 2015).

Another perspective on the components of journalistic excellence, one that emphasizes the evolution of excellence, is based on the theory of a practice, developed by MacIntyre (2007). MacIntyre argues the standards of excellence of a practice evolve as participants pursue excellence, acting with virtue. Their achievements inherent in the practice have been conceptualized by MacIntyre as internal goods.

MacIntyre's (2007) theory has informed several media scholars' discussions of the development of excellence. Lambeth (1992) was first to apply the theory in journalism, followed by Aucoin (2005), Borden (2007), and Craig (2011). Aucoin (2005) applied the theory to understand the practice of investigative journalism, which overlaps with the work of data journalism. Craig (2011) used MacIntyre's framework to examine excellence in online journalism. According to Borden (2007), internal goods of journalism include knowledge and inquiry, among others. For investigative journalism, the internal goods include telling the whole story, emphasizing issues of public importance, telling truth, and making a difference (Aucoin, 2005).

Lambeth (1992) viewed computer-assisted reporting as an extension in the meaning of excellence in journalism. He argued many social problems were uncovered through this type of journalism. Craig (2011) also supported this view, saying new technologies—such as those that provide the means for sophisticated visual presentations of data—provide the capacity for a continued redefinition of the meaning of excellence in journalism.

MacIntyre's (2007) theory suggests as social practices (e.g., journalism) evolve, so does their understanding of excellence. Lambeth (1992) noted the pursuit of standards of excellence "sets in motion a dynamic by which a practice's very capacity to achieve excellence can be systematically elevated and extended. As a result, new concepts of the goods and ends involved in a practice emerge" (pp. 73-74). MacIntyre's theory provides an important conceptual foundation for understanding the place of the journalistic discourse analyzed in this study. This kind of discourse can play a part in the enactment of the practice's role as what Borden (2007) calls a "moral community." Professional discourse "that sustains, repairs, and extends standards of excellent journalism" is one of the ways the practice manifests itself as a moral community (p. 87).

Research Question

This study addresses the following research question:

RQ1: How did the moral community of journalists on the NICAR-L listserv discuss the ethical issues connected with posting of gun permit data?

Background

The publication of the gun permit maps and database prompted strong public reaction, primarily focusing on privacy and safety concerns. News about the maps quickly went viral on social media, partly because criticism and links were posted on several conservative blogs (Horowitz, 2012). *Journal News* readers also responded on its website, and the newspaper reported it had been inundated with phone calls "claiming publication of the database put their safety at risk or violated their privacy" (Horowitz, 2012). More than 1,000 people locally, nationally, and internationally responded to a *New York Times* blog about the incident (Goodman, 2012). The responses were about evenly divided between those opposing and favoring publication of the maps. Opponents were mainly concerned about the invasion of their privacy and possible harm to themselves and their families from gun thieves and public ostracism of them as gun owners. Many respondents said the newspaper acted irresponsibly and questioned its news judgment. Conversely, many respondents lauded the newspaper for the courage to publish the map and show the extent of gun ownership in their community, despite pressure from the gun supporters and legislators. They often questioned why gun supporters would not welcome publication of this information if they felt so strongly about their Second Amendment rights and backed the newspaper's First Amendment rights.

Bloggers responded by posting online the names, addresses, and phone numbers of all employees at the newspaper (Goodman, 2013a). After employees and even a chief executive at parent company Gannett were threatened, the newspaper hired armed guards. Initially, it requested gun permit data from another county and planned to expand the map to it (Goodman, 2013b; Haughney, 2013). About a month after the map was posted, the newspaper kept the map with its data points but removed the links to the personal information of the permit holders. The newspaper's editor said the personal information was removed to comply with the spirit of a new law quickly passed by New York legislators and signed by the governor (Goodman, 2013b). The law prohibits the release of information about gun permit holders for 120 days and allows other permit holders to request the removal of their names and addresses. It was sponsored by the area's state assemblyman.

Although the 2012 case prompted the greatest attention, another New York newspaper reporter's Freedom of Information request for a gun permit database for three counties sparked a public outcry in 2008. The Glen Falls *Post-Star's* editor published an editorial on February 4 in response. "We here at *The Post-Star* find ourselves in the unusual position of responding to the concerns of our readers about something that has not even been published in our newspaper or Web site," the editorial began ("The Post-Star Addresses Pistol Permit Controversy," 2008). The editor said the request "drew great concern from members of gun clubs and people with gun permits in general, a concern we totally understand" ("The Post-Star Addresses Pistol Permit Controversy,"

2008). In addition, the editor assured readers "this newspaper takes its community service responsibility seriously when it comes to sensitive and private information."

Method

The controversy over the White Plains newspaper's actions prompted discussion among professional journalists nationwide, including members of the NICAR-L. Investigative Reporters and Editors, the first and largest organization of investigative reporters in the world, founded the institute in 1989 to train journalists in computer-assisted reporting (Aucoin, 2005). The listers on NICAR-L are among the top data journalists in the field. They include data journalists from the *New York Times, Washington Post*, *Chicago Tribune*, *Dallas Morning News*, *Boston Globe*, *National Public Radio*, *Reuters*, *ProPublica*, and the Center for Public Integrity. Listers also include endowed chairs in journalism from major universities and winners of the Pulitzer Prize and the Philip Meyer Award, named after the author of *Precision Journalism* and given annually for the best use of social science methods in data journalism. From 2000 to 2010, NICAR-L had about 1,200 subscribers. As data journalism has expanded both nationally and internationally, its membership grew quickly to about 1,500 in 2012 when the clickable map was posted and to 2,185 in March 2015 (Ted Peterson, personal communication, March 25, 2015). In MacIntyre's (2007) perspective, they are distinctively qualified to discuss data journalism because they are participants who are pursuing excellence within the practice of journalism.

To find relevant posts, one researcher searched the listserv's archives for posts about "gun permits." Although seven strings discussing gun permit data were found, starting in 2007, it was decided to examine three discussion threads because they mainly focused on ethical principles. The first string involved 72 posts from 36 journalists during February 5 to 6, 2008, about the case in Glen Falls. The last two threads involved the 2012 publication of the story and clickable maps. The first thread involved 39 posts from 26 listers during December 24 to 27, 2012, and discussed the public reaction to posting the story and map. The second one from a follow-up discussion about 2 months later (February 26-27, 2013) involved eight mainly lengthy posts from seven listers.

The first thread was started by a journalist from the *Washington Post* and involved journalists from throughout the United States, including those from the *New York Times* and *Wall Street Journal*, and one from Canada. The second thread was started by a journalist from the *Journal News* informing them about the story. Although journalists nationwide and an international news service participated, journalists from the East Coast contributed the most, including the *New York Times*. The follow-up third thread involved mainly in-depth posts from U.S. journalists, including one from *ProPublica* and one from South America.

A qualitative content analysis was conducted on the posts in the three threads to understand the details of the ethical perspectives of the writers of these posts. The protocol for the analysis was developed using Altheide and Schneider's (2013) qualitative media analysis approach, which has been used in many other studies (listed by Altheide & Schneider, 2013; see, for example, Gormly, 2004).

Following their guidelines, the protocol was drafted listing several categories to guide data collection. The categories constituted the frames of discussion, which Altheide and Schneider defined as "the focus, a parameter or boundary, for discussing a particular event" (p. 53).

In identifying the frames, the researchers focused on emphases in the major theories of ethics, such as duty-based (deontological) and consequence-based (consequentialist) perspectives. However, the authors also left room for other categories to emerge in further reading. The categories were as follows:

1. Freedom versus responsibility
2. Journalistic purpose
3. Privacy
4. Verification
5. Consequences
6. Alternatives

The first four categories all relate to elements of deontological ethics (e.g., Patterson & Wilkins, 2014; Plaisance, 2009) because they involve journalistic duties. Consequences, both positive and negative, clearly relate to consequentialist ethics (e.g., Patterson & Wilkins, 2014; Plaisance, 2009) but also relate to the duty of minimizing harm. Alternatives are raised in the ethical perspective of Sissela Bok (Bok, 1978), who proposed an approach to decision making that included seeking alternatives that will not raise ethical issues (Patterson & Wilkins, 2014).

The unit of analysis was the post. Each post was analyzed for all six categories, which Altheide and Schneider (2013) also referred to as variables, based on definitions established for them. If one text segment within a post met the criteria of more than one variable, it was coded for multiple categories. After reading the texts closely to identify frames, the researchers identified themes—"the recurring typical theses that run through a lot of the reports" (p. 53)—through additional close reading. The "Findings" section presents these frames and key themes, along with specific discourse—detailed examples from the language of the posts—that reflect the themes.

Findings

The findings focus on the frames discussed in the "Method" section and the themes that appeared most frequently in relation to each. (See Table 1 for a summary of the frames and themes.) The frame of freedom versus responsibility is grouped with journalistic purpose because of their relationship with each other, and privacy is discussed with verification for the same reason.

Freedom Versus Responsibility and Journalistic Purpose

The freedom versus responsibility frame includes considerations of freedom of action and responsibility in action, especially possible tension between the two. It also includes whether the information serves a journalistic purpose because it relates so

Table 1. Lists of Frames and Themes.

Frames	Themes
Freedom versus responsibility and journalistic purpose	1. Data should not be posted only because they are public and can be posted. 2. Journalists should have a newsroom discussion about whether to post the data. 3. Requesting data without knowing what to do with them is not responsible and can backfire with the public and editors.
Privacy and verification	1. Risks to a person's private life should be considered when deciding whether to post data online. 2. Excluding certain private information from an online database is ethically justified. 3. Posting data online increases privacy concerns because online information is easier to access than paper records. 4. Posting private information is likely to foster a quicker, angrier backlash. 5. Verification of information is important because personal information can quickly become outdated.
Consequences	1. Publication of names and addresses may put gun permit holders and their families at risk to thieves. 2. Publication of names and addresses angers law-abiding citizens and can lead to legislative action to close access to these records. 3. Reporters may be at risk if their editors abandon them in response to public pressure. 4. Public backlash could result in quick removal of the information from the website. 5. Government workers might leak information to organizations such as gun clubs and undermine journalistic efforts. 6. Names and addresses become inaccurate for some permit holders if data are online for a while.
Alternatives	1. Journalists should try innovative ways to notify gun permit owners that their names and addresses will appear online. 2. Data could be used to show trends, instead of using individual names and addresses. 3. Gun permit data should be joined with other databases to be meaningful and serve a larger journalistic purpose.

closely to issues of responsibility. One of the most frequently mentioned ethical themes in the study was whether gun permit data should be posted online just because they are public data and can be posted. "Yes, it is public information but so are a lot of things that are not ethical or responsible to publish," wrote one lister. Another wrote, "We also have every right to publish them in a clickable map for the world to very easily see. The question is: should we? We retain a responsibility to be, well, responsible."

Many listers also linked posting the information to whether it served a larger public purpose and was accompanied by a story putting the data into context. "I feel I should be enthusiastic about more data, but I don't see what's being done with the data that serves a better conversation on firearms possession," wrote one lister. Another wrote, "Shouldn't it be fair game if we add some sort of public service 'value-added' component? In other words, just the list itself doesn't really serve the public interest . . . (well, in the readers' eyes anyways)." A third wrote,

> I don't see the news value at all in publishing that map. If the point is how commonplace gun permits are, then I would argue the map actually obscures that point. I don't see the inherent public need to know that, for example, my next door neighbor is legally permitted to own a gun.

Others disagreed. "You simply want to know who owns guns. Why do you want to know? Different people, different reasons," wrote one lister. Another wrote,

> There are no special rules for handling gun permit data as much as some groups might wish there were. Publishing it wholesale the way we do so many other data sets or using it for analysis, they are both valid.

A second theme was journalists should have a newsroom discussion about whether to post the data. One lister wrote, "Every data set is like a human source. . . . You weigh whether to publish the information you get from it, in what context and to what end." Another wrote,

> My point is that in the rush to push data onto the web, there has to be an honest debate about competitive and strategic reasons NOT to put data in the hands of consumers—a debate that can feel awkward for many of us who spend so much time championing the public's right to government information.

A third theme was that requesting data without knowing what to do with them is not responsible and can backfire with the public and editors. One lister wrote,

> But sometimes we're like the barking dogs after the pickup truck when it comes to data with guns and it's pretty evident here, from the ed's note that they haven't fully formulated what to do with the info, if they got it.

Another wrote, "This really aggravates me. This kind of stuff reinforces the idea with editors and the public that reporters do nothing but waste time on fishing expeditions."

Privacy and Verification

Privacy implies loss of control by the people whose gun permit data were presented over who sees the gun permit information or the context in which it is presented. Verification implies the need to confirm the accuracy of information in the gun permit

database. The two concepts are linked because publishing inaccurate information can lead to invasion of privacy, and many journalists and their employers seek to verify it before it is published. The main theme was journalists should consider the risks to a person's private life when deciding whether to post data online. One lister wrote, "And if we're going to make something public, or publicize that it's public, then we need to weigh the value of publishing against the risks posed to the people included in the story/database/graphic." Another wrote,

> That is exactly the problem with this map and other stories that believe extreme or radical transparency is good enough to provide meaningful information to its readers/users. They include possible huge violations of regular folks lives (whether you like that they own guns or not) and expose them to things like retaliation, social condemnation, public humiliation and other possible punishment from the community.

The second theme was that excluding certain private information from online databases is ethically justified. One wrote, "I think leaving home address off is good judgment. I'd also leave off a DOB [date of birth] so the identity theft nuts don't freak out." Another wrote, "For instance, we did some foreclosure stuff a few months back. While I had address info on the Web site, I didn't see any reason to put people's names up there."

A third theme was that posting information online, especially in a clickable map, makes it more accessible to people than paper records, increasing privacy concerns. One poster summed it up this way:

> That's why I am against seeing marriage & divorce records slapped up on news sites. It's fine on official gov sites because you have to make an effort to find them, but do you really want your daughters & sisters stalked because some jackass who thinks she ought to be his went to a news site & found out she was single?

A fourth theme was posting private information is likely to foster a quicker, more angry backlash. One lister wrote,

> But that same database also includes their dates of birth and their home addresses, which is invaluable in matching the data against other databases. Are those fields public? Yes. Would I favor putting them online and risking a backlash? Hell no.

A fifth theme was verification of information is important because personal information can quickly become outdated. For example, a gun permit holder could move. This situation makes it especially important to verify the information in the database.

Consequences

A substantial amount of discussion also related to various possible consequences of publishing the map and database, mostly negative but a little positive.

One of the most prominent themes in the entire discussion was gun permit holders and their families may be at risk to thieves and others because their names and

addresses were posted online and their guns are valuable. A few posters thought knowing someone had a gun could deter thieves, who might fear it would be used on them. One lister worried about risk of theft wrote,

> It's not the threat of harm that gun owners are fearful of . . . it's the theft while they're not at home. Guns are valuable, a great lure for thieves who once inside, get the guns and whatever else they can get because they know that registered handgun owners probably have a few other items worth something.

One post quoted a former police officer on Facebook: "I'm not always home to protect my home and family so they are in danger by having my information posted publicly as is every retired officer."

However, another post offered an opposing view:

> I also think the worst-case scenarios posited by opposing lawmakers and advocates are unlikely and also in themselves not enough reason to withhold access to the data (as many, many other kinds of records would be justified to withhold under that standard).

A second prominent theme was publication of the names and addresses of permit holders angers law-abiding citizens and can lead to a variety of consequences, including legislative action to close access to these records. The gun lobby has a closer connection to lawmakers, it was argued, which can result in quick action. One poster wrote, "The other thing is that the more data we put online involving private citizens the more those citizens will push to have that data exempted from public records." Another one, emphasizing the influence of pro-gun interests, said,

> The gun lobby is exactly that: a lobby. Gun owners and gun rights groups give campaign donations to lawmakers. They have staffers working full-time in the state capitol to bend the ears of lawmakers. That's the side lawmakers hear, and they act on it.

Another theme focused on risk to journalists—that reporters are at risk if their editors abandon them in response to public pressure about the gun permit data and also may be at risk from angry users. One person expected no support from managers for this kind of journalism: "God help that pioneering reporter when he writes his story and his editors hang him out to dry." Regarding fear from the outside, another wrote about a previous incident in which "some loonies on our message board posted my address and threatened my family."

Among other themes related to consequences were concerns that the public backlash could result in quick removal of the information from the website, that government workers might leak information to organizations such as gun clubs and undermine journalistic efforts, and that names and addresses would become inaccurate for some permit holders if the data are online for a while. On the last point—an issue common across data journalism—one poster wrote,

> These are questions of scale but also of archiving—the gun map will be around and easily found on search a decade from now, when the facts may have changed and the technical knowledge of how to make corrections to the data in short supply.

Alternatives

The alternatives frame focused on possible approaches that might avoid or minimize ethical problems. One theme within this frame was journalists should try innovative ways to notify gun permit owners that their names and addresses will appear online. These approaches were needed because some media outlets had a policy of notifying anyone whose names will appear online. Furthermore, the large number of gun permit holders in the database made it impractical to notify everyone personally. One journalist argued,

> If we don't reach the parties involved, we must explain in the story what efforts were made to do so. Obviously this is *really hard* in a story like this. One could imagine mailing postcards to all of those people, but that seems expensive, impractical, and slow.

Another journalist suggested the possibility of polling a small subset of the people whose names might appear to see what they think.

A second theme centered on the idea that aggregate data could be used instead of individual gun permit data to avoid using individual names and addresses and to show trends. One post raised this question: "Would the interactive have been less compelling if it omitted the names of the individuals? What if it blurred/distorted their addresses so you knew the general area, but not the specific address." A related approach another person suggested was to group the permits into zones such as census areas and color-code them based on density per thousand.

The third theme was the gun permit data should be joined with other databases to be meaningful and serve a larger journalistic purpose. For example, it could be joined with a database of felons to determine whether felons are illegally granted gun permits. One argument for this was based on precedent: "Previous stories usually joined the data to other data for the stories and meaning." Another comment suggested digging deeper to provide an analysis of weapons per capita based on area income or urban versus suburban versus rural location.

Discussion

This study has addressed ethical issues connected with data journalism as perceived by data journalists and, in particular, the ethics of reporting and visualization of sensitive public data online with details such as names and addresses. The publication of the interactive maps and database of gun permits in New York was particularly sensitive coming in the wake of the Sandy Hook school shootings. It prompted strong public reaction, especially to the maps that made it easier to get private information.

A qualitative content analysis examined how this controversy over publication of gun permit data and related issues were addressed in discussion on the listserv of the

National Institute for Computer-Assisted Reporting. Analysis of 119 posts in three discussion threads shed light on the concerns of the participants, whose organization includes experienced data journalists nationally and worldwide.

The analysis of frames and themes through a close reading of the discourse in the threads found the discussion was framed in several ways, touching on both considerations of ethical duties and consequences. One prominent frame, freedom versus responsibility, centered on the issue of whether gun permit data should be posted just because the data are publicly available and can be posted. This frame also related to considerations of the purposes of journalism and what public needs are actually met by publication of the data. The frame of privacy, a major concern of the public, suggested a loss of control over who sees permit information or the context in which it is presented. That frame was closely tied to another frame of verification because of the prospect for invasion of privacy of other individuals if inaccurate information is published. The frame of consequences, also reflected in public concerns about safety, focused primarily on concerns about negative outcomes such as increasing risk to thieves and rousing the gun lobby against access to these records. (It turned out that gun theft did not actually increase as a result of publication of the data ["Gun Permit Spree Tied to Permit Map Never Materialized," 2013].) The frame of alternatives centered on approaches to the handling of this kind of data that might avoid or minimize ethical problems.

Scholarly Contribution

These findings connect with previous literature in several ways. Literature on the meaning and potential of data journalism (e.g., Anderson, 2011) implied that it could help maximize the ethical potential of journalism, especially by increasing the capacity of investigative journalism to dig deep and get at truth. The findings of this study, however, highlight some of the challenges in using data to provide depth and truthfulness—partly because fear of consequences can prompt a backlash against its use, as well as because of problems of verification of public data sets and risks (actual or feared) of invasion of privacy. Literature on ethical issues and challenges in data journalism is also relevant to the findings of this study. Meyer (2002) noted that data "can create illusions of certitude," and concerns about accuracy of the gun data highlight the problem of making such a powerful presentation of data that may not be fully accurate or will surely become inaccurate with time (Messner & Garrison, 2007). In addition, the study resonates with the concerns raised by Daniel et al. (2010) about incomplete data and with their suggestion that private data could be aggregated and de-identified.

The most important connection with previous studies is with the stream of work examining MacIntyre's (2007) theory of practices and their evolution, carried into journalism study by Lambeth (1992), Aucoin (2005), Borden (2007), and Craig (2011). In MacIntyre's perspective, the members of NICAR, mostly current or former practitioners of data journalism or related work, are distinctively prepared to offer nuanced insight into the ethics of data journalism because they are bearers of the standards of excellence of the practice. As Borden (2007) argued, the practice functions as a moral

community, and one of the ways this happens is through "discourse that sustains, repairs, and extends standards of excellent journalism" (p. 87). The discourse sustained standards of excellence in that it highlighted broad frames—such as freedom versus responsibility, privacy and verification, and consideration of consequences—that are longstanding considerations in journalism that will inevitably accompany the doing of public affairs journalism, whether data journalism or not. The discourse repaired standards as community members on the listserv questioned using data merely because they are publicly available, a concern particularly connected with freedom versus responsibility and journalistic purpose, and raised a number of specific concerns about consequences. This kind of critical evaluation is part of what Borden called solidarity in a moral community involving sanctioning of those who violate standards (though not all thought that standards had been violated).

As for the extension of standards, it is too much to claim that these discussion threads alone changed the standards for data journalism. However, a search showed these were the only ones on the listserv to focus on the ethics of gun permit data, so for this group with expertise in data journalism, the threads represent an important site for the development of standards. The numerous themes that emerged represented concerns specific to data journalism and gun permit data in particular. Beyond that, within two of the threads, specific alternatives—proposals for better practice—emerged in the course of the discussion. This is notable because alternatives represent ways to refine and sharpen standards of excellence.

The work of sustaining, repairing, and extension of standards of excellence reflected in this discourse illustrates how online discussion can provide a space for members of the practice of journalism to function as a moral community and wrestle with ethical challenges they face in pursuing excellent work and how best to address them. The functioning of this moral community contributes to the development of the practice.

Practical Recommendations

Beyond its scholarly contribution, this study points to practical suggestions for journalists who want to do data journalism ethically. The Poynter Institute for Media Studies has created protocols of questions for addressing difficult areas of ethical decision making (e.g., McBride, 2012) but nothing recent related to data journalism (K. McBride, personal communication, October 9, 2015). The frames identified in the posts for this study suggest several considerations reflected in the questions below. These concerns overlap with those in the Poynter protocols because they are ongoing ethical issues, but the questions also include specifics about data journalism that reflect the themes identified in this study. The protocol is addressed to newsroom settings directly, but it is also relevant to discussion online among individual journalists thinking through these issues:

1. Whenever the data contain personal information, a discussion about whether to post them online must take place in the newsroom.
2. Questions to consider:

a. Does the information serve a journalistic and public purpose? To what extent? The data must at least serve both of these purposes to be posted online.
b. Who could be harmed by the information? To what extent? Are there risks to a person's private life from elements of the data? What is the potential impact of data that may be erroneous or out of date?
c. Are there alternatives that would maximize the public purpose such as combining with information from other databases? Are there alternatives that would minimize harm, such as aggregating personal data instead of using individual names and addresses?
d. Can the data be verified? Have reasonable steps been taken to verify the accuracy of the data? Can people in the database be notified before publication? What can be done to enable correction of data errors identified after publication?

3. The final decision should be made with all of these considerations in mind, with journalistic purpose and minimizing harm having the highest priorities.

These recommendations may also be useful in journalism courses on ethics and on data journalism.

Limitations and Future Research

This study is limited by the fact that it grows primarily out of one case based on the interactions of a relatively small number of professional journalists. In addition, interviews could have shed deeper light on their thinking about the ethical considerations. Future research could dig deeper into these ethical questions through in-depth interviews with data journalists, particularly combined with surveys to provide a broader view of the scope of opinion on these and other topics across the field. The protocol of questions developed from the frames and themes identified in this study could provide a template for evaluating the ethical standards of data journalists through interviews and surveys, as well as qualitative or quantitative content analyses of their work itself and discourse about that work.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

Altheide, D. L., & Schneider, C. J. (2013). *Qualitative media analysis*. Los Angeles, CA: SAGE.

Anderson, C. W. (2011). *Notes towards an analysis of computational journalism*. Social Science Research Network. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2009292

Aucoin, J. (2005). *The evolution of American investigative journalism*. Columbia: University of Missouri Press.

Berkowitz, D., & Eko, L. (2007). Blasphemy as sacred rite/right: "The Mohammed cartoons affair" and maintenance of journalistic ideology. *Journalism Studies*, *8*, 779-797.

Bogart, L. (2004). Reflections on content quality in newspapers. *Newspaper Research Journal*, *25*, 40-53.

Bok, S. (1978). *Lying: Moral choice in public and private life*. New York, NY: Random House.

Borden, S. L. (2007). *Journalism as practice: MacIntyre, virtue ethics and the press*. Burlington, VT: Ashgate.

boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, *15*, 662-679.

Boyle, D. (2001). *The sun of our discontent: Why numbers make us irrational*. New York, NY: Texere.

Bradshaw, P. (2015). Data journalism. In L. Zion & D. Craig (Eds.), *Ethics for digital journalists: Emerging best practices* (pp. 202-219). New York, NY: Routledge.

Burgoon, J. K., Burgoon, M., & Atkin, C. (1982). *The world of the working journalist*. New York, NY: Newspaper Advertising Bureau.

Cairo, A. (2016). *The truthful art: Data, charts, and maps for communication*. San Francisco, CA: New Riders.

Carter, E. L., & Lee, L. T. (Eds.). (2016). Information access and control in an age of big data [Special issue]. *Journalism & Mass Communication Quarterly*, *93*(2).

Coddington, M. (2015). Clarifying journalism's quantitative turn: A typology for evaluating data journalism, computational journalism, and computer-assisted reporting. *Digital Journalism*, *3*, 331-348.

Cohen, S., Hamilton, J., & Turner, F. (2011, October). Computational journalism. *Communications of the ACM*. Retrieved from http://cacm.acm.org/magazines/2011/10/131400-computational-journalism/fulltext#

Craig, D. A. (2011). *Excellence in online journalism: Exploring current practices in an evolving environment*. Los Angeles, CA: SAGE.

Culver, K. B. (2013, February 22). Where the Journal News went wrong in mapping gun owners. *MediaShift*. Retrieved from http://mediashift.org/2013/02/where-the-journal-news-went-wrong-in-mapping-gun-owners053/

Daniel, A., & Flew, T. (2010, November 15-16). *The Guardian reportage of the UK MP expenses scandal: A case study of computational journalism*. Paper presented at the Communications Policy & Research Forum, Sydney, Australia. Retrieved from http://eprints.qut.edu.au/38701/2/38701.pdf

Daniel, A., Flew, T., & Spurgeon, C. (2010). The promise of computational journalism. In K. McCallum (Ed.), *Media, democracy and change: Refereed proceedings of the Australian and New Zealand Communications Association Annual Conference, Australian and New Zealand Communication Association* (pp. 1-19). Canberra, Australia: Australian and New Zealand Communication Association.

Fairfield, J., & Shtein, H. (2014). Big data, big problems: Emerging issues in the ethics of data science and journalism. *Journal of Mass Media Ethics*, *29*, 38-51.

Flew, T., Spurgeon, C., Daniel, A., & Swift, A. (2011). The promise of computational journalism. *Journalism Practice*, *6*, 157-171.

Garrison, B. (1998). *Computer-assisted reporting* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

Gladney, G., Shapiro, I., & Castaldo, J. (2007). Online editors rate web news quality criteria. *Newspaper Research Journal*, *28*, 55-69.

Goodman, J. D. (2012, December 26). Outcry over newspaper's map of handgun permit holders. *The New York Times*. Retrieved from http://cityroom.blogs.nytimes.com/2012/12/26/a-newspaper-publishes-names-of-gun-permit-holders-sparking-outrage/?_r=0

Goodman, J. D. (2013a, January 18). Newspaper takes down map of gun permit holders. *The New York Times*. Retrieved from http://www.nytimes.com/2013/01/19/nyregion/newspaper-takes-down-map-of-gun-permit-holders.html

Goodman, J. D. (2013b, January 2). Newspaper that put gun permit map online hires armed guards. *The New York Times*. Retrieved from http://www.nytimes.com/2013/01/03/nyregion/putnam-officials-keep-gun-permit-records-from-journal-news.html?_r=0

Gormly, E. (2004). Peering beneath the veil: An ethnographic content analysis of Islam as portrayed on the 700 Club following the September 11th attacks. *Journal of Media and Religion*, *3*, 219-238.

Gray, J., Chambers, L., & Bounegru, L. (2012). *The data journalism handbook: How journalists can use data to improve the news*. O'Reilly Media. Retrieved from http://datajournalismhandbook.org/1.0/en/introduction_2.html

Gun permit spree tied to permit map never materialized. (2013, June 15). *Journal News*. Retrieved from http://www.lohud.com

Hamilton, J. T., & Turner, F. (2009). *Accountability through algorithm: Developing the field of computational journalism*. Retrieved from http://web.stanford.edu/~fturner/Hamilton%20Turner%20Acc%20by%20Alg%20Final.pdf

Haughney, C. (2013, January 7). After pinpointing gun owners, paper is a target. *The New York Times*, p. A1.

Hayes, A. S., Singer, J. B., & Ceppos, J. (2007). Shifting roles, enduring values: The credible journalist in a digital age. *Journal of Mass Media Ethics*, *22*, 262-279.

Hindman, E. B., & Thomas, R. J. (2013). Journalism's "crazy old aunt": Helen Thomas and paradigm repair. *Journalism & Mass Communication Quarterly*, *90*, 267-286.

Horowitz, J. (2012, December 27). Paper's map of gun permit owners stirs backlash. *The Washington Post*, p. C03.

Houston, B. (1999). *Computer-assisted reporting: A practical guide*. Boston, MA: St. Martin's Press.

Houston, B. (2010). The future of investigative journalism. *American Academy of Arts & Sciences*, *139*, 45-56. doi:10.1162/daed.2010.139.2.45

Jonsson, P. (2012, December 14). School shootings in Connecticut cast somber mood across nation. *The Christian Science Monitor*. Retrieved from http://lexisnexis.com

Kovach, B., & Rosenstiel, T. (2007). *The elements of journalism: What newspeople should know and the public should expect*. New York, NY: Three Rivers Press.

Krueger, K. (2014, July 7). How J-schools are tackling the demand for data journalism skills. *MediaShift*. Retrieved from http://www.pbs.org/mediashift/2014/07/how-j-schools-are-tackling-the-demand-for-data-journalism-skills/

Lambeth, E. B. (1992). *Committed journalism: An ethic for the profession* (2nd ed.). Bloomington: Indiana University Press.

Lewis, S. C., & Westlund, O. (2015). Big data journalism. *Digital Journalism*, *3*, 447-466.

Liptak, A., & Peters, J. W. (2013, November 19). Congress and the courts weigh new attempts to scale back N.S.A. spying. *The New York Times*, p. A16.

MacIntyre, A. (2007). *After virtue* (3rd ed.). Notre Dame, IN: University of Notre Dame Press.

Mayer-Schonberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work and think*. Boston, MA: Houghton Mifflin Harcourt.

McBride, K. (2012). Questions to consider before publishing autopsy reports. *Poynter*. Retrieved from http://www.poynter.org/latest-news/everyday-ethics/186250/questions-to-consider-before-publishing-autopsy-reports/

Merrill, J. C. (1968). *The elite press: Great newspapers of the world*. New York, NY: Pitman Publishing.

Messner, M., & Garrison, B. (2007). Journalism's "dirty data" below researchers' radar. *Newspaper Research Journal*, *28*, 88-100.

Meyer, P. (2002). *Precision journalism: A reporter's introduction to social science methods*. New York, NY: Rowman & Littlefield.

Patterson, P., & Wilkins, L. (2014). *Media ethics: Issues and cases* (8th ed.). New York, NY: McGraw-Hill.

Plaisance, P. L. (2009). *Media ethics: Key principles for responsible practice*. Thousand Oaks, CA: SAGE.

Podesta, J., Pritzke, P., Moniz, E. J., Holdren, J., & Zients, J. (2014). *Big data: Seizing opportunities, preserving values*. Retrieved from http://www.whitehouse.gov/issues/technology/big-data-review

The Post-Star addresses pistol permit controversy. (2008, February 4). *Post Star*. Retrieved from http://poststar.com/news/latest/the-post-star-addresses-pistol-permit-controversy/article_1ba7c108-8116-5940-8139-d4ec01e5d55e.html

Reese, S. D. (1990). The news paradigm and the ideology of objectivity: A socialist at the Wall Street Journal. *Critical Studies in Media Communication*, *7*, 390-409.

Thomas, R. J., & Hindman, E. B. (2015). Confusing roles, uncertain responsibilities: Journalistic discourse on Juan Williams, NPR, and Fox News. *Journalism & Mass Communication Quarterly*, *92*, 468-486.

Vallance-Jones, F., & McKie, D. (2009). *Computer-assisted reporting: A comprehensive primer*. Don Mills, Ontario, Canada: Oxford University Press.

Westin, A. (2003). *Best practices for television journalists*. Fair Press Project. Retrieved from http://www.freedomforum.org/publications/BestPractices/BestPractices.Westin.pdf

Worley, D. R. (2012, December 24). The gun owner next door: What you don't know about the weapons in your neighborhood. *The Journal News*. Retrieved from http://www.lohud.com

Author Biographies

David Craig is the associate dean for academic affairs and a professor in the Gaylord College of Journalism and Mass Communication at the University of Oklahoma. The primary focus of his teaching and research is media ethics. He is the author of *Excellence in Online Journalism: Exploring Current Practices in an Evolving Environment* (Sage, 2011) and *The Ethics of the Story: Using Narrative Techniques Responsibly in Journalism* (Rowman & Littlefield, 2006). He earned a BS in journalism from Northwestern University, an MA in communication from Wheaton College, and a PhD in journalism from the University of Missouri–Columbia.

Stan Ketterer is an associate professor of journalism in the School of Media and Strategic Communications at Oklahoma State University. His major research interest is using social science methods in investigative reporting. He has been a speaker and panelist at the national

Computer-Assisted Reporting Conference and conducted workshops on the topic. His research interests also include international journalism and sports journalism. He has published 15 articles and book chapters. Ketterer earned a BA, MEd, and PhD from the University of Missouri–Columbia.

Mohammad Yousuf is an adjunct instructor in the Gaylord College of Journalism and Mass Communication at the University of Oklahoma. His dissertation focused on issues relating to strategic management of online media. His research interests include strategic management of online media, media economics, media ethics, journalistic use of social media, data journalism, and news media and democracy. His scholarship reflects 8 years of work experience with national and international media in Bangladesh as a reporter and copy editor. He earned a Master of Social Sciences from the University of Dhaka, Bangladesh, and a PhD from the University of Oklahoma.

View publication stats