Renegotiating the Privacy Ethics of Digital Genealogy Databases: An Intersectional Analysis of Technology, Criminality, and DNA

While not the primary focus of this presentation, we will approach some of the crimes of the Golden State Killer, many of which are sex-based offenses. The East Area Rapist began attacking east Sacramento in the summer of 1976, including the communities of Rancho Cordova and Danville. Police hold this perpetrator responsible for assaulting at least 50 women in a series of attacks that lasted until July 1979. The Visalia Ransacker terrorized Visalia, California, a community just south of Fresno, through a series of burglaries and vandalisms in 1974-75. The Original Night Stalker is one of the most prolific serial offenders in California state history. He operated in Ventura County and Santa Barbara County, north of Los Angeles, but he also targeted victims in Orange County, near Irvine. He is responsible for at least ten murders in these areas from October 1979-May 1986. Through DNA-matching technology using the digital genealogy database GEDmatch.com, police confirmed that these three perpetrators were in fact the same perpetrator, Joseph James D’Angelo.

In April 2018, Sacramento County sheriffs arrested D’Angelo, aka The Golden State Killer (McNamara, 2018), at the home he shared with his daughter and grandchild. By arresting D’Angelo authorities identified a serial sexual offender who terrorized California for decades. However, the surveillance tactics and DNA-matching technology authorities used in the identification and apprehension of D’Angelo created questions concerning the ethical implications of user actions on digital genealogy databases. This rhetorical act changed the way we think about the triangulation of technology, genetics, and criminality; moreover, this case serves as a primary point to consider the potential for divergent ethics guiding corporations, government organizations, and individual users using the data generated on these sites.

It is important to note, I think, that while I identify one exigency for this project as the arrest of D’Angelo, there are multiple factors which could be identified as the exigent cause for this phenomenon. The purpose of focusing on D’Angelo is to make concrete connections to the impact digital genealogy databases have on the evolution of police tactics. This work, then, follows in the footsteps of Carolyn Miller (1979), Steven Katz (1992), Angela Haas (2007), and a bevy of other scholars who have called for a humanistic approach to technical communication as a field and as a discipline, as well as rhetoricians Jeff Grabill and Stacey Pigg (2012), who acknowledge the *messy* nature of identity performance in digital spaces. Part of the messiness, or complexity, here is that DNA-matching technologies have varying implications for people of different races and cultures. The impact of digital identity formation and curation on a genealogy database, and the entire spectrum of DNA-matching technologies, impacts a Hasidic Jew in Brooklyn in a much different way than it impacts a detained child at the US-Mexico border. So, yes, we must ask ourselves, what are the implications of this phenomenon on police tactics, but also, what are the implications of this phenomenon on our culture(s)?

The purpose of this project, then, is to understand user actions on digital genealogy databases as I work toward a public ethical framework for using these sites. The problem I explore is the (un)ethical use of genealogy databases by individual users, police and government agencies, as well as public and private corporations. The intersection of DNA-as-evidence and digital genealogy databases serves as a revolutionary moment for digital rhetoric scholars concerning investigation into ethical responsibility, collaborative digital identity construction, and the potential for participatory investigation (amateur sleuthing). While these niche communities grow in popularity, our understanding of casting blame and finding justice through the application of these digital tools continues to evolve as the function of these sites evolves. We must consider the ethical implications of using digital genealogy databases to catch criminals; moreover, how do these sites protect user information in the age of Big Data when the purpose of the community is to share as much information as possible and to discover as much information as possible? DNA-matching technology is not a perfect science, as highlighted by Katie Worth in 2015 and Lauren Kirchner in 2017. Using DNA-matching technology is a rhetorical act which deserves heightened analysis as minority oppression pervades American culture. We must consider how these communities function as inherently racist and classist sites which privilege certain users.

Within this context, issues regarding ethical user responsibility abound, and rhetoric and composition is poised to address these concerns. Whether the user is a student performing a school project, a parent searching out a child previously put up for adoption, a police agency uploading DNA to identify a criminal, or a health insurance company seeking additional data to set premiums, we must work to ensure users consider the literacies (Selber, 2004) needed to navigate these sites and use data ethically. As we work toward a public ethical framework for utilizing digital genealogy databases we can begin by considering the conceptualization of *consent* through case studies of the User Agreements for these sites. I identify User Agreements as sites which fit Gross’s (1996) contextual model, places to research to understand how audience’s cultural, political, and socioeconomic conditions, along with their ability to use technological resources, affect consent.

In reconsidering ethical approaches to data collection on these sites, I emphasize the notion we should contend with the policies that shape online experiences (Edwards, 2018). Reframing User Agreements interrogates the role of consent when enacting critical rhetorical action. Moreover, we are contributing to the field of rhetoric and composition by expanding research regarding the ethics of data collection and identity formation within these unique digital ecosystems. But, while our contributions are within this field, there is much we can glean by application of methods from other disciplines as we construct an ethical framework.

It might be helpful to consider the conceptualization of *consent* in the User Agreements guiding digital genealogy databases from an interdisciplinary perspective. Using the free, downloadable software *AntConc*, version 3.5.8, I learned that *consent* is the 128th most used word in a corpus comprised of the User Agreements of GEDmatch.com, ancestry.com, and 23andme.com. The term *consent* maintains a frequency of 18 times used. This data, from an admittedly, but intentionally, small corpus, suggests that, to the companies which own and operate digital genealogy databases and to potential users, like police agencies, *consent* is not as important to them as, say, *privacy*, which is the 87th most used word in the corpus, with a frequency of 28 times used. *Consent* certainly is not as important as the term *data* which is the 25th most used word in the corpus, with a frequency of 129 times used. This indicates that the genealogy companies are much more interested in user data than with ethical user privacy and ethical user consent.

It is critical for users to understand that genealogy databases are companies, and that these companies are more interested in user data than privacy or consent. Users must understand that *privacy* works against the goal of the company and the database, to form a digital identity and to build an online community. In many ways, so does *consent*. Digital genealogy databases, as well as the partnerships between DNA-matching technologies and law enforcement agencies need public oversight; an ethical framework for users navigating these sites works to achieve this goal. In initiating a multi-pronged, interdisciplinary response which promotes more ethical outcomes for users of these sites by filling the gap between ethical consent and user action, we can transform practices on genealogy databases by enacting an ethical framework which fosters responsibility concerning consent at its core.

Turning back to scholarship focused in rhetoric and composition, particularly digital rhetoric, we will find work by James P. Zappen (2005), Doug Eyman (2015), and Angela Haas (2018) helpful in our endeavor. Eyman calls for scholars of digital rhetoric to combine multiple theoretical frameworks to employ during research and analysis of rhetoric in the digital sphere. In taking up his call, Zappen’s (2005) work serves in a roundabout way as a model for Eyman’s text a decade later, to focus digital rhetoric on four major areas, including the use of rhetorical strategies in production and analysis of digital text, identifying characteristics, affordances, and constraints of new media, formation of digital identities, and the potential for building social communities. Of course, while each of the areas of digital rhetoric highlighted by Zappen (2005) are afforded exploration, coupling his methodological framework with Haas’s (2018) work toward a digital cultural rhetoric provides a robust path forward for rhetorical analysis focused on digital genealogy databases and establishes what I refer to as an Identity-Attentive Digital Cultural Rhetorical Methodology.

Haas (2018) posits many attributes of digital cultural rhetorical studies as she works toward a digital cultural rhetoric. Her effective, yet not exhaustive, paradigm is applicable to the work I am doing here by responding to these attributes concerning the areas of digital rhetoric outlined by Zappen (2005). Haas (2018) defines digital rhetoric as the digital negotiation of information—and its historical, social, economic, and political contexts and influences—to affect change and outlines the following attributes of a digital cultural rhetoric which requires a negotiation—an interfacing—between bodies, identities, rhetoric, and technology: digital rhetoric requires interrogating the politics of digital interfaces, studying digital rhetorics in relation to/with specific communities and cultures of practice, examining the relationship between older and new technologies; valuing diverse bodies; and reassessing access.

With an Identity-Attentive Digital Cultural Rhetorical methodological framework in mind, we can pair an ethical framework, like virtue ethics, to better understand the actions taking place on digital genealogy databases. What we find helpful from a virtue ethics framework in constructing a public ethical framework for users of digital genealogy databases is that there are not rules to follow, or desired outcomes to reach, but, rather exemplars to emulate (Gallagher, 2018). Exemplars are people who take action, make choices, and have a sense of proper motivation with respect to a particular disposition (Gallagher, 2018). Indeed, this is quite productive given the dynamic range of User Agreements for these sites. But, how do we identify exemplars at work within these unique digital ecosystems?

To begin, we can analyze the practices of the community, as well as practices of individual users to identify ones with ethical character that has its origins not only in their individualism but also in a wider sense of social co-consent, a concept deriving from the Aristotelian idea of *telos* (Gallagher, 2018). Here, then, is the real crux of the application of virtue ethics to the actions of users of genealogy databases: patience. This will take time—beyond my lifetime, perhaps—but virtue ethics gives us a start, a leg up, even, in answering the questions John Duffy (2018) poses: “What kind of person do I want to be? How should I live my life? What does it mean to be a good person?” So, for users of digital genealogy databases, they must ask themselves: what kind of user do I want to be? How should I live my digital life? What does it mean to be a good user? Complex questions, for sure, with divergent reactions; yet, with a more defined conceptualization of consent, one which focuses on and challenges assumptions concerning the co-creation of digital identity among users, answers to these questions become a little less messy.

The issue of whether or not the tactics used by police to identify D’Angelo are ethical is a much larger issue which not only concerns police tactics and surveillance, but also the medical field and insurance companies, all at the intersection of Big Data. By extending collaborative discussions that promote intersectional resistance to marginalization and oppression, we work towards a public ethical framework for users building their identity, and the identity of others, on digital genealogy databases. There are many questions concerning how digital genealogy databases make use of the data users upload to their sites. If these companies are subsidiaries of large corporations, what information are they sharing? What is the motivation to share specific information? For example, what are the implications of these companies sharing individual medical histories with insurance companies?

We must continue to press our police agencies for ethical transparency regarding their tactics and motives for solving crimes and catching criminals. We must hold them responsible for their actions. We must continue to complicate the function of online genealogy databases, considering how these sites could potentially be used for surveillance. We must consider the rhetorical velocity and trajectory of our actions on these sites, specifically the data we share, as well as the potential actions of other users. We must consider how ethical work is done in these digital spaces. Online genealogy databases are comprised of infinite amounts of information, and are primed to aid police agencies, Big Data, and other public and private organizations. A public ethical framework is one way to center user ethics as this technology continues to evolve.

Works Cited

Edwards, D. (2018). Circulation gatekeepers: Unbundling the platform politics of youtube’s content id. Computers & Composition: An International Journal, 47, 61-74.

Eyman, D. (2015). *Digital rhetoric: Theory, method, practice*. Ann Arbor, MI: U of Michigan P.

Duffy, J., Gallagher, J., & Holmes, S. (2018). Introduction. *Rhetoric Review,* 37, 321-327.

Gallagher, J. (2018). Enacting virtue ethics. *Rhetoric Review,* 37, 379-384.

Grabill, J. & Pigg, S. (2012). Messy rhetoric: identity performance as rhetorical agency in online public forums. *Rhetoric Society Quarterly*, 42 (2), 99-119.

Gross, A.G. (1996). The rhetoric of science. Cambridge, MA: Harvard UP.

Haas, A.M. (2007). Wampum as hypertext: An American Indian intellectual tradition of multimedia theory and practice. *Studies in American Indian Literatures Series 2*, 19 (4), 77-100.

--. (2018). Toward a digital cultural rhetoric. In J. Alexander and J. Rhodes (Eds.), *The Routledge Handbook of Digital Writing and Rhetoric* (412-22).

Katz, S. (1992). The ethic of expediency: Classical rhetoric, technology, and the holocaust. College English, 54 (3), 255-275.

Kirchner, L (2017, September 4). Traces of crime: How new york’s dna techniques became tainted. *The New York Times*. Retrieved from https://www.nytimes.com/2017/09/04/nyregion/dna-analysis-evidence-new-york-disputed-techniques.html.

McNamara, m. (2018). I’ll be gone in the dark: One woman’s obsessive search for the golden state killer. New York, NY: Harper.

Miller, C. (1979). A humanistic rationale for technical writing. *College English*, 40 (6), 610-617.

Selber, S.A. (2004). *Multiliteracies for a digital age*. Carbondale, IL: Southern Illinois UP.

Worth, K. (2015, June 24). The surprisingly imperfect science of dna testing. *Frontline*. Retrieved from https://www.pbs.org/wgbh/frontline/article/the-surprisingly-imperfect-science-of-dna-testing-2/.

Zappen, J. P. (2005). Digital rhetoric: Toward an integrated theory. *Technical Communication Quarterly*, 14, 319-325.