Hashtag Archiving

Tara L. Conley

In this chapter I introduce the term *hashtag archiving* to describe a human-centered approach to capturing discourse. Although this analysis cites relevant literature on hashtag research, along with reference cases, steps, and suggestions for archiving hashtag data, it is not definitive. Rather, it is a starting point for those with the foresight to memorialize hashtag data. Hashtag archivists recognize that history turns into myth, communication technologies transform, interfaces disappear, meaning constructed through social media use changes, and interest in hashtag research may eventually wane. Hashtag archivists also understand that memories trick us. Throughout the process of collecting and organizing data, hashtag archivists become storytellers by piecing together narratives from metadata. No matter the technique or theory used to inform gathering hashtag data, it is therefore necessary for hashtag archivists to incorporate reflective practices when building archives that capture the stories of our time.

The Story of the Hashtag

The # symbol has an uncertain history. Accounts of its origin are often vague and even mythical with respect to its typographical lineage and name. It is an interwoven history characterized by an evolution of symbols, spoken languages, myths, and codes that can be traced across time periods from the Roman Empire, to the Enlightenment, to the present day, as well as across technical fields such as cartography and computer programming. In the late fourteenth century, according to punctuation historian Keith Houston (2013), "the abbreviation 'lb' for *libra* entered English" (42). Centuries later, *lb* took on many different forms as a result of "carelessly rushing pens" (43). One notable example, Houston illustrates, occurred when Sir Isaac Newton hastily scribbled ℔, or *libra*, the abbreviation for "pound in weight." This was one of the earliest documentations of the # symbol. A similar Latin term, *pondo* (loosely meaning "to weigh"), was going through a transformation of its own in spoken language. The Latin term *pondo* evolved into the Old English phrase *pund*, which eventually became the modern-day term *pound*. It was across these moments in time, when the histories of typography and spoken language collided, or, as Houston (2013, 43) describes, "When *libra* and *pondo* were reunited," that the # symbol (or pound sign) was born.

In the US, contemporary perceptions of the pound sign (# symbol) can be traced back to the early telecommunications and Internet eras. In telecom the # symbol took on yet another name, the octothorpe. The "tortuous history" of the octothorpe marks a starting point for

scholars to analyze the # symbol's significance in communication technologies and as a cultural cue of modern discursive practices (Conley 2017; Houston 2013, 48; Salazar 2017). As communication technologies evolved, so, too, did the # symbol. Take, for instance, another modern-day origin story that occurred on August 23, 2007, when programmer Chris Messina used the # symbol for the first time on Twitter. He tweeted: "How do you feel about using # (pound) for groups. As in #barcamp [msg]?" (@chrismessina, August 23, 2007). Messina was harking back to 1980s cyberculture, when the # symbol was used across online communities as a way to identify channels of interest (Salazar 2017, 22). Messina's use of the # symbol in 2007 adds yet another layer to the historical tapestry of the hashtag. This moment in time further clarifies the significance of the # symbol in terms of how information is called up and organized and how discursive practices emerge.

Hashtag Archiving and Uncertain Etymologies

The etymologies of hashtags on social media can be as uncertain as the history of the # symbol itself. Yet this uncertainty is perhaps one of the reasons why researchers, journalists, media makers, and activists are drawn to moments when hashtags appear, trend, and become the focus of mass media and popular culture. Hashtags help the tellers build stories about the worlds in which we live. For this reason I propose the term *hashtag archiving* as a human-centered approach to capturing discourse in an era of uncertain hashtag data.

Hashtag archiving is a process of capturing and preserving social media data assigned to the visual (#) and nonvisual (U + 0023) dimensions of code that also require interpretive analysis and collaboration. Hashtag archiving involves annotating, indexing, and curating in order to build repositories—whether across interactive media platforms or in the form of text-based documentation—for public retrieval. In 2013 I published the website www.hashtagfeminism.org in order to document and archive feminist discourse across Twitter. This platform is an example of a hashtag data repository that serves as a public resource. Gathering such data sets is often a makeshift process. It involves collecting data across multiple platforms and channels, including manual methods (e.g., using Excel), fee-for-service analytics platforms, and the crowdsourcing of data sets from other resources. One example of a collaborative research tool and consortium that crowdsources Twitter data sets is www.docnow.io. That said, hashtag archiving is not synonymous with digital archiving in the sense of reimaging content on the web. Rather, the process of hashtag archiving locates digitally born content in order to analyze discursive practices, social movements, and storytelling across online and off-line spaces.

Hashtag archiving is also an explicit approach to social justice that incorporates feminist and antiracist frameworks. Punzalan and Caswell (2016, 30–33) argue for more critical approaches to archiving that include archival pluralism, ethics of care, and non-Western ways of knowing. They state these approaches in efforts to confront the authoritarian, white, hegemonic structures that perpetuate systemic inequalities in the field of archiving. In arguing for a more explicit acknowledgment of social justice, Punzalan and Caswell also point out several key characteristics, three of which are most relevant to hashtag archiving:

(1) the inclusion of underrepresented and marginalized sectors of society; (2) the reinterpretation and expansion of archival concepts that tend to favor dominant cultures and state and corporate entities; (3) the development of community archives (27–29).

When considering hashtag archiving in the context of the inclusion of underrepresented and marginalized sectors of society, I turn to the issue of origination, which is worthy of further discussion. I use the phrase *origin story* throughout this chapter so I do not privilege time and numbers over context and culture but emphasize attribution. When documenting the origins of hashtag campaigns and movements, I emphasize the labor of Black and nonwhite queer, women, femme, and trans people across online media. Their work has too often been erased across white feminist and liberal media landscapes. For instance, I have been critical of media professionals with public online and off-line platforms minimizing Black women's work online. Most notably, I have discussed white liberal media's early erasure of Tarana Burke's off-line advocacy work in founding the Me Too movement, which Burke also points out (Burke 2018; Conley 2018). Also, before Twitter was considered the primary mechanism for using hashtags to promote awareness and organize around social and political causes, I pointed out white women's appropriation of Black feminist thought in the blogosphere (Conley 2013a; Florini 2013). These issues are not new; however, they reflect similar instances of erasure and appropriation that characterize second-wave feminism and neoliberal thought in the latter half of the twentieth century (Conley 2017, 26–27). So indeed, origination matters in hashtag archiving to the extent that the unseen labor of those who contribute to shaping culture and society also matters.

Ethical Considerations for Hashtag Archiving

Annotating Protest: The Case of #Ferguson

In a lengthy footnote to their seminal piece on Ferguson, Yarimar Bonilla and Jonathan Rosa (2015, 12) discuss the ethical implications of citing tweets from the early moments when an eighteen-year-old Black man named Michael Brown was shot and killed in Ferguson, Missouri, during the summer of 2014. Bonilla and Rosa cite, without identifying, a Twitter user who tweeted: "I just saw someone die" (4). They provide a rationale for their decision not to identify the Twitter user: "We have thus thought carefully about when to quote, cite, or paraphrase from Twitter posts. We have used real names of Twitter users when discussing tweets that went 'viral' or were featured in mainstream media reports. However, when quoting or paraphrasing from unreported tweets, we have chosen not to reproduce the username of the author—erring on the site [*sic*] of privacy at the expense of offering proper attribution" (12).

In this case Bonilla and Rosa's decision to prioritize anonymity over attribution reflects an anthropological approach to conducting Twitter research that they call *hashtag ethnography*. They do not locate Twitter threads showing images and video of Michael Brown lying dead in the street. Along with images and video, the tweet referenced was attached to #Ferguson through retweets, saved on archived Twitter time lines, and amplified by mainstream news

reports. These events raise important questions for researchers and archivists documenting cases similar to #Ferguson when moments of violence and trauma are immediately accessible.

Although it is beyond the scope of this chapter, it is worth noting that along with #Ferguson, there are ethical issues to consider when archiving #MeToo and instances of sexual violence. As scholarly literature on #MeToo grows, so does concern about archiving trauma. In addition, the Me Too hashtag has inspired adjacent hashtags, such as #MeTooIndia, #MeTooK12, and #MeTooMosque, which carries further ethical implications for archiving transnational stories of sexual violence across the web. In outlining this case study on #Ferguson, I suggest that archivists of #MeToo might consider the ethics of annotating protest movements with stories about personal and community trauma.

In building a corpus for analyzing the story of protest movements, one might ask: To what extent, if at all, does documenting a time line of events, from the early moments of injustice to the rise of social and political protest, contribute to a hashtag archive for the public?

For efforts to determine ethical approaches to annotating the rise of digital protest movements, I propose the following recommendations:

- Provide an interpretation of "early-moment" data and why certain information is documented. Bonilla and Rosa offer additional commentary in a footnote to contextualize an original tweet about Michael Brown's death. Hashtag archivists' annotations of early moments should be central to analysis. Annotations can be presented in various forms (hypertext, image, video, audio) and by way of multimedia annotation platforms—for example, by using www.thinglink.com and www.scalar.me/anvc/scalar/.
- Next, explain the political and social context of the moment and why it is necessary to prioritize anonymity in an era of increased federal and law enforcement surveillance of protesters and others who document and publicly denounce state-sanctioned violence across digital and social media. I provide an example that illustrates an explanation of context in a 2013 online case study I conducted about Renisha McBride, a young woman killed in Dearborn, Michigan (Conley 2013b).
- Assess the sources are already accessible, usually through news reports, and decide whether it is necessary to cite or not cite, as Bonilla and Rosa chose to do, in order to err on the side of anonymity. Explain a rationale for citing or not citing and why it matters to the archive.
- Finally, continue to reassess archived hashtag data over time to check for changes in digital content and inaccuracies (hashtag data can easily be changed or erased). This process can be accomplished through manual methods, fee-for-service data analytics platforms, and the crowdsourcing of data from collaborators.

Bonilla and Rosa's article was the first to introduce hashtag ethnography as a practice for examining social movements in Twitter research. Hashtag ethnography informs hashtag archiving in terms of its emphasis on reflective practices for memorializing the early moments of social and political movements online. One also learns from Bonilla and Rosa's

work that the hashtag researcher is, in fact, an archivist.

Indexing the Signifying Monkey: The Case of #BlackTwitter

The # symbol did not create Black Twitter. As with Ferguson, the # symbol located moments and discursive practices emerging through Twitter. André Brock and Sarah Florini are among the early scholars of Black Twitter who have analyzed it as a form of Black oral tradition and Black cultural discourse, also known as *signifyin'* (Brock 2012, 530; Florini 2014, 226). In citing Henry Louis Gates's work on the signifying monkey, Brock and Florini offer a framework for understanding Black Twitter as counterpublics (Gates 1983, 1988):

> Black hashtag signifying revealed alternate Twitter discourses to the mainstream and encourages a formulation of Black Twitter as a "social public": a community constructed through their use of social media by outsiders and insiders alike. (Brock 2012, 530)
>
> On Black Twitter, signifyin' often functions as a marker of Black racial identity by indexing Black popular culture. One example is the popular hashtag game in "Hip hop" circles signifyin' on the R&B singer and rapper Drake. Hashtags such as #DrakePunchlines or #FakeDrakeLyrics mock Drake's lyrical techniques, such as his use of the truncated metaphor (i.e., a phrase immediately followed by an associated word or phrase). This style is sometimes referred to as "hashtag rap" because it mirrors a common use of hashtags on Twitter. (Florini 2014, 227)

Brock and Florini provide a means for understanding #BlackTwitter as a way to locate Black Twitter counterpublics. #BlackTwitter not only marks Black cultural discourse but also locates "sites of resistances [that] allow for double-voiced and encoded communication" (Florini 2014, 226). That is to say, among these counterpublics, Black Twitter users construct meaning through referents and cues that the dominant culture cannot easily decode. In this case, hashtag archiving #BlackTwitter expands upon traditional archival concepts that tend to favor dominant culture(s) and instead considers Black Twitter counternarratives as integral to our collective memory (Punzalan and Caswell 2016, 29).

It is worth noting here that the study of Black Twitter is part of a much larger and burgeoning body of work on hashtags as sites to examine publics and counterpublics. Some of this work includes examining, for instance, the sociopolitical practice of hashtag(ging) (Myles 2018); the role of hashtags in producing and amplifying social movements (Jackson 2016; Jackson and Foucault Welles 2016; Kuo 2018); hashtags as sites to locate affinity spaces (Khoja-Moolji 2015; Walton and Oyewuwo-Gassikia 2017); and hashtags as modes of storytelling (Conley 2019; Yang 2016).

Archiving counterpublics also raises important questions about indexing encoded cultural practices—namely, to what extent can discursive identities based on shared referents of articulations, performances, humor, wordplay, and style be classified, if at all (Brock 2018,

1017)? To address this question, hashtag archivists should first build an index of their own practices and approaches. Consider these suggestions for taking an inventory of one's orientation to hashtag data:

- When indexing community practices through counqugh counterpublics such as #BlackTwitter, hashtag archivists must accept they are lurkers and visitors, whether they consider themselves part of the community or not. Ask: How do archivists construct meaning about users and user practices when lurking across and among various platforms and interfaces? Include these reflections as part of the archive.
- Hashtag archivists should embrace a social justice model of archiving as presented by Punzalan and Caswell. Namely, archivists must consider the development of community archives with, for, and in service of others. Ask: In what ways can hashtag archivists facilitate collaboration and participation in efforts to build repositories of encoded communication?
- Hashtag archivists must avoid the desire to generalize users and their practices online when hashtags are unable to articulate cultural identities. Ask: What are the limitations of hashtags in documenting culture? What are hashtags unable to account for and locate, even with the most sophisticated tools and methods for analysis?

Certainly, answers to these questions may never fully be satisfied. For this reason, hashtag archivists should be comfortable with the uncertainty of indexing discursive practices using hashtag data.

References

ɔnilla, Yarimar, and Jonathan Rosa. 2015. "#Ferguson: Digital protest, hashtag ethnography, and the racial politics of social media in the United States." *American Ethnologist* 42:4–17. https://doi.org./10.1111/amet.12112.

ock, André. 2012. "From the Blackhand side: Twitter as a cultural conversation." *Journal of Broadcasting and Electronic Media* 56 (4): 529–549. https://doi.org./10.1080/08838151.2012.732147.

ock, André. 2018. "Critical technocultural discourse analysis." *New Media and Society* 20 (3): 1012–1030. https://doi.org./10.1177/1461444816677532.

ırke, Tarana. 2018. "#MeToo founder Tarana Burke on the rigorous work that still lies ahead." *Variety*, September 25, 2018. https://variety.com/2018/biz/features/tarana-burke-metoo-one-year-later-1202954797/.

ɔnley, Tara L. 2013a. "An open letter to Amanda Marcotte." *Feminist Wire*, March 4, 2013. http://www.thefeministwire.com/2013/03/an-open-letter-to-amanda-marcotte/.

onley, Tara L. (website). 2013b. "Tracing the impact of online activism in the Renisha McBride case." Accessed May 19, 2019. https://taralconley.org/media-make-change/blog/2013/tracing-the-impact-of-online-activism-in-the-renisha-mcbride-case.

onley, Tara L. 2017. "Decoding Black feminist hashtags as becoming." *Black Scholar* 47 (3): 22–32. http://dx.doi.org/10.1080/00064246.2017.1330107.

onley, Tara L. 2018. "Framing #MeToo: Black women's activism in a white liberal media landscape." *Media Ethics* 30 (1). https://www.mediaethicsmagazine.com/index.php/browse-back-issues/210-fall-2018-vol/3999237-framing-metoo-black-women-s-activism-in-a-white-liberal-media-landscape.

onley, Tara L. 2019. "Black women and girls trending: A new(er) autohistoria teoría." In *This Bridge We Call Communication: Anzaldúan Approaches to Theory, Method, and Praxis,* edited by Leandra Hinojosa Hernandez and Robert Guiterrez-Perez, 231–256. Lanham, MD: Lexington Books.

orini, Sarah (website). 2013. "White feminists: Step your game up." July 2013. http://www.sarahflorini.com/?p=130.

orini, Sarah. 2014. "'Tweets, tweeps, and signifyin': Communication and cultural performance on 'Black Twitter.'" Special issue. *Television and New Media* 15 (3): 223–237. https://doi.org./10.1177/1527476413480247.

ites, Henry Louis. 1983. "The blackness of blackness: A critique of the sign and the signifying." *Critical Inquiry* 9 (4): 685–723.

ites, Henry Louis. 1988. *Signifying Monkey: A Theory of African-American Literary Criticism.* London: Oxford University Press.

ouston, Keith. 2013. *Shady Characters: The Secret Life of Punctuation, Symbols, and Other Typographical Marks.* New York: W. W. Norton.

ckson, Sarah J. 2016. "(Re)imagining intersectional democracy from Black feminism to hashtag activism." *Women's Studies in Communication* 39 (4): 375–379. http://dx.doi.org/10.1080/07491409.2016.1226654.

ckson, Sarah J., and Brooke Foucault Welles. 2016. "#Ferguson is everywhere: Initiators in emerging counterpublic networks." *Information, Communication and Society* 19 (3): 397–418. https://doi.org/10.1080/1369118X.2015.1106571.

noja-Moolji, Shenila. 2015. "Becoming an 'intimate publics': Exploring the affective intensities of hashtag feminism." *Feminist Media Studies* 15 (2): 347–350. https://doi.org/10.1080/14680777.2015.1008747.

io, Rachel. 2018. "Racial justice activist hashtags: Counterpublics and discourse circulation." *New Media and Society* 20 (2): 495–514. https://doi.org/10.1177/1461444816663485.

yles, David. 2018. "'Anne goes rogue for abortion rights!' Hashtag feminism and the polyphonic nature of activist discourse." *New Media and Society* 21 (2): 507–527. https://doi.org/10.1177/1461444818800242.

nzalan, Ricardo L., and Michelle Caswell. 2016. "Critical directions for archival approaches to social justice." *Library Quarterly* 86 (1): 25–42. https://doi.org/10.1086/684145.

lazar, Eduardo. 2017. "Hashtags 2.0: An annotated history of the hashtag and a window into its future." *Icono 14* 15 (2): 16–54. https://doi.org/10.7195/ri14.v15i2.1091.

alton, Quenette L., and Olumbunmi Basirat Oyewuwo-Gassikia. 2017. "The case for #BlackGirlMagic: Application of a strengths-based, intersectional practice framework for working with Black women with depression." *Journal of Women and Work* 32 (4): 461–475.

ng, Guobin. 2016. "Narrative agency in hashtag activism: The case of #BlackLivesMatter." *Media and Communication* 4 (4): 13–17. http://dx.doi.org/10.17645/mac.v4i4.692.