

Social Sentinel

Regan Hultquist

## Introduction

Modern technology is used for many lifestyle conveniences as well as being used to save lives. In the early 2000s, these two roles of technology merged when Social Sentinel (now Navigate360 Detect) was created. Since 2015, U.S. universities have used Social Sentinel, a technology monitoring company that tracks students' social media behavior for the stated purpose of identifying intentions to harm themselves or others. 37 universities are confirmed to have used Social Sentinel, although the true number of institutions may be much higher. The company argues that if used correctly, the technology can help save lives.

The idea for the technology originated in the early 2000s when a young police sergeant told the then police chief at the University of Vermont at Burlington, Gary Margolis, information regarding a series of petty thefts, a sexual assault, and a protest planned at the executive offices of the university.<sup>1</sup> When Margolis asked the sergeant where he got the information, he was shown a new online site: Facebook. Margolis realized at that moment that the campus police needed “a digital beat cop – ‘someone’ with an ear to the social landscape.”<sup>2</sup>

During his final year as police chief in 2008, Margolis began working to develop the technology. In 2012, he announced Campus Sentinel, an app that aggregated crime stats on college campuses. By 2014, Margolis had an app that kept tabs on online activity, rebranding the service as Social Sentinel, a social media threat alert service.<sup>3</sup>

---

<sup>1</sup> Arijit Douglas Sen and Derêka Bennett, “Tracked: How Colleges Use AI to Monitor Student Protests,” Pulitzer Center, September 20, 2022, <https://pulitzercenter.org/stories/tracked-how-colleges-use-ai-monitor-student-protests>.

<sup>2</sup> Arijit Douglas Sen and Derêka Bennett

<sup>3</sup> Ari Sen and Derêka K. Bennett, “Tracked: How One Company's AI Surveillance Tool Helps Colleges Monitor Protests,” The Dallas Morning News (The Dallas Morning News, September 20, 2022), <https://interactives.dallasnews.com/2022/social-sentinel/>.

The app model is fairly simple: experts in linguistics, mental health, and public safety worked together to build a library of more than 450,000 words and phrases that set off its algorithm's alarm bells.<sup>4</sup> Margolis has said that the software scans approximately one billion posts every day, from a variety of social media platforms. Social Sentinel is pitched to schools as a tool to help prevent school shootings and student suicides. To take advantage of the service, schools pay an agreed-upon subscription fee for a certain amount of time. For example, the University of North Texas signed a \$15,000-a-year contract with Social Sentinel in 2015.<sup>5</sup> Shawsheen Valley Technical High School in Massachusetts, on the other hand, paid \$10,000 for the same service.<sup>6</sup> The company is not transparent about the specific pricing model on its website, but it appears that institutions pay between \$1 and \$2 per student for the service annually.

A 2022 investigation completed by *The Dallas Morning News* and the UC Berkeley Graduate School of Journalism showed that, despite publicly saying the service was not a surveillance tool, Social Sentinel representatives promoted the tool to universities for “mitigating” and “forestalling” student protests.<sup>7</sup> In 2017, reporters found that the Kennesaw State University police department had listed a criminal intelligence analyst job posting; the role included monitoring Social Sentinel for “crime, civil unrest, protest” and more. The report sparked outrage as people dubbed the use of Social Sentinel a violation of individuals’ first amendment rights.

David Smitherman, a marketer for Social Sentinel, reached out to a UNC-Charlotte administrator to endorse the service. In 2017, he sent a whitepaper from the company titled “Demonstrations and Protests: Using Social Media to Gather Intelligence and Respond to Campus Crowds.” (See Exhibit A for the full whitepaper). The document outlines how Social Sentinel’s

---

<sup>4</sup> Lynn Jolicoeur and Lisa Mullins, “To Detect Threats and Prevent Suicides, Schools Pay Company to Scan Social Media Posts,” WBUR (WBUR News, March 22, 2018), <https://www.wbur.org/news/2018/03/22/school-threats-suicide-prevention-tech>.

<sup>5</sup> Arijit Douglas Sen and Derêka Bennett

<sup>6</sup> Lynn Jolicoeur and Lisa Mullins

<sup>7</sup> Ari Sen and Derêka K. Bennett

technology can be used to monitor demonstrations both before and after they occur. It also mentions that Michigan State University had already used the service to monitor protests.<sup>8</sup> *The Dallas Morning News* obtained emails from 2015 and 2016 that show the whitepaper was sent to at least three other colleges considering the service as well.

As a student wrote in the Harvard Crimson, “there’s a fine line between watching and watching out for.”<sup>9</sup> Many are debating on which side of the line universities using Social Sentinel fall. Is Social Sentinel’s service a case of proactive policing or an invasion of privacy?

### **Surveillance for Student Safety?**

With so much public distaste regarding Social Sentinel following the publication of the investigative report, how did the company become popular in the first place? Social Sentinel’s growth in popularity in the early 2010s had a lot to do with the increase in student suicides and high-profile school shootings at that time.

Suicide is the second leading cause of death among people ages 10-24. The rate of suicide for people of those ages has been increasing every year since 2007; according to the Centers for Disease Control and Prevention (CDC), the suicide rate increased by 60% between 2007 and 2018.<sup>10</sup> Social worker and associate professor at Simmons University Elana Premack Sandler found that many people who post online about committing suicide do not receive help, even though the post may be seen by a significant number of people.<sup>11</sup> Social Sentinel argues that its service can help address this growing problem. By flagging potentially harmful posts and alerting monitors, individuals can receive the help they need.

---

<sup>8</sup> Arijit Douglas Sen and Derêka Bennett

<sup>9</sup> “Social Sentinel and the Creeping Surveillance University: Opinion,” *The Harvard Crimson*, October 18, 2022, <https://www.thecrimson.com/article/2022/10/18/editorial-surveillance-social-sentinel/>.

<sup>10</sup> “Even before Covid-19 Pandemic, Youth Suicide Already at Record High,” *UC-Davis Health*, March 6, 2022, <https://health.ucdavis.edu/news/headlines/even-before-covid-19-pandemic-youth-suicide-already-at-record-high/2021/04>.

<sup>11</sup> Elana Premack Sandler, “Can Social Media Help Prevent Suicide?,” *Psychology Today* (Sussex Publishers, April 6, 2009), <https://www.psychologytoday.com/us/blog/promoting-hope-preventing-suicide/200904/can-social-media-help-prevent-suicide>.

Supporters of Social Sentinel also argue that it can help prevent violence on school grounds, particularly school shootings. According to a federal report that examines crime and safety in schools across the United States, school shootings in 2020-21 soared to the highest number in over two decades.<sup>12</sup> In June 2022, a doctorate student at Marymount University studied how effectively mass shootings could be predicted by monitoring social media. Analyzing 500 Twitter posts shortly before and after five mass shootings in recent U.S. history, both manually and using AI, Nnaemeka Ekwunife found that about 66 percent of the tweets pointed to the possibility of an incident or described the incident in almost real-time.<sup>13</sup> It may be no surprise then that Texas has implemented digital surveillance technologies, like that offered by Social Sentinel, in more school districts than any other state while also being the state with the most school shootings over the past decade.<sup>14,15</sup>

However, while the company's advertising may be effective, some argue that the technology itself is not. For instance, even with Social Sentinel's services, school shootings take place. It is confirmed that the Uvalde school district purchased monitoring services from Social Sentinel in 2019-2020, although it is unclear whether the subscription was still in effect in May 2022, when the Robb Elementary School shooting took place.<sup>16</sup> Whether or not the subscription was active in 2022, Social Sentinel faced extreme scrutiny for Uvalde having contracted with the company previously.

---

<sup>12</sup> Donna St. George, "School Shootings Rose to Highest Number in 20 Years, Federal Data Says," The Washington Post (WP Company, July 4, 2022), <https://www.washingtonpost.com/education/2022/06/28/school-shootings-crime-report/>.

<sup>13</sup> Nicholas Munson, "Alumnus Completes Study on Predicting Mass Shootings from Social Media," Marymount University, June 7, 2022, <https://marymount.edu/blog/alumnus-completes-study-on-predicting-mass-shootings-from-social-media/>.

<sup>14</sup> Corin Faife, "After Uvalde, Social Media Monitoring Apps Struggle to Justify Surveillance," The Verge (The Verge, May 31, 2022), <https://www.theverge.com/2022/5/31/23148541/digital-surveillance-school-shootings-social-sentinel-uvalde>.

<sup>15</sup> Sharon Lurye, "States with the Most School Shootings," U.S. News & World Report (U.S. News & World Report, May 27, 2022), <https://www.usnews.com/news/best-states/articles/2022-05-27/states-with-the-most-school-shootings>.

<sup>16</sup> Peter D'Auria, "In the Wake of Texas School Shooting, a Vermont-Founded Company Draws Scrutiny," VTDigger, June 7, 2022, <https://vtdigger.org/2022/06/07/in-the-wake-of-texas-school-shooting-a-vermont-founded-company-draws-scrutiny/>.

Another problem is that the technology only draws upon data found in the public domain of social media. According to the technology news website The Verge, 98% of the data Social Sentinel collected in 2019 came from Twitter because of its conventions as a mostly public platform, even though younger demographics use Instagram, Snapchat, and even Facebook more.<sup>17</sup> When potentially harmful posts are on private Instagram, Snapchat, or Facebook accounts, Social Sentinel’s technology does not flag them. While this feature arguably makes the service less invasive, it also makes the service less effective.

These are several criticisms of Social Sentinel’s stated use of preventing suicides and school shootings. However, the service was largely accepted until *The Dallas Morning News* published its report outlining the service’s role in monitoring student protests. This raised arguably more severe ethical questions.

### **Protests at the University of North Carolina**

There is evidence that at least 37 U.S. universities have used Social Sentinel, with some using it beyond the intended purpose of preventing suicides and school shootings. One of these universities was the University of North Carolina-Chapel Hill (UNC). In 2018, Social Sentinel was used in monitoring protestors opposing the Confederate statue Silent Sam.<sup>18</sup> (See Exhibit B for photos of the protest). Silent Sam is a bronze statue of a Confederate soldier by Canadian sculptor John A. Wilson, which stood on McCorkle Place of UNC from 1913 until 2018. McCorkle Place is the “front door” of the university, meaning that the statue was given “a position of honor.”<sup>19</sup> The statue was the subject of protests for decades until it was finally pulled down by protestors on August 20, 2018.

---

<sup>17</sup> Corin Faife

<sup>18</sup> Aisha Baiocchi, “University Used Social Media Monitoring in Silent Sam and pro-Choice Protests,” The Daily Tar Heel, September 28, 2022, <https://www.dailytarheel.com/article/2022/09/university-social-sentinel>.

<sup>19</sup> “Silent Sam,” Wikipedia (Wikimedia Foundation, November 7, 2022), [https://en.wikipedia.org/wiki/Silent\\_Sam](https://en.wikipedia.org/wiki/Silent_Sam).

One of the authors of *The Dallas Morning News* investigative article was UNC alum Ari Sen, who attended the university from 2016 to 2020. Sen was a student at UNC during the Silent Sam protests. In an interview with current UNC students, Sen described how the protests led him to Social Sentinel: “I started talking to the activists,” he said. “And what I was hearing from them was, ‘we feel like we’re being put on a list. We feel like we’re being surveilled.’”<sup>20</sup> He started requesting documents from the university and got thousands of pages of documents back, including a contract for the Social Sentinel service. He found that in 2018, around the time of the protests, UNC Police entered specific keywords related to Silent Sam into the Social Sentinel service to gather information about the statue protests.<sup>21</sup>

After Sen published articles about UNC’s use of Social Sentinel in monitoring student protests, UNC issued a statement:

The University uses Social Sentinel to identify threatening or concerning public social media posts surrounding events and campus activities that require campus public safety. The University does not use the service to monitor student email. The resource is an industry best practice, only uses public information, and does not collect private personal information. The University coordinates with and relies on assistance from state and federal law enforcement agencies when appropriate to ensure campus safety. The University’s current contract with Navigate360 ends in October [2019] and the contract will not be renewed.<sup>22</sup>

According to the documents obtained by Sen, UNC used geofencing and the Social Sentinel’s Local+Lists features during the Silent Sam protests in 2017 and 2018.<sup>23</sup> Geofencing describes technology that creates a virtual geographic boundary and triggers a response when a

---

<sup>20</sup> Aisha Baiocchi

<sup>21</sup> Aaron Keck, “Social Media Tracking Tool, Used By UNC, Raises Speech and Privacy Concerns,” Chapelboro.com, September 22, 2022, <https://chapelboro.com/news/unc/social-media-tracking-tool-used-by-unc-raises-speech-privacy-concerns>.

<sup>22</sup> Cullen Browder, “Seven North Carolina College Campuses Have Secretly Tracked Social Media Posts of Students, Protesters,” WRAL.com (WRAL News, September 20, 2022), <https://www.wral.com/artificial-intelligence-policing-exposed-on-seven-north-carolina-college-campuses/20484988/>

<sup>23</sup> Aisha Baiocchi

cell phone enters that area. Social Sentinel uses Local+Lists to give localized alerts, using a customizable AI search engine to scan public posts for specific language and phrases and consequently alert monitors.<sup>24</sup>

Emails show that geofencing for specific events and using local algorithms were two features of Social Sentinel that originally persuaded UNC to employ the service. Uncovered contracts show that UNC paid \$24,500 annually to use Social Sentinel’s service between the fall of 2016 to 2019. Documents show that UNC used Social Sentinel’s geofencing and algorithmic services in August 2017 to acquire statistics about mentions of Silent Sam and went as far as to single out content. Names and graduating years of staff, faculty, and alumni that tweeted about the statue and its removal were found listed in emails.<sup>25</sup>

When North Carolina Representative Graig Meyer (D-Orange County) read UNC’s contract with Social Sentinel, he found the wording extremely concerning. He said, “When you don’t have a solid policy that regulates the way that technology intrudes on private lives then you’re going to have things that intrude in a way that is threatening or harmful to people.”<sup>26</sup>

## **Ethical Questions**

Much of the discomfort and tension brought up when Social Sentinel’s activities came to light in October mirror the discomfort and tension that arose after Edward Snowden’s leak in 2013. The words “creepy,” “Big Brother,” and “stalking” are used regularly by those who worry about their personal information being collected online.<sup>27</sup> The investigative report again brings to mind

---

<sup>24</sup> Aisha Baiocchi

<sup>25</sup> Ari Sen, “UNC Campus Police Used Geofencing Tech to Monitor Antiracism Protestors,” NBC News (NBC News, December 21, 2019), <https://campussafetycommission.web.unc.edu/wp-content/uploads/sites/20749/2020/01/UNC-campus-police-used-geofencing-tech-to-monitor-antiracism-protestors.pdf>.

<sup>26</sup> Cullen Browder

<sup>27</sup> Lee Rainie and Maeve Duggan, “Privacy and Information Sharing,” Pew Research Center (Pew Research Center, January 14, 2016), [https://www.pewresearch.org/wp-content/uploads/sites/9/2016/01/PI\\_2016.01.14\\_Privacy-and-Info-Sharing\\_FINAL.pdf](https://www.pewresearch.org/wp-content/uploads/sites/9/2016/01/PI_2016.01.14_Privacy-and-Info-Sharing_FINAL.pdf).

the question: how much privacy, if any, should be given up in the name of public safety and security? Who should that privacy be given up to? The government, a university?

Protests can potentially threaten public safety, causing injury or loss of life, particularly if they get out of hand. Students assume universities have their best interests in mind and will support their intellectual endeavors, including peaceful protests. Where is the line between keeping students safe and hosting free expression?<sup>28</sup> Margolis has said, “we are not a surveillance tool; we are not a monitoring tool; we’re not an investigative tool.”<sup>29</sup> If Social Sentinel is none of these things, then what is it?

### **Are Public Posts Private?**

The fact that the service only draws from public posts on social media raises other questions. If the posts are in the public domain, is there an invasion of privacy? According to Sen, there are two concerns here: “one, we found that no college affirmatively told students about this technology...there was a real lack of transparency. And...what the experts I talked to really feared was, one, that students were going to be chilled, that they were going to be afraid to express themselves online. And two: police might be able to track networks...and we don’t know what the police department might do with that information later on.”<sup>30</sup>

Several studies show that digital surveillance can hurt mental health, even if that surveillance is only of publicly available data.<sup>31</sup> Experts say that “even collecting publicly posted information may harm students’ privacy and speech rights.”<sup>32</sup> Since 2019, records show that Social

---

<sup>28</sup> “When Safety Becomes Surveillance Colleges Track Students Using AI Designed to Monitor Mental Health,” The Batch (Deeplearning.ai, November 9, 2022), <https://www.deeplearning.ai/the-batch/colleges-track-students-using-ai-designed-to-monitor-mental-health/>.

<sup>29</sup> Lynn Jolicoeur and Lisa Mullins

<sup>30</sup> Aaron Keck

<sup>31</sup> Saumya Kalia, “What Is a Constant Lack of Digital Privacy Doing to Our Mental Health?,” The Swaddle, January 26, 2022, <https://theswaddle.com/what-is-a-constant-lack-of-digital-privacy-doing-to-our-mental-health/>.

<sup>32</sup> Arijit Douglas Sen and Derêka Bennett



Sentinel has offered colleges the ability to monitor private student emails.<sup>33</sup> A large portion of people agree that this is a violation of privacy even if looking at public posts is not. However, many companies already do both of these things. In fact, most of those companies go beyond monitoring the content of emails; for example, Gmail links the following data to you: purchases, location, contact info, photos or videos, audio data, search history, user ID, and device ID, among other types of data.<sup>34</sup> Are these two types of monitoring different?

On one hand, universities may physically act on the information that they receive from the data collected by sending police to protest sites, whereas tech giants like Google use the data in less tangible ways: by targeting advertisements. Another difference is that most online companies get consent before collecting data. Whether that consent is true consent is up for debate. In fact, ProPrivacy.com, a digital privacy group, found that only 1% of subjects in a social experiment read ‘terms and conditions.’ The other 99% of respondents agreed to absurd things in the terms and conditions like permission to give their mother full access to their browsing history and the name rights of their first-born child.<sup>35</sup> Informed consent says that “participants should be presented with relevant information in a comprehensible format and then should voluntarily agree to participate.”<sup>36</sup> Do the terms of the agreement include relevant information in a comprehensible format? This is debatable.<sup>37</sup> Given the situation with Social Sentinel, is this question even an applicable one? Should consent be an important part of the discussion regarding Social Sentinel, or is consent not necessarily an issue given that they are using publicly available data?

---

<sup>33</sup> Ari Sen and Derêka K. Bennett

<sup>34</sup> Serena O'Sullivan, “Gmail Reveals the Massive Amount of Data It Has on You,” Komando.com, May 6, 2021, <https://www.komando.com/security-privacy/data-google-has-on-you/780537/>.

<sup>35</sup> Tim Sandle, “Report Finds Only 1 Percent Reads 'Terms & Conditions',” Digital Journal, January 29, 2020, <https://www.digitaljournal.com/business/report-finds-only-1-percent-reads-terms-conditions/article/566127>.

<sup>36</sup> Ivanfanta, “Do We Actually Agree to These Terms and Conditions?,” Data Science W231 - Behind the Data Humans and Values (UC-Berkeley, July 20, 2021), <https://blogs.ischool.berkeley.edu/w231/2021/07/09/do-we-actually-agree-to-these-terms-and-conditions/>.

<sup>37</sup> Ivanfanta

## Disparate Impact

Another troubling aspect of Social Sentinel’s technology is that it may disproportionately flag posts by people of color.<sup>38</sup> First of all, marginalized groups may feel intensified effects of social media monitoring. Students who are already at a higher risk of running into the law or being surveilled may distrust social media monitoring. This could lead to increased self-censorship on platforms. The variability in the accuracy rates for social media monitoring tools, especially for minority speakers (e.g. people posting in languages other than English or people using vernacular associated with a subgroup) may also contribute to this. Even when these flagged posts are reviewed by humans, English-speaking reviewers might still misunderstand the meaning of these posts.<sup>39</sup>

A serious problem is that Social Sentinel and other social media monitoring companies treat their algorithms as proprietary. It is unclear what training data they use in their models, or whether that training data has been audited for racial bias. In 2017, some UMass-Amherst researchers put African American Vernacular English into language identification machine learning models. While it should identify the language as English, one of the models flagged the language as Danish with 99 percent confidence.<sup>40</sup> This shows that AI language-processing models are not always reliable with non-Anglicized English text and may exhibit biases from training data. Social Sentinel claims on their website, “We don’t perpetuate any biases.” However, it is very difficult to promise this if the underlying models have these sorts of biases built in.

---

<sup>38</sup> Lizzie O’Leary, “Why Expensive Social Media Monitoring Has Failed to Protect Schools,” Slate Magazine (Slate, June 4, 2022), <https://slate.com/technology/2022/06/social-media-monitoring-software-schools-safety.html>.

<sup>39</sup> Michael German et al., “Social Media Monitoring in K-12 Schools: Civil and Human Rights Concerns,” Brennan Center for Justice, November 22, 2022, <https://www.brennancenter.org/>.

<sup>40</sup> Will Knight, “AI Programs Are Learning to Exclude Some African-American Voices,” MIT Technology Review (MIT Technology Review, April 2, 2020), <https://www.technologyreview.com/2017/08/16/149652/ai-programs-are-learning-to-exclude-some-african-american-voices/>.

## Looking to the Future

In 2022, Social Sentinel rebranded itself Navigate360 Detect. While the company called this an independent move, it likely had to do with the bad press surrounding Social Sentinel after *The Dallas Morning News*' report. On its website, the company says it is working to develop its AI further. Universities have reported that a significant amount of alerts that they receive from the Social Sentinel service were false alerts. It flags posts including song lyrics, Bible verses, and obvious jokes. On social media, people tend to speak very sarcastically and ironically. Hyperbolic language can be difficult for machine learning models to catch. Right now, only 0.005% of scans are actionable. The company hopes to refine the technology to eliminate these types of false positives.

There is a larger political conversation to be had surrounding gun regulation and mental health programming. Right now schools are desperate to do something to protect their kids, whether that's school safety drills, monitoring services like Navigate360 Detect, or physical security like metal detectors or security guards. The Uvalde shooting showed that all of those initiatives can fail and do fail. Therefore, a different conversation about what measures are effective and what new approaches can be implemented are needed to tackle these problems.

The technology itself is not invasive, but it can be used for invasive ends. Navigate360 now has an 'Ethical and Responsible AI' section on its website, which seems to be a step in the right direction. However, as the investigation showed, the company has been less than transparent. Beyond schools, digital surveillance is also used by law enforcement, the government, and even private companies on their employees.<sup>41</sup> Digital technology is not designed to do harm, yet a failure

---

<sup>41</sup> Steven Feldstein and David Wong, "New Technologies, New Problems – Troubling Surveillance Trends in America," Just Security, August 6, 2020, <https://www.justsecurity.org/71837/new-technologies-new-problems-troubling-surveillance-trends-in-america/>.

to establish clear, enforceable regulations regarding the operation of new surveillance technologies will fail to protect people's rights. More must be done to balance the desire to protect students using AI and the need to protect privacy interests and the right to free speech.

## Appendix:

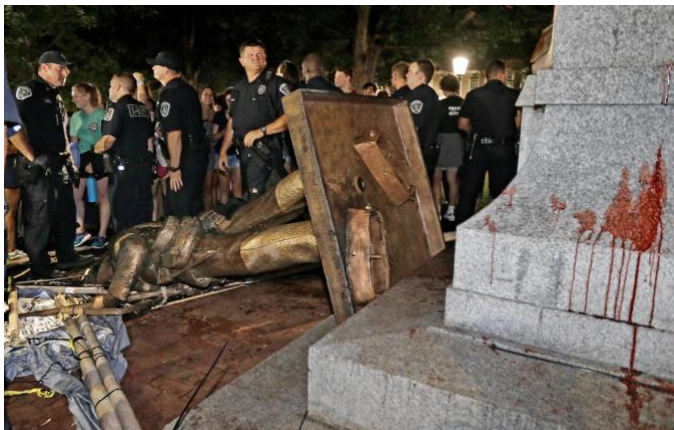
### Exhibit A: Social Sentinel Whitepaper

Source: [https://www.documentcloud.org/documents/22274805-ssi\\_hed\\_protest\\_whitepaper](https://www.documentcloud.org/documents/22274805-ssi_hed_protest_whitepaper)

This whitepaper from the company titled “Demonstrations and Protests: Using Social Media to Gather Intelligence and Respond to Campus Crowds” outlines how Social Sentinel can be used to monitor before and during demonstrations. This whitepaper was sent to at least four universities, likely more.

### Exhibit B: Photos of UNC’s Silent Sam Protests

Below are some photos of the protests against the Silent Sam protest on UNC’s campus in August 2018. It is confirmed that UNC Police used Social Sentinel’s services during these protests to gather information regarding the demonstration and the people involved.



Levenson, Michael. “Toppled but Not Gone: U.N.C. Grapples Anew with the Fate of Silent Sam.” The New York Times. The New York Times, February 14, 2020. <https://www.nytimes.com/2020/02/14/us/unc-silent-sam-statue-settlement.html>.



Keck, Aaron. “Social Media Tracking Tool, Used By UNC, Raises Speech and Privacy Concerns.” Chapelboro.com, September 22, 2022. <https://chapelboro.com/news/unc/social-media-tracking-tool-used-by-unc-raises-speech-privacy-concerns>.