Trans Safe Spaces on Social Media

George Birch

April 19, 2021

Abstract

Transgender social media users face extraordinary risks and threats online, yet also reap many benefits from social media use. To understand how encryption technology may or may not be able to benefit trans social media users by providing an added layer of security, we provided eight transgender Facebook users with an encryption tool to be used on Facebook and conducted semi-structured interviews with five of the participants after they had used the tool for a few weeks. These interviews covered both the participants' backgrounds on social media as well as their experiences with the tool. We analyzed these interviews to understand the accomplishments and failings of the security tool to create a sense of safety online, and discovered a wide range of experiences that reflected the complex ways that a sense of safety can be both built and damaged for a vulnerable population.

Introduction

In the United States, over 1.4 million people identify as transgender, meaning their gender identity defers from the one they were assigned at birth. People who identify as transgender are even more prevalent in younger age ranges, which are also groups more likely to be active on social media. As a result of this as well as factors that will be discussed more below, trans people have an active presence in online communities, and a wealth of research explores this phenomenon.

Lerner et al. [19] provided a comprehensive snapshot of how trans people enjoy many benefits from social media usage but also face severe threats of harm online and resultingly go about online interactions with an unusually high degree of risk and emotional stress. Because of this, we can expect that creating safer online experiences for this population has the potential to ameliorate these stresses and anxieties while also allowing better access to the benefits of social media use. One suggestion proposed by Lerner et al. was to further explore how technological solutions that harness the power of crytography could accomplish these goals.

This introducing a promising area of research since cryptographic technologies have the power to provide strong security while also having the potential to be easily implemented on top of the medium provided by existing social media platforms. While the companies behind major social media platforms are often deaf to calls for better privacy and security for their users in interest of profits, open source and/or community developed security software has the potential to provide for vulnerable users' needs without compromise.

The goal of such technology should not simply be to protect trans users, but also make them *feel* protected so that they can engage in social media usage openly and freely. This kind of experience that is desired is most succinctly expressed as that of a "safe space", which in this work is defined as a space where users are free from physical, verbal, and emotional harm (especially that relating to their identity) as well as stress or worry of it, and furthermore where they can speak and organize freely.

In this work, we explore in particular the effects for trans users of a security tool that encrypts communications over a social media platform, in this case Facebook. The tool used is an original implementation described below that borrows its design from the plethora of research studying the use of cryptography on social media. The eight participants of the study, in two groups, were left to use the tool to communicate with each other on Facebook for a period of a few weeks. At the end of this period, semi-structured interviews were conducted on five volunteering participants about their experiences with the tool as well as background on their typical social media usage.

This work explores the findings of these interviews, which cover participants' understandings and perceptions of risk on social media, successes with the security tool, struggles with the security tool, and more. In our analysis, we identify how the risk models participants employ to understand threats of harm online inform their attitudes towards the security tool, the ways that knowledge and understanding of online privacy and security facilitate or disrupt mechanisms to protect trans users, and discuss the potential for security technology to adapt to the rapidly changing social media landscape and needs of trans users.

Related Works

This study builds off of prior research both in the experiences and needs of transgender people online as well as in mechanisms to increase privacy and security on social media. In this section, a brief overview will be given of relevant prior work in both these areas.

Trans Communities on Social Media

Online spaces provide unparalleled opportunity for groups of people largely excluded from mainstream public spaces to connect, and transgender communities are no exception. The availability of these spaces for transgender people has been shown to provide many important benefits. For example, social media can greatly reduce stress during gender transition by facilitating support from one's network [14], be used to make connections with strangers who have similar experiences [23], make important health information accessible [2], provide sites to engage in activism [19], and even help keep transgender users physically safe [23].

Before changes to the site in 2018, Tumblr in particular epitomized many of these benefits. By allowing mutability of identity, separation from offline social networks, posting of erotic content, and much more, the site became a valuable space for transgender users [15, 7]. There even exists a social media site made specifically for transgender users called Trans Time, demonstrating the importance of this medium of connection for many transgender people [12].

Trans Security and Privacy Challenges

Despite this host of examples of the beneficial potential of social media platforms for transgender users, risks and harms stemming from their usage are also experienced disproportionately by these users. Trans users often experience stress while using social media, and platform design can both directly do harm and facilitate harmful usage by other users [14].

Harmful aspects of platforms include a variety of characteristics that were unlikely to have had their impact on trans users considered by cisgender designers. For example, Haimson et al. [13] found that the often permanent or difficult to destroy record of one's identity created by platforms – such as photos, gender markers, names, etc. – can become problematic for trans users who are transitioning and wish to remove these public traces of a past identity. Bivens and Haimson [6] explored how assumptions about gender are embedded into platforms, such as when gender categorization systems (which are almost impossible to avoid online because of their usefulness in generating advertising revenue) disallow non-binary identities or the ability for users to update their present gender identity. Erotic content or medical educational content relating to trans bodies or surgery is frequently banned by platforms seeking to create "family friendly" brand images who categorize it as pornographic, creating a perception among users that platforms could betray their trans communities and effectively purge them at any time, such as the sentiment found in response to Tumblr by Haimson et al. [15].

While platforms themselves in isolation can do harm to trans users, it is often the ways in which they create environments where trans people are susceptible to harm from other users that make them unsafe. Lerner et al. [19] and Scheuerman et al. [23] found that trans social media users report feeling unsafe due to the fear of attacks such as blackmailing, doxxing (the malicious public disclosure of private information about a person), hateful content, being outed (have their queer identity revealed) to people who know them offline, being fetishized for their trans identity, being stalked online or in real life, and more. In Scheuerman et al.'s study, trans social media users reported these threats coming from strangers from outside their online circles as well members of their own online communities, and even content being passively delivered to them by platforms could be harmful.

It is important to note that while these may seem like ubiquitous negative consequences of being online in the present day, there are specific aspects of platform design that facilitate the potential for harm. For example, the searchability of trans users by abusive users makes attacks easy [23], and many trans users specifically consider their online visibility to be a risk factor [19]. This is made worse by what Alexander Cho [7] refers to as "default publicness," a design stance driven by common business models that builds platforms around information about users being public and searchable, making maintaining some degree of anonymity or privacy at the best require active steps to be taken by the user and at worst be impossible. Hanckel et al. [16] note that the extreme promotion of visibility inherent in the design of many platforms also promotes the visibility of transphobic or otherwise hateful content to users.

Social Media Security and Privacy Enhancements

While the above section demonstrates that trans social media users are at risk of a large variety of harm, there has also already been substantial research into how modifications to the social media experience can meet privacy and security concerns of a general user population. Because technology design facilitates or prevents abusive interactions for trans users [23], we can expect that some of these design modifications might help alleviate the dangers and difficulties that trans social media users face.

Some research in this body of work examines how the surface level design of websites could be different to push users to better maintain their privacy. Fang and LeFevre [9] developed and user tested a

tool that uses machine learning technology to generate an accurate understanding of a user's community and then gives the user recommendations on how to adjust their privacy settings accordingly. Wang et al. [25] similarly developed and user tested a tool that "nudges" Facebook users to reconsider sharing a post by reminding them of who their audience is or simply by delaying the post to prevent impulsive sharing. Approaches such as these are not just hypotheticals – social media users already take their privacy concerns into their own hands by employing their own tactics such as using nicknames of pseudonyms instead of their real names [24] and using different websites for different purposes to control their audiences [26]. Among trans users on Facebook in particular, some utilize strategies to limit their sharing of information such as maintaining lists of trusted friends to allow post viewing access to, unfriending or blocking other users, or even deleting their account [14]. The site Trans Time referenced above can also be viewed as a privacy and security enhancing measure developed by people in the trans community.

Unfortunately, such surface-level approaches have limitations. Social media companies are reluctant to implement design changes that decrease users' sharing and the openness of their information since this directly contradicts their typically data collection and advertising-based business models. Furthermore, even if users adopt these types of strategies themselves, they don't make fundamental changes to a social media ecosystem that leaves troves of personal information in the hands of technology companies that users may not trust, and they often require a high degree of constant vigilance and knowledge of the mechanisms of sites just to maintain a basic mitigation of vulnerability. Accordingly, a large body of research examines how users can guarantee their privacy and security on social media sites by wielding the ultimate tool for this purpose – cryptography. A variety of encryption schemes designed to be embedded into social media usage have been imagined and even implemented. These employ a variety of tactics, including advanced group management or attribute-based encryption to dynamically control sharing, hiding the fact that encryption is taking place to avoid detection by disapproving platforms, decentralized storage of content, and more. Typically, these tools embed themselves into the user experience in the form of applications and web browser extensions that perform encryption and decryption of content as the user posts and receives content on a particular social media website. Examples of such technology include Scramble [4], Hummingbird [8], Frientegrity [10], NOYB [11], EASiER [17], DECENT [18], FaceCloak [20], Cachet [22], Persona [3], and HITC [1]. This study uses an original implementation of an encryption scheme based on that of Ion et al. [5] in their paper "For Some Eyes Only: Protecting Online Information Sharing". This scheme was chosen because of its lack of platform dependence, ability to hide the fact that encryption is taking place, and flexibility of implementation. The simplified version implemented in this study will be thoroughly detailed below.

While it has been established by Lerner et al. [19] that some trans users are aware of the benefits of encrypted communication and use tools such as Signal (an encrypted messaging service popular among populations such as activists), there has not yet been usability studies to see if any of the above encryption tools for use on social media websites would be useful to trans users. In fact, research on these tools typically focuses just on their security capabilities, and rarely are there usability studies for even a general population. Given the simultaneous importance of social media usage and prevalence of risks involved in it for many trans people, there is substantial reason to hypothesize that some might benefit from social media encryption software.

Encryption System

The encryption system is a modified version of that developed by Ion et al. in their 2012 paper "For Some Eyes Only: Protecting Online Information Sharing". Our study exists primarily to understand the benefits of general social media encryption tools for transgender users, and thus this section serves to establish a basic understanding of the encryption tool and the differences between it and Iulia et al.'s original that it is based on rather than a complete description of its functioning. For a detailed security analysis, please refer to their paper.

Threat Model, Assumptions, and Goals

Transgender social media users may have a wide variety of individual privacy concerns and fears in regard to their sharing of online information. To help alleviate these, this tool protects all plaintext posted to a social media platform of choice from being viewed by an attacker who has control over the communication channels used to post, an attacker who hacks into the user's account or otherwise gains unwanted access, and even the platform itself. However, we assume that the attacker does not have access to the users' private keys or their physical devices as well as the software that runs on them.

Platforms often have rules that would disallow the posting of ciphertext, such as those controlling lengths of posts, those disallowing posting of encrypted material, or even automatic spam detection and removal systems. Therefore, it is necessary that encrypted posts to the social media platform appear as real posts, referred to onwards as "dummy texts".

To preserve the functionality of the social media platform, the encryption system still uses the platform as a medium to make posts, leave comments, and interact with other users in any way. Users are able to make text posts, which also indirectly allows for

the posting of images by using a third-party host and posting the link (note that this does not provide encryption of the image file but still protects the association of the post with the image link). Since the system needs to be used on both the sending and receiving ends, it is suited for situations in which a group of users agree together to use the tool and then use a secure channel to establish secret keys.

Overview of Design

To accomplish these goals, the system functions through a Chrome web browser extension that acts as a middleman between the user and the social media platform (in the case of this study, Facebook), generating a dummy text for the user to actually post on the platform when they wish to make a post, and replacing incoming dummy texts with the decrypted plaintext they are associated with. These dummy texts are not themselves ciphertexts (at least not in the typical sense of the word), but rather placeholders that map to the actual ciphertexts, which are hosted on a third-party platform. This third-party platform does not need to be trusted as it can only access encrypted data, and in this study the data is stored on Google Firebase.

Sending Posts

To make a post, the user types the post into the web browser extension's pop-out dialogue, which returns a dummy text to the user for them to actually post onto the social media platform. To do this, the extension first selects a dummy text by grabbing the first 248 characters of a random Wikipedia page (this length was chosen to prevent Facebook from disrupting the system by hiding part of the post). Additionally, the ciphertext is prepared by encrypting the plaintext using AES-256 in CBC mode using a private key stored locally on the user's device. Finally, these two pieces are linked by uploading the ciphertext to Google Firebase, filed under the SHA-256 hash of the dummy text.

Receiving Posts

The web browser extension automatically decrypts posts when they appear on the user's webpage. For this simple implementation, the dummy texts are denoted by having "@@" at the beginning and end of them to make their detection easier, though in a non-research environment it would be wise to use a more subtle detection technique and remove this indicator that makes the dummy text stand out.

When the browser extension sees a dummy post, it first grabs the ciphertext that it represents by computing the SHA-256 hash of the dummy text and using that to lookup the ciphertext on Google Firebase. Note that any attacker could perform this step without the user's secret key, but, since this only leads to the ciphertext, the only problem it causes is revealing the fact that the dummy text is actually some encrypted post. Once the ciphertext has been obtained, the extension again uses AES to decrypt it and seamlessly replaces the dummy text with the plaintext on the user's webpage.

Key Management

Since social media interactions involve sharing posts with sizeable and dynamic groups of users, key management and encryption/decryption access is often a major challenge in social media encryption tools. Since this study seeks to focus on the user experience of an encryption tool in a controlled environment, a simplified key management scheme is employed. Since participants use the tool in Facebook groups with static membership, all members of each group are simply provided with a symmetric group key used for both encryption and decryption. This scheme would cause problems in real world scenarios where group membership is dynamic, but this is outside the scope of this study. These keys are generated and distributed by the researcher.

Methods

Study participants were provided with the encryption tool as a Chrome web browser extension, given time to use it in groups, and then optionally interviewed over the phone. The period given to use the tool was 2.5 weeks for one group and 3 for the other, with the discrepency being due to technical issues with installation that delayed one group more than the other. The interviews were recorded and then transcribed and coded for analysis.

Recruitment and Demographics

Participants were recruited through postings on Oberlin College related Facebook pages. They were encouraged to sign up in groups (for example a friend group or a club/organization), but were offered to be placed in groups with other participants if they wished to sign up by themselves. The study was advertised as open to anyone who identifies as transgender with no restrictions for non-binary, genderqueer, or other identities.

In total, 9 participants signed up for the study, with 6 signing up as a group and 3 signing up as individuals, who were then placed into a group together. One participant from the first group dropped out during the study. Five participants opted into the optional interview portion of the study at the end.

At the beginning of the study, all participants were required to fill out a survey covering demographics as well as their typical social media usage. This survey emphasized open responses to reflect the inap-

plicability of mainstream categorizations of identity that queerness often entails. A table of the results is shown below. All participants were Oberlin College students.

	Gender	Sexual Orientation	Race and/or Ethnicity	Age
P1	Questioning... non-binary/demigirl/genderflux?	Queer/bisexual	White	21
P2	queer and/or trans	queer	White	21
P3	non-binary	queer	White	21
P4	lol idk	queer	White	21
P5	genderfluid	asexual	Latinx and/or Hispanic; White	19
P6	Questioning/Feminine	Bisexual	White	20
P7	Non-Binary	Questioning	White	23
P8	Non-Binary	Bisexual	White	21

Interview Methodology

Interviews were semi-structured, with a base set of questions designed to gain context into the participant's background experiences on social media, especially those to do with privacy, security, and their transgender identity, as well as their experience using the tool. Questions were designed to be open-ended, and participants were encouraged to explore their own personal thoughts, feelings, and understandings about any relevant subjects that arose. For the portion of the interview asking about experiences with the social media tool, questions were worded neutrally and negative responses were welcomed to help alleviate any potential pressure participants felt to report positive experiences with the security tool. All five interviews were conducted by the author over the phone due to Covid-19 precautions. The interviews ranged in length from 14 to 35 minutes, with an average length of 23 minutes.

Analysis

The author analyzed the interviews by listening back through them, taking thorough notes and beginning to mark codes that represented recurring concepts. After two passes of this process, the codes were reorganized and grouped into categories. Finally, each category was reviewed by revisiting each interview segment marked with a code in that category, and overarching themes and conclusions for the category were established out of this.

Findings

Interview questions covered not only participants' experiences with the study's security tool, but also background on their social media use and relationship with online privacy and security. Unsurprisingly, participants expressed a wide range of opinions and attitudes in both areas, reflecting the diverse and unique experiences of trans individuals online. In this section we explore some of the most major themes that arose from the interviews, beginning with participants' social media backgrounds.

Background

To determine whether or not the encryption tool in the study met the privacy and security needs of participants, it is necessary to begin by understanding what those needs are. Lerner et al. [19] provide the most comprehensive and recent look into the transgender social media experience through a privacy and security lens, covering trans users' goals for technology use, perceptions and understandings of threats, and actions to mitigate them. Much of the motivation for this study is based off of the findings of the Lerner et al. study and those authors' recommendation of cryptographic technology as a possible tool to help trans social media users, and thus background questions in this study were based on their findings.

These questions focused on participants' goals for social media use as well as their privacy and security needs, though interviewees also frequently delved into deep and nuanced frameworks of perceiving, understanding, and combating online threats.

Goals for Technology Use

Lerner et al. identified activism, representation/visibility, and political discourse as major goals for trans social media users' online activity. All three of these appeared among this study's interviewees, with activism in particular standing out prominently. Almost all (4/5) participants interviewed reported engaging in activism on social media, both with and without a focus on transgender issues specifically.

All four of these participants mentioned engaging in activism on Instagram in particular, which they utilized to access and share mutual aid networks, petitions, and educational material on social issues. While often recognizing limitations of this model of activism, a common sentiment among interviewees was that this platform was in particular useful for being on the receiving end of resources and actions being compiled and organized by other activists. One participant explained:

> *A lot on Instagram, people will post petitions... I follow a lot of environmen-*

tal accounts that talk about indigenous sovereignty and how to help with that. I think it's not necessarily organizing, but receiving information from organizers. (P4)

Adversaries

In understanding the threats that trans social media users face online, Lerner et al. identified common adversaries that were their source. Many of these were echoed by study participants, often less by being directly addressed as adversaries but mentioned in passing as individuals or groups who they needed to hide information from. One participant[1], who uses a preferred name that differs from their given one on Facebook and Instagram explained:

> *I've very selective about who follows me on both of those. Some family members just can't see that. (P9)*

Despite not appearing in the Lerner et al. study, employers appeared multiple times (2/5) among study participants as adversarial figures. This disparity between the present study and that of Lerner et al. can easily be explained by the fact that all 5 interviewees in this one were college students, a population likely to be especially concerned with being hired in the future. This sentiment demonstrates that being visibly trans online to almost anyone can be a threat to one's livelihood. Other adversaries pose even deeper and consequential threats. For example, multiple participants (2/5) brought up surveillance and action by the state when asked generally about their privacy and safety concerns on social media. One participant explained:

> *I'm less so worried about my physical safety in regards to other citizens, but more from police, state apparatus, stuff like that... When Trump was president, there was always kind of this underlying fear that, like I didn't think trans were going to get rounded up or something like that, but there was always this weird feeling for me that it kind of could happen or it could be weirdly criminalized somehow. I never really understood why, but I was always nervous about that. (P3, Nonbinary person)*

This sort of vague and unplaceable fear of harm and how it could come through technology was not remotely uncommon. In fact, almost all participants (4/5) brought up some sort of fear of the technological unknown, despite this subject never being directly asked about. For these participants, a great amount of worry about their safety and privacy on social media did not come from mechanisms of harm they were familiar with, but rather ones they confessed to not understanding or even being aware of. As P4 simply put it, "Honestly I'm wary of everything". The feeling that information about you is being extracted and used in ways you don't even know about is not unfamiliar to anyone who has been staying informed about the digital world over the past decade, but for trans social media users this phenomenon poses a much more existential threat.

Some threats experienced by participants were by contrast very concrete. Echoing the findings of Lerner et al.'s study, some participants (2/5) specifically brought up doxxing (the malicious release of personal and identifying information about someone online) as an active and commonplace threat throughout the online trans community.

Risk Models

In their study, Lerner et al. identified a few prominent "risk models" – frameworks of perceiving, understanding, and responding to threats of harm online – that guide their participants navigation of social media. Understanding these among participants of this study was especially important since whether or not the encryption tool increased their sense of safety and security is likely to be largely a function of how well it responds to their model(s).

While multiple of the models identified by Lerner et al. appeared within the 5 participants interviewed in this study, that of visibility stood out as by far the most pronounced, with every interviewee (5/5) indicating it at some point. In this model, being very visible or vocal as a trans person online is a primary driving force of risk, and by staying quiet or "laying low", as one participant said, you can avoid being harmed. As one participant put it:

> *I've never really been worried that someone is going to, like, hunt me down. But mostly because I haven't been super vocal about [trans topcis]. (P9)*

Despite not being recognized as a category by Lerner et al., another that appeared ubiquitously (5/5) was risk as a function of platform. In this model, participants understood a large source of risk to be the treatment of data by a particular social media platform, whether in terms of sale or exchange of that data with other corporations or government entities, or in terms of ease of access by the public. Another way to frame this model is as one of audience. *Who* is able to see information is a central concern, with the platform the information is on being a determining factor.

[1]This participant did not associate their interview data with their responses to the demographic survey, and as a result will be referred to as P9

Frequently, this model was introduced while discussing Facebook in particular, which was a common subject since participants were asked how the privacy/security experience in the study differed from a typical Facebook privacy/security experience. One participant explained:

> *I get freaked out often by how easily accessible stuff is particularly on places like Facebook. (P3, Non-binary person)*

Other risk models experienced by participants included ones based on particular racial or gender identities as risk factors (2/5), as well as technical proficiency and understanding (3/5).

These models of risk informed how participants adapted their actions to remain safe online. Every participant (5/5) talked about defenses they employ in their normal social media usage outside of the study, with some of the most common being restricting who follows them on various platforms (2/5) as well as the use of encryption-based security software (3/5).

Experiences with the Encryption Tool

Participants reported a variety of experiences with using the encryption tool, often with detailed understandings of what led them to have these.

Safe Space?

Almost all participants (4/5) reported that they overall felt that the study Facebook group was a safe space, with the final one describing it as a "safe-ish space". In some cases, this was clearly aided by the additional security provided by the study's security tool. When asked whether or not the group was a safe space, one participant responded:

> *Definitely. I think... because I knew that what we said couldn't get seen and processed by Facebook. (P9)*

However, at times it was also unclear how much of the sense of the group being a safe space, if any, could be directly attributed to the benefits of the security tool. Most participants (4/5) mentioned that the membership of the group was a major factor in it being a safe space. One of the study groups consisted of participants who knew each other beforehand and signed up together, and to multiple of them this was an important element of the safety they felt:

> *It was my friends... that's the big reason [the group was a safe space]. I don't think it hurt to have the encryption in place, but more so it was a safe place because it was people I knew pretty well (P3, Non-binary person)*

In the other group, where the participants did not know each other beforehand, common identity fulfilled a similar role in establishing trust and creating a sense of safety:

> *Having the commonality of all being part of the study and all being trans people, and especially all being Obies [Oberlin College students] I think is probably the biggest thing, having that common ground made me feel better. (P5, Non-binary person)*

Making a Difference

All participants (5/5) either said that they found the tool made a difference and was useful or that it would be in other situations or for other people. For those who did experience an effect of the tool themselves, this was often the improved emotional experience this study hoped to create:

> *I do think [I would have felt differently] without the security tool. I don't think that I would've felt as safe or open in the group if it didn't have that security in place. (P3, Non-binary person)*

Participants who said that the tool didn't make a difference for them personally during the study but might in other situations cited reasons that included the fact that the content they were sharing in the group wasn't content that they were especialy concerned about:

> *I think it would've been different if it was during Black Lives Matter, or rather the protests last summer, or if we were planning [a protest] against the school. (P4)*

Not Making a Difference

In some cases, by contrast, participants noted ways in which the encryption specifically did not make a difference. One participant related this not to a particular failing of the software to meet a need, but rather the technological paranoia discussed above:

> *I just feel that when there's that air that everyone can see everything, it doesn't go away when you use encryption, that air of suspicion. (P4)*

In contrast, though, in one case the security provided by the tool simply didn't answer the privacy and security concerns of the participant:

> *That there was a tool of any kind did make it feel a little safer, but I feel like the biggest barrier to using a Facebook group is the fact that it's on Facebook for me, and I feel like that's hard to get away from (P6, Gender questioning / feminine)*

Unlike P4, whose unmet concern was not concrete, P6 cited Facebook's practices of using cookies to track users across other websites and their ability to discern information about users based on trends of activity rather than content as the reason why the tool didn't make a bigger impact. This attitude reflects the risk model based on platform discussed above and how it presents challenges to tools like the one in this study.

A Lack of Understanding

Some participants (3/5) brought up their lack of understanding of how the tool worked being noticeably detrimental to their experience with it. While the participants interviewed overall had a high degree of familiarity with and faith in cryptographic technology (3/5 reported using it outside of the study), the fact that it was still a black box to them presented an obstacle to being able to trust it more. One participant, while still going to their study group's Facebook group to look at posts, almost never used it themself. Talking about why they didn't use it much, they explained:

> *I feel like I wasn't really super sure how it worked, I just knew that if you turned it on or had it then it'll make just make it more secure. That was my basic understanding... I had it on, but I didn't interact with it. (P6, Gender questioning / feminine)*

User Design, Practicality, and Technical Challenges

Frequent downsides to the encryption tool that appeared could be categorized as dislike of the user design, the impracticality of using the tool, and miscellaneous technical issues. While these are largely specific to the particular tool used in this study as well as its implementation, the fact that these types of complaints are detrimental remains significant and generalizeable to the greater category of privacy/security enhancing software.

One participant described how even just small extra steps involved in using the tool created enough of a nuisance that it disrupted the flow of the user experience:

> *Sometimes I would have to refresh the page in order to see the content that was posted, which is just more than I would usually have to do and I'm kind of lazy. (P5, Genderfluid person)*

These sorts of annoyances slowed down the use of Facebook enough for this participant that their interest in using the platform was actually dampened overall by the tool:

> *I think I might have posted more if [the group] hadn't had a security tool, just because I'm lazy and the extra step[s]... actually turned me from posting a bit. (P5, Genderfluid person)*

This sentiment highlights the need for security tools that are not just highly functional but also get close enough to matching the extremely well optimized use flow of a site like Facebook that they are actually enjoyable to use.

An especially common (4/5) frustration with the encryption tool was simply that participants didn't like using it with the medium of a Facebook group. Multiple of these participants noted how the type of communication that they felt the tool was useful for was better suited to a group chat of some type instead. One participant, who used the study Facebook group for planning social events, explained:

> *I personally feel like Facebook groups are unwieldy for scheduling social interaction or communicating easily. (P6, Gender questioning / feminine)*

This participant reported using a wide variety of social media platforms including Facebook in their daily life, but with a strong preference towards Twitter and Reddit.

Technical issues were also present among participants. One case demonstrated how these can extend beyond mild inconveniences and can actually undermine the emotional sense of security provided by software. This participant had multiple technical difficulties throughout the study, and reflecting back on them remarked:

> *I think I would have felt safer if we all knew how to use [the tool]. (P4)*

These findings reflect a key point often overlooked in studies that exclusively explore the security capabilities of software rather than user experiences – security software is only as valuable as how safe it makes users feel, and this can easily be determined by mundane hiccups in the user experience.

Discussion

In this section, we discuss the implications of the findings. Through doing so, we hope to build a more complete understanding of why crytographic technology both succeeds and fails to improve the social media experiences of trans people, and furthermore how this goal can be better accomplished in the future.

Risk Models

Reflecting the results of Lerner et al.'s study, participants demonstrated that they perceive and mitigate risk through models of understanding that can both enable them to make more informed decisions but also complicate the process of meeting their needs. In some cases, these models were understood as heuristics that simplified complex situations, such as when participants indicated that they could reduce the risk of harm simply by being less visible and vocal online instead of having to trace through the potential implications of every action they take. At other times, these models could be fully fleshed out understanding of how people can be harmed, such as that of P6, to whom any activity on Facebook was dangerous because of Facebook's use of cookies to track users across sites as well as discern personal information just from engagement trends.

In either case, the model(s) employed by an individual inform their perception of risk in a given situation online, and thus privacy and security tools are only useful if they not only protect users but also provide emotional relief through answering to these models. Returning to the two models just discussed and using the encryption tool in this study as an example, the tool is likely to be very useful to someone whose risk model is largely based on visibility since it hides the content they post online, but it may be nearly useless to someone whose risk model is primarily based on platform even though it works the same to both users and protects them exactly the same.

Privacy and Security Literacy

The presence of situations where security software successfully protects users but isn't able to provide a feeling of safety because of their risk models, where security software doesn't protect users but makes them think that it does because of their risk models, and where a lack of understanding of software just generally turns users away from it suggests the potential for better online privacy and security education to make significant improvements to the online experiences of trans people. As we saw, in some ways this is already happening – 3/5 participants brought up using encryption-based security software outside of the study, and an impressive array of knowledge about protecting yourself online was demonstrated throughout the interviews. Of course, there is potential for this to be biased data given that there were a small number of participants, and people who sign up for an online privacy and security study may be more likely to already have interest in the subject. Regardless, though, it was apparent that these discussions are already being had within trans communities online.

Since education in the trans community in this area could still be furthered, a potential topic for future research could be whether or not online privacy and security education (classes, seminars, informational documents, etc.) create a difference in trans social media users' sense of safety online without changes being imposed on their behavior. Given that sometimes participants of this study connected feeling unsafe with particular (accurate) understandings of online privacy and security risks, there may also be a substantial difference in the effect of education that focuses on techniques to remain safe and private versus education that focuses on risks.

Flexible Technology

One of the reasons the security tool design chosen for this study was selected was that it is especially flexible – the core mechanism could be applied to any social media site or webpage, or even text message chats. Results from this study further support the advantages of this approach.

For example, widespread sentiment among participants that platforms are mistrustful underscores the need for platform-independent solutions. Given the business model of mining and sometimes selling user data for advertising purposes that many social media platforms employ as well as the often shady or relatively hidden methods they go about this with, it is unsurprising that many participants expressed this sentiment. One participant even said they were mistrustful of Signal, an end-to-end encrypted messaging service. While one argument that could be made in response to this would be that there needs to be a demand for platforms to have better data practices and be more open and honest about them, the unfortunate reality is that it is simply more profitable to have weak privacy considerations, and U.S. government regulations have been slow to respond. All of this suggests a need for users with especially strong privacy and security needs, such as many trans users, to take meeting these needs into their own hands with tools that do not rely on the support of platforms and instead function independently.

Another reason the results of this study support the need for flexible, platform-independent tools is the rapidly changing social media ecology. Despite being a relatively new platform that only came to prominence over the past few years, 5/8 participants who filled out the initial demographic survey reported using Tik Tok. By contrast, only 3/8 reported using Tumblr despite it being a major hub of online trans communities prior to the changes to the site in 2017. To be able to consistently and stably meet the privacy and security needs of trans users, then, it is important for technologies to be flexible and adaptable to constantly changing sets of popular platforms.

Technology for Good and Evil

Technologies that are designed and developed for one community or purpose are often co-opted by others, such as when the Myanmar military utilized Facebook for the Rohingya genocide [21]. Unfortunately, these

misuses are often difficult or impossible to stop, presenting technology designers with the ethical dilemma of whether a given technology should exist at all if it has the potential for misuse.

With the rise of sentiment among the American far-right that major social media platforms like Facebook censor conservative speech, the adoption of security technologies such as the one in this study by extremist hate groups becomes a major concern. Already, platforms such as Parler are springing up and gaining prominence among those who views violate terms of service on other platforms for spreading hate, misinformation, or violence. Encryption tools like the one in the study could easily allow people to spread harmful content on major platforms while avoiding content policies put in place to protect people.

However, ceasing the development of these technologies for such reasons would prevent vulnerable users, such as trans people, to reap their benefits. Additionally, as one study participant noted while discussing the subject at the end of their interview, hate speech is already allowed to thrive online without these protections. Examples like the Rohingya genocide as well as the development of an alt-right hub on 4chan demonstrate that social media uses that directly do harm are already rampant out in the open. By contrast, this speech taking place over encrypted communications would close it off to the outside world, neutering its ability to indoctrinate new followers into hateful ideologies.

Of course, these reasons do not eliminate the need for concern over how security technologies can be used. It does, however, encourage further research for the benefit of those who are most vulnerable.

Limitations of the Study

A major limitation of the study was the lack of demographic diversity present among participants. All 8 were Oberlin College students of ages ranging from 19 to 23, and all interviewees identified as white (with one additionally identifying as hispanic/latinx). Given the differences in technological experiences across generations as well as the presence of race as a major risk factor for trans people online (as discussed by some participants), this lack of diversity could impose limitations on the generalizeability of the results to transgender people across the U.S.

Additionally, the fact that the study examines the experiences users had with only one particular security tool leaves open the possibility that they were largely determined by idiosyncrasies of the tool itself and are not representative of trans experiences of social media encryption tools more generally. Supporting the possibility that participants would have different experiences with a different tool are the variety of technical difficulties and user design challenges that participants discussed as impactful to their overall experience.

Conclusion

Trans social media users face an exceptional threat of harm online. To protect them while allowing social media to continue to provide benefits to them, it is important to explore technology solutions that both provide the actual privacy and security they need and ensure a sense of safety that allows online activity to be comfortable. To work towards this goal, we provided trans social media users for a few weeks with a security tool that encrypts their communications on Facebook and then interviewed some of them on their experiences with it. These interviews revealed complex and varied experiences with social media that gave participants unique and individual attitudes towards the security tool. Among those reported were both positive benefits as well as obstacles, which together painted a picture of how encryption technologies might succeed and fail to provide users with safe spaces.

In particular, the findings suggested potential for some trans social media users to find benefits from taking their privacy and security into their own hands with similar security software, with one of the greatest obstacles to this being a lack of privacy and security knowledge causing confusion and diminishing the effectiveness of tools. For all participants, the tool failed to be perfect in at least one regard and was effective in another, demonstrating the highly individual and nuanced experiences of trans social media users that reject the possibility of any one-size-fits-all solution.

Acknowledgements

The author would like to thank the participants who were so kind to share their deeply personal experiences, as well as Professor Roberto Hoyle for providing guidance throughout the project.

References

[1] Ahmed Abdulla and Spiridon Bakiras. *HITC: Data Privacy in Online Social Networks with Fine-Grained Access Control*. Pages: 134. May 2019. DOI: 10.1145/3322431.3325104.

[2] Laima Augustaitis. "Online Transgender Health Information Seeking: Facilitators, Barriers, and Future Directions". en. In: (2021), p. 14.

[3] Randy Baden et al. "Persona: an online social network with user-defined privacy". en. In: (), p. 12.

[4] Filipe Beato, Markulf Kohlweiss, and Karel Wouters. "Scramble! Your Social Network Data". en. In: *Privacy Enhancing Technologies*. Ed. by David Hutchison et al. Vol. 6794. Series Title: Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 211–225. ISBN: 978-3-642-22262-7 978-3-642-22263-4. DOI: 10.1007/978-3-642-22263-4_12. URL: http://link.springer.com/10.1007/978-3-642-22263-4_12 (visited on 10/18/2020).

[5] Filipe Beato et al. "For some eyes only: protecting online information sharing". en. In: *Proceedings of the third ACM conference on Data and application security and privacy - CODASPY '13*. San Antonio, Texas, USA: ACM Press, 2013, p. 1. ISBN: 978-1-4503-1890-7. DOI: 10.1145/2435349.2435351. URL: http://dl.acm.org/citation.cfm?doid=2435349.2435351 (visited on 10/18/2020).

[6] Rena Bivens and Oliver L. Haimson. "Baking Gender Into Social Media Design: How Platforms Shape Categories for Users and Advertisers". In: *Social Media + Society* 2.4 (Oct. 2016). Publisher: SAGE Publications Ltd, p. 2056305116672486. ISSN: 2056-3051. DOI: 10.1177/2056305116672486. URL: https://doi.org/10.1177/2056305116672486 (visited on 09/09/2020).

[7] Alexander Cho. "Default publicness: Queer youth of color, social media, and being outed by the machine". en. In: *New Media & Society* 20.9 (Sept. 2018), pp. 3183–3200. ISSN: 1461-4448, 1461-7315. DOI: 10.1177/1461444817744784. URL: http://journals.sagepub.com/doi/10.1177/1461444817744784 (visited on 09/11/2020).

[8] Emiliano De Cristofaro et al. "Hummingbird: Privacy at the Time of Twitter". In: *2012 IEEE Symposium on Security and Privacy*. ISSN: 2375-1207. May 2012, pp. 285–299. DOI: 10.1109/SP.2012.26.

[9] Lujun Fang and Kristen LeFevre. "Privacy wizards for social networking sites". en. In: *Proceedings of the 19th international conference on World wide web - WWW '10*. Raleigh, North Carolina, USA: ACM Press, 2010, p. 351. ISBN: 978-1-60558-799-8. DOI: 10.1145/1772690.1772727. URL: http://portal.acm.org/citation.cfm?doid=1772690.1772727 (visited on 09/14/2020).

[10] Ariel J Feldman et al. "Social Networking with Frientegrity: Privacy and Integrity with an Untrusted Provider". en. In: (), p. 16.

[11] Saikat Guha, Kevin Tang, and Paul Francis. "NOYB: privacy in online social networks". en. In: *Proceedings of the first workshop on Online social networks - WOSP '08*. Seattle, WA, USA: ACM Press, 2008, p. 49. ISBN: 978-1-60558-182-8. DOI: 10.1145/1397735.1397747. URL: http://portal.acm.org/citation.cfm?doid=1397735.1397747 (visited on 10/17/2020).

[12] Oliver L Haimson et al. "Trans Time: Safety, Privacy, and Content Warnings on a Transgender-Specific Social Media Site". en. In: 4 (2020), p. 27.

[13] Oliver L. Haimson et al. "Digital Footprints and Changing Networks During Online Identity Transitions". en. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. San Jose California USA: ACM, May 2016, pp. 2895–2907. ISBN: 978-1-4503-3362-7. DOI: 10.1145/2858036.2858136. URL: https://dl.acm.org/doi/10.1145/2858036.2858136 (visited on 09/13/2020).

[14] Oliver L. Haimson et al. "Disclosure, Stress, and Support During Gender Transition on Facebook". en. In: *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing - CSCW '15*. Vancouver, BC, Canada: ACM Press, 2015, pp. 1176–1190. ISBN: 978-1-4503-2922-4. DOI: 10.1145/2675133.2675152. URL: http://dl.acm.org/citation.cfm?doid=2675133.2675152 (visited on 09/13/2020).

[15] Oliver L. Haimson et al. "Tumblr was a trans technology: the meaning, importance, history, and future of trans technologies". en. In: *Feminist Media Studies* (Oct. 2019), pp. 1–17. ISSN: 1468-0777, 1471-5902. DOI: 10.1080/14680777.2019.1678505. URL: https://www.tandfonline.com/doi/full/10.1080/14680777.2019.1678505 (visited on 09/20/2020).

[16] Benjamin Hanckel et al. "'That's not necessarily for them': LGBTIQ+ young people, social media platform affordances and identity curation". en. In: *Media, Culture & Society* 41.8 (Nov. 2019), pp. 1261–1278. ISSN: 0163-4437, 1460-3675. DOI: 10.1177/0163443719846612.

URL: http://journals.sagepub.com/doi/10.1177/0163443719846612 (visited on 09/20/2020).

[17] Sonia Jahid, Prateek Mittal, and Nikita Borisov. "EASiER: encryption-based access control in social networks with efficient revocation". en. In: *Proceedings of the 6th ACM Symposium on Information, Computer and Communications Security - ASIACCS '11*. Hong Kong, China: ACM Press, 2011, p. 411. ISBN: 978-1-4503-0564-8. DOI: 10.1145/1966913.1966970. URL: http://portal.acm.org/citation.cfm?doid=1966913.1966970 (visited on 10/11/2020).

[18] Sonia Jahid et al. "DECENT: A Decentralized Architecture for Enforcing Privacy in Online Social Networks". In: *arXiv:1111.5377 [cs]* (Dec. 2011). arXiv: 1111.5377. URL: http://arxiv.org/abs/1111.5377 (visited on 10/16/2020).

[19] Ada Lerner et al. "Privacy and Activism in the Transgender Community". en. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Honolulu HI USA: ACM, Apr. 2020, pp. 1–13. ISBN: 978-1-4503-6708-0. DOI: 10.1145/3313831.3376339. URL: https://dl.acm.org/doi/10.1145/3313831.3376339 (visited on 09/21/2020).

[20] Wanying Luo, Qi Xie, and Urs Hengartner. "FaceCloak: An Architecture for User Privacy on Social Networking Sites". en. In: *2009 International Conference on Computational Science and Engineering*. Vancouver, BC, Canada: IEEE, 2009, pp. 26–33. ISBN: 978-1-4244-5334-4. DOI: 10.1109/CSE.2009.387. URL: http://ieeexplore.ieee.org/document/5283227/ (visited on 10/18/2020).

[21] Paul Mozur. "A Genocide Incited on Facebook, With Posts From Myanmar's Military". en-US. In: *The New York Times* (Oct. 2018). ISSN: 0362-4331. URL: https://www.nytimes.com/2018/10/15/technology/myanmar-facebook-genocide.html (visited on 04/18/2021).

[22] Shirin Nilizadeh et al. "Cachet: a decentralized architecture for privacy preserving social networking with caching". en. In: *Proceedings of the 8th international conference on Emerging networking experiments and technologies - CoNEXT '12*. Nice, France: ACM Press, 2012, p. 337. ISBN: 978-1-4503-1775-7. DOI: 10.1145/2413176.2413215. URL: http://dl.acm.org/citation.cfm?doid=2413176.2413215 (visited on 10/16/2020).

[23] Morgan Klaus Scheuerman, Stacy M. Branham, and Foad Hamidi. "Safe Spaces and Safe Places: Unpacking Technology-Mediated Experiences of Safety and Harm with Transgender People". en. In: *Proceedings of the ACM on Human-Computer Interaction* 2.CSCW (Nov. 2018), pp. 1–27. ISSN: 2573-0142, 2573-0142. DOI: 10.1145/3274424. URL: https://dl.acm.org/doi/10.1145/3274424 (visited on 09/09/2020).

[24] Zeynep Tufekci. "Can You See Me Now? Audience and Disclosure Regulation in Online Social Network Sites". In: *Bulletin of Science, Technology & Society* (), p. 36.

[25] Yang Wang et al. "A field trial of privacy nudges for facebook". en. In: *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*. Toronto, Ontario, Canada: ACM Press, 2014, pp. 2367–2376. ISBN: 978-1-4503-2473-1. DOI: 10.1145/2556288.2557413. URL: http://dl.acm.org/citation.cfm?doid=2556288.2557413 (visited on 09/09/2020).

[26] Xuan Zhao, Cliff Lampe, and Nicole B. Ellison. "The Social Media Ecology: User Perceptions, Strategies and Challenges". en. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. San Jose California USA: ACM, May 2016, pp. 89–100. ISBN: 978-1-4503-3362-7. DOI: 10.1145/2858036.2858333. URL: https://dl.acm.org/doi/10.1145/2858036.2858333 (visited on 09/14/2020).