"I feel physically safe but not politically safe": Understanding the Digital Threats and Safety Practices of OnlyFans Creators

Ananta Soneji
Arizona State University

Vaughn Hamilton
Max Planck Institute for Software Systems

Adam Doupé
Arizona State University

Allison McDonald
Boston University

Elissa M. Redmiles
Georgetown University

Abstract

OnlyFans is a subscription-based social media platform with over 1.5 million content creators and 150 million users worldwide. OnlyFans creators primarily produce intimate content for sale on the platform. As such, they are distinctly positioned as content creators and sex workers. Through a qualitative interview study with OnlyFans creators (n=43), building on an existing framework of online hate and harassment, we shed light on the nuanced threats they face and their safety practices. Additionally, we examine the impact of factors such as stigma, prominence, and platform policies on shaping the threat landscape for OnlyFans creators and detail the preemptive practices they undertake to protect themselves. Leveraging these results, we synthesize opportunities to address the challenges of sexual content creators.

1 Introduction

The exponential growth of OnlyFans since its founding in 2016 has transformed it into a prominent social media platform with over 1.5 million creators, who primarily create erotic content, and 150 million "fans" [71]. OnlyFans creators, like other professional content creators, navigate a complicated landscape of content creation, monetization strategies, and fan management, and need to engage in extensive off-platform advertisement to draw followers. The prominence that comes with being a content creator, regardless of the platform, brings specific security and privacy risks such as obsessive fans, online harassment, and stalking [30, 91].

OnlyFans creators face additional sources of risk stemming from sex work stigma [30]. Those who create and sell sexual content are sex workers; they exchange erotic services for money [69]. However making, distributing, or even possessing certain kinds of sexual content is legislated in a variety of ways across different countries [84]. Although millions of people worldwide work in the sex industry, sex workers globally face significant stigma, discrimination, violence, and harassment [23, 67, 82]. Therefore, OnlyFans' sexual content

creators must manage risks such as censorship of sexual expression, platform precarity [30], refusal of service from banking institutions [77], and digital account loss [15]. Despite these risks, OnlyFans provides a level of flexibility, accessibility, and control that draws creators to the platform [47].

Drawing on the contextual risk factors outlined by Warford et al. [99], OnlyFans creators are at-risk—"anyone who experiences heightened digital-safety threats"—due to their *prominence* as well as the *stigma* associated with sex work. With this work, we respond to the call of Thomas et al. [90]—"responding to each class of hate and harassment requires a unique strategy"—by identifying the online risks to and protective measures taken by sexual content creators.

Through semi-structured interviews with 43 OnlyFans creators,[1] we investigate their online risks and the protective strategies they employ to mitigate them. In particular, we recruited both participants who were new to creating sexual content and those with previous sex industry experience, capturing a broad spectrum of experience and stages of learning about digital safety. We investigate three research questions:

RQ1: How well does the existing online hate and harassment framework [90] capture the threats experienced or perceived by sexual content creators?

RQ2: How do contextual risk factors [99]—stigma and prominence—relate to the threats faced by sexual content creators?

RQ3: What is the role of platforms in sexual content creators' digital safety?

We analyze our interviews with OnlyFans creators through the lens of Thomas et al.'s online hate and harassment framework [90]. We do so for two primary reasons, to: (i) support a growing effort to use unifying language and frameworks [99] across research of hate, harassment, and safety issues and (ii) validate the framework in the context of the combination of contextual risk factors we study—prominence and stigma. This approach further allows us to juxtapose our participants'

---

[1]Hereafter, when we use the term 'OnlyFans creators,' we refer to creators producing sexual content on the platform.

experiences with those of previously studied populations: general internet users [100], at-risk users [99] and professional non-sexual content creators [81, 91].

We find that OnlyFans creators contend with two existing threats from Thomas et al.'s online hate and harassment framework—toxic content and content leakage—and two sources of risk not included in the Thomas et al. framework: platform precarity and censorship (e.g., shadowbanning[2] and deplatforming[3]). Whereas prior work finds that non-sexual content creators find toxic content to be the most prominent threat [91], our participants were more concerned about platform-related risks than they were about toxic content and content leakage. In response, we find participants diversified the platforms they used and practiced self-censorship to avoid being deplatformed.

The OnlyFans creators we study face many of the same toxic content attacks as non-sexual content creators [91]. While they use some of the same strategies to address such toxic content—"ignoring haters" and engaging in self-censorship [91]—our participants also establish strong boundaries with their fans, extensively use platforms' blocking features to counter harassment, and draw strength and support from the broader community of sexual content creators.

Given the intimate nature of content produced as part of sexual content creation, our analysis emphasizes the prevalent and seemingly unavoidable threat of intimate content leakage—constituting both content theft and non-consensual distribution of intimate content where the creator did not intend to share it [68]—a consequence our participants often associated with the online sharing of explicit content.

While prior work on non-sexual content creators finds a reactive approach to digital safety [81], we find that sexual content creators are uniquely proactive and prepared to defend against digital threats. *All* of our participants adopted at least one protective practice before joining OnlyFans, compared to only one-third of participants in prior work focused on non-sexual content creators. This preparedness is vital due to the essential requirement for cross-platform promotion to earn money on OnlyFans. Without internal promotion mechanisms on OnlyFans, creators must drive traffic to their page from other platforms. This increases creator visibility and subjects them to heightened risks such as outing and doxxing[4] [81]. Participants point to the role of sex work stigma in influencing their proactive stance toward digital safety.

---

[2]Shadowbanning refers to platforms silently suppressing a user and/or their content, reducing their visibility. Users may be unaware they have been shadowbanned until they experience unusually low engagement.

[3]Deplatforming refers to the removal of a creator's account from a platform, often without transparency or recourse [16].

[4]Doxxing refers to revealing someone's real name, home address, workplace, phone number, or other personal information for harassment and financial harm.

2   OnlyFans Background

We begin with an overview of the platform, creator interactions with the platform and fans, and noteworthy platform policy changes impacting creator experiences.

OnlyFans and creators. OnlyFans is a fast growing (70% per month [38]) social media platform with two types of user: creator and fan. All users must confirm that they are over 18, with additional identity verification checks for creators[5]—name, date of birth, address, government photo ID, selfie, social media handles, payment details—and consent checks when content features more than one creator.

Creators can host both free and paywalled pages. Users must have an OnlyFans account to view free pages. Creators can also generate income through pay-per-view content, custom content requested by individual fans, paid direct messages—where creators can set a per-message fee or flat fee to set up messages for their fans [14]—and fan tips. Creators earn 80% of all payments. OnlyFans creators manage their payouts and are responsible for chargebacks from fans, a model termed "digital patronage" [18]. OnlyFans limits the type of content creators can share and sets guidelines for language in private messages [72].

Notably, the platform lacks internal discoverability, necessitating creators to drive traffic internally (when creators promote each other) or externally (from social media platforms or personal websites) [6], fostering "multi-platform practices" [44]. This boosts the platform's public visibility, and OnlyFans' brand recognition has further grown by "celebrity hype" [47] and extensive mainstream news media coverage [78], particularly during the COVID-19 pandemic. Additionally, Hamilton et al. argue that OnlyFans' unique design has not only spurred its growth but has also attracted a diverse range of sexual content creators, including those with existing sex industry experience and newcomers [47].

Policy changes and known creator risks. Due to its association with erotic content [42], OnlyFans creators are exposed to the prevailing stigma attached to sex work [53] and the challenges stemming from platform policies. Notably, Bhalerao et al. point out that platform policies, created without subject matter knowledge or sufficient understanding of its effects, can inadvertently harm sex industry workers [13], as demonstrated by a proposed sexual content ban by OnlyFans in August 2021. The platform attributed the ban to banking-related issues [42] and despite being reversed, resulted in creators suffering financial and subscriber losses [8,13]. The utilization of online platforms by sex workers for both work and daily activities is significantly affected by policies that are stigmatizing and overly restrictive [13]. These policies may lead to deplatforming and reduced platform accessibility, highlighting the precarious nature of sex work platforms [5, 13, 17]. Such circumstances can have profound financial and emotional im-

---

[5]https://onlyfans.com/transparency-center/verification

pacts on marginalized content creators [4]. Furthermore, the structural design of OnlyFans necessitates off-platform advertising, subjecting creators to the policy changes and associated risks of other platforms.

Additionally, since OnlyFans uses third-party services for user verification and monitoring for minor content on the platform, breaches of OnlyFans or its third-party services are sources of risk for creators [85]. In a case study, Ford highlights the absence of preventive measures against screenshots or screen recording on OnlyFans, creating a risk of unauthorized distribution of creators' material, constituting both service theft and image-based sexual abuse [34].

3 Related Work

We discuss prior literature on online risks and safety practices of content creators and sex workers.

Content creators. Across social media platforms, content creators produce content for financial benefits [56, 93], self-expression [19, 65, 70, 73], socialization [65, 73, 93], and popularity (e.g., to create "media brands") [26]. However, prior work explored the multifaceted challenges that content creators face. For instance, Uttarapong et al. investigated the harassment experiences of women and LGBTQ streamers on Twitch, revealing difficulties in coping real-time with negativity from fans [96]. While Jhaver et al. describe how YouTube creators benefit from using collaborative keyword filtering to manage hate and harassment [52], community self-moderation requires significant emotional and relational work [96]. Further, such labor may not fall equally. Haimson et al. found that content moderation may lead to disproportionate censorship of trans people and certain racial identities (e.g., Black), despite following platform policies [43].

In a survey of 135 social media creators, Thomas et al. explored online risks, including targeted attacks, hate speech, threats of violence, impersonation, account hijacking, stalking, and false reporting [91]. To cope, creators often resort to self-censorship or may choose to leave a platform entirely. Samermit et al. in a recent interview study with 23 content creators, found that prominence, social norms, marginalization, and access to sensitive resources were reasons for online risks, including toxic content, harms to family, stalking, privacy risks, impersonation, and account hijacking or false reporting [81]. To cope, creators maintained privacy, employed moderation techniques, and sought advice from other creators, but found platform support and legal aid hard to reach.

Sex workers. Sex work encompasses a diverse spectrum of sexual labor, including stripping, pornography, and outdoor full-contact sex work [60]. Digitally-mediated sex work, defined as "the Internet-mediated exchange of sexual commodities and/or services" [53, 82], is performed by sex workers who organize and market their labor through digital platforms [12, 53, 75]. This form of sex work may offer increased

safety, a more appealing labor environment, better wages, efficient advertising, and client vetting methods [27, 53].

Digitally-mediated sex work includes both digital-mediation of in-person work (e.g., digital communication with clients for in-person sex work) and online-only sex work like OnlyFans or web-camming "in which cam models sell interactive computer-mediated sex online" [54]. Jones and Campbell, respectively, discuss challenges in digitally-mediated in-person sex work, namely lack of control on digital identities [53], under- or non-payment for services, online harassment including persistent unwanted contact, verbal abuse, surveillance, and hesitance in seeking law enforcement help against harms given the non-existent legalization of crimes on digital platforms [23]. Sanders et al. discover that U.K.-based sex workers mitigate such digital risks by using aliases, filtering clients, and utilizing social media to connect with friends and ensure their safety [82]. In a recent study, McDonald et al. highlighted that in-person sex workers faced risks coming from clients, deficient legal protections, and hostile digital platforms, which participants mitigate by covering for a friend or colleague, vetting clients, and managing digital identities [67]. Strohmayer et al. studied how sex workers use digital technologies to report harassment such as boundary pushing and violence, and use such tools to proactively avoid violent or disrespectful clients [87, 88].

In analyzing the largest UK-based adult entertainment platform, Rand [75] finds that although digital platforms enable flexibility and choice, sex workers become dependent on those platforms. As a consequence, workers may pay higher intermediary fees and must constantly diversify and engage with consumers to stay profitable in a competitive market.

The traffic to online adult websites grew [39] during the COVID-19 pandemic outbreak and associated lockdowns, providing a source of income for both sex workers and those without prior sex work experience [45, 47, 78]. Hamilton et al. found that the in-person sex workers who shifted to online-only sex work during the COVID-19 pandemic experienced increased digital exposure and faced platform loss (deplatforming), stalking, harassment, and content theft [45]. Rubattu et al. performed a media analysis post COVID-19 pandemic, finding that those new to sex work experienced outing and harassment online [78], echoing pre-pandemic observations [53, 54, 82].

Apart from the aforementioned challenges, sex workers also face risks stemming from different legislatures and social norms across countries. Sex work legality varies from country to country and even within individual countries. Prior work finds that American legislation and social attitudes lead platforms (including payment processors) to censor sexual expression and sex work to shield from liability and protect their corporate image—regardless of the legal status of the censored expression or labor in the U.S. or in the user's country [1, 9]. To do so, platforms may use algorithmic shadow-banning[2] [15, 92], place blanket bans in their terms of service

(ToS) against the use of their services for any sex work-related activities [9, 13, 15, 16], or over-censor potentially objectionable content [4]. Stardust et al. theorize that sex work stigma is a key driver for such policies and practices and argue for anti-discrimination protections for sex workers [86].

Beyond censoring expression and work, prior work also finds that sex workers themselves may be systematically identified and de-platformed based purely on their identity as sex workers, regardless of whether they are using the platform for that work [9, 16, 67]. Such censorship can lead to diminished financial stability, safety, and community access [7, 17].

OnlyFans, a UK-based platform, operates where many types of sex work are legal and primarily facilitates pornographic services, which are legal in the US. Our work contributes to the broader literature on both professional content creation and online sex work as we investigate creator security and privacy on OnlyFans, a web-based social media platform focused primarily on sexual content and with a unique combination of attributes including lack of internal searchability that requires cross-platform advertising and interaction, a greater level of mainstream visibility than other sexual-content focused platforms [45, 47], and a variety of subscription, paywall and direct interaction models of creator-fan interaction [79].

4 Methods

We conducted 43 semi-structured interviews with OnlyFans creators in two cohorts: 1) at the onset of the COVID-19 pandemic and 2) between September and October 2021. Additional questions regarding the impacts of policy changes were incorporated into our questionnaire in response to OnlyFans' announcement and subsequent retraction of the ban on explicit content in August 2021 (§ 2). Here, we provide an overview of our participant recruitment, data collection, and data analysis approach, along with the ethical practices followed throughout this multi-phase interview study.

4.1 Participant Recruitment

We recruited participants with a variety of backgrounds including those with and without prior experience in sex work on social media, through the researchers' personal networks, and for the second phase of recruitment, with printed flyers in coffee shops, grocery stores and on university campuses in multiple U.S. cities.

The link and QR code on the recruitment graphic led participants to a screening survey in Qualtrics (Appendix A). We employed Qualtrics' quota function with filters for age, gender, and ethnicity to achieve a diverse sample, encompassing various identities that could impact participants' experiences on the platform. Given that OnlyFans does not publish demographic data about its creators and the inherent limitation of smaller sample sizes in interview studies, our aim was to prioritize a diverse sample rather than pursuing a "representative"

one. We therefore used the quota function to ensure demographic diversity amongst those we recruited. Qualification for the study was limited to those over 18 and currently or recently (within 6 months) creating content on OnlyFans. Participants that qualified through the screening survey questions were invited to self-schedule an interview through Calendly, which allows for anonymous sign-up.

Participant demographics. We had a diverse sample of participants across many identities with a median age of 30 years ($\sigma = 5.37$); self-reported genders[6] of woman (25), non-binary (15), man (4), and trans (3); and self-reported race/ethnicity of: white alone (26), Black (5), Asian (4), Hispanic and white (3), Asian and white (1), and other (4). 26 of 43 participants were disabled (3 choose not to answer). Of the first cohort of participants, all were LGBTQ+ and we did not ask the second cohort to specify. We did not ask the first cohort about their highest level of education or their income on OnlyFans. Most of the second cohort had at least some college education but no degree, with the rest having additional qualifications. Among those who disclosed their OnlyFans income earned to date, the reported income ranged between $135 and $332,000 ($\sigma = \$83,974.24$). Out of 43 participants, 24 had prior sex work experience, 19 did not, and three had recently left the platform before the interview.

Not all of our participants disclosed their number of followers; but those who did disclose ranged from creators with tens to those with tens of thousands of fans. The web-use skills [48] of our 43 participants ranged from 1.33 to 5 with a mean of 3.57 ($\sigma = 1.21$).

4.2 Data Collection

Before starting the interview, interviewers explained the study goals and answered any questions regarding confidentiality, the research team and data use, and reminded participants that they could skip any question. We then re-affirmed consent to record. We used a semi-structured interview protocol, starting with background questions about their prior works and their comparisons to OnlyFans experience, content creation, business model, safety and privacy, impacts of changes in ToS, platform use, and overall OnlyFans experience. The relevant interview questions for this paper can be found in Appendix A. All interviews were conducted online in English via video, voice, or chat. They were transcribed and any identifying information was removed. Participants were compensated $50 through PayPal or Amazon gift card.

4.3 Data Analysis

Study researchers employed thematic analysis [20] to gain insights into the online experiences of OnlyFans creators. After becoming familiar with the collected data, the primary

---

[6]We report genders following the HCI Guidelines for Gender Equity and Inclusivity: www.morgan-klaus.com/gender-guidelines.html

coder inductively created an initial codebook [25] from a random selection of five transcripts. This initial codebook was then cross-referenced with the online hate and harassment framework [90] to add any missing themes and align the terminology of the inductive codes with the framework where applicable. Two coders used this codebook to independently analyze five new transcripts and test the coverage of codes. They met to discuss code applications extensively and made necessary refinements to the codebook, mainly developing the codes into themes. After independently coding three more transcripts, the coders achieved a substantial inter-coder reliability score for Cohen's Kappa ($\kappa > 0.7$) [58] and refined the codebook until any minor disagreements were resolved. Coders used the final codebook to independently code the remaining transcripts, evenly dividing them between both.

We primarily report our findings qualitatively using participant quotes, shared experiences, and incidents to describe and interpret the themes that emerged from the interviews. While reporting, we occasionally provide counts (i.e., how many participants reported each risk or safety practice) of overall themes to demonstrate the prevalence of certain patterns and ensure transparency in our data analysis.

4.4 Ethics

The research procedures were approved by the institutional ethics boards of the researchers. At every stage of the research, we undertook great care to provide the highest possible level of protections for participants' data privacy and anonymity. We used end-to-end encrypted platforms for the interviews (paid Webex) and no systems required legal information or login. The interviews were scheduled using Calendly and we provided advice on signing up with an encrypted email address if preferred. The payment options could both be anonymous (Amazon gift card) or to any address (PayPal).

Research justice. We employed a sex worker to transcribe the interviews. A copy of the published research will be sent to all participants who requested it. We also employed OnlyFans creators to work as consultants on designing the study (including the interview protocol) and they circulated the recruitment graphic within their networks. The authors' commitment to research justice contributes to their trust capital as researchers of a population with research fatigue. This trust capital is essential in such research with a marginalized community.

Positionality. To design the study, recruit participants, and write the interview protocol, the research team included OnlyFans creators as consultants. The data analysis and writing of this paper were carried out by researchers who are scholars of technology and sex work but who are not OnlyFans creators. The OnlyFans creator consultants who assisted in the study design were not participants.

Limitations. OnlyFans is an extremely competitive market. Whilst we felt our participants were very honest with us, it is

important to note that due to the nature of the work and the industry in general, some participants might have preferred to omit relevant information. OnlyFans has a global reach, and has creators and fans from many countries. We conducted interviews with creators only from the Global North in English, which limits generalizations regarding other territories. By focusing on our participants' shared (positive or negative) experiences, some insights might have been missed from those with contrasting opinions, however our participants' varied experiences on the platform might have balanced out the bias. Lastly, the participant demographics include creators with genders or ethnicities which were not reflected in the interview or research team, limiting our insight into that data.

5 Results

After analyzing interviews with 43 OnlyFans creators through the lens of the online hate and harassment framework [90], we now delve into their perceived and experienced risks and how they defend against those risks.

5.1 Platform Risks

Participants were most concerned (29 participants) about risks from platform precarity and censorship.

Platform precarity. The policy of banning sexual content from OnlyFans, which was later reversed (§ 2), had a lasting impact on our participants. The impacts of this ToS change were not only financial: "people were not re-billing or re-subscribing" (P11), but also emotional:

> *When you deplatform creators you take away their community and that's violence. You are isolating people and as we all know from this last year isolation leads to severe mental health consequences.*

P9 reported that the change in ToS and resulting media coverage elevated the platform's visibility, leading to increased harassment toward creators: "we didn't even get to find out from the platform itself, we found out from news outlets and I think it also exposed us to a lot of harassment on the internet from people who don't like sex work."

OnlyFans explained the ToS change as being a result of pressure from financial institutions [77]. Some participants shared this view, explaining that "banking institutions are now deciding what is moral and what is not moral to purchase... it directly impacts people on OnlyFans" (P11). Others ascribed intentionality and responsibility to OnlyFans. 14 participants felt that OnlyFans "is uncomfortable with being a platform that hosts sex work" (P12), "is trying really hard to say no, there are no sex workers... it's mostly chefs and fitness influencers" (P39), and was "cleaning up their image" by banning sexual content (P26). P22 felt that "what happened with the ToS revealed that OnlyFans simply does not believe sex workers to be their primary creative force, which I think

they are." After the ban, 14 participants—new to creating sexual content—expressed feeling less safe due to the increased platform precarity. P6 stated:

> *In terms of like the decisions that the platform's making... they just use sex workers and then will sell them out. I feel physically safe you know, but like not as politically safe as I did at first if that makes sense.*

Platform censorship. Since OnlyFans lacks internal discoverability, creators work cross-platform to drive fans to their accounts. These platforms may remove, block or restrict the content they share, shadowban their accounts, or deplatform creators, even if the platform's ToS are not violated.

Twitter does not ban pornography and other forms of consensually produced adult content when marked "sensitive" [35]. But P32 experienced shadowbanning on Twitter for "overly promoting OnlyFans." Although platforms by and large hide the criteria for shadowbanning content [4], our participants have identified behaviors that seem to impact their reach. They perceive that the likelihood of shadowbanning while promoting on other platforms increases when creators using automation tools are perceived as bots (P27), when they use hashtags often (P25, P31, P33, P38), or send the same message with an OnlyFans link too many times in a period of time (P23). Creators become aware of their content being shadowbanned from platforms only when they observe less engagement or income generation. P23 described how fear of shadowbanning creates cascading promotional challenges because "other colleagues don't tend to retweet you when you're promoting your OnlyFans, probably out of fear of getting shadowbanned themselves."

Prior work on professional non-sexual content creators [81] notes that creators are highly aware that violation of platform norms comes at a high cost—losing access to the platform and its associated audience. Our participants similarly feared violating OnlyFans' and other platforms' ToS. 13 of the 22 participants in our second cohort reported that they had either read or skimmed through the ToS. P9 explained their proactivity in carefully reading the ToS: "they [OnlyFans] often ban people and they [the creators] don't get their pending balance. It's not a risk that I'm willing to take."

Even when a post does not violate ToS, censorship occurs due to factors such as co-moderation (where other users participate in moderation by blocking and reporting other people) [15], discrimination against expression of sexuality, sexual identity [15], poor implementation of not-safe-for-work (NSFW) moderation [92], or when creators post content that is legally sensitive—deemed criminal in some countries [4] or perceived as advocating against policies and laws [15].

In line with prior work on racial, gender and body size variance in platform censorship of sexual expression [36, 43, 66], participants underscored the variance in censorship across creators. P39 explained: "I was so heavily censored... I had a photo of me and my friend that was clothed, and we were kissing, and we were two Asian women, I'm South Asian, she's Asian, and that got deleted right away... similar photos of my white counterparts weren't censored... more explicit photos." Conversely, P30 described his experience:

> *My risk of being censored is a lot lower as a male performer, and as a white performer. So I'm quite amazed my Instagram is still available and intact while other people have lost their account a thousand times in that same time period.*

Participants tried to hypothesize about the cause of such censorship, pointing to co-moderation (P42: "I'm trying to figure out if it's someone reporting me or if it's Instagram itself" and P9: "TikTok is huge for that [censorship]... since they have a user-based report system") and suppression of sexual expression (P27: "I feel like the algorithm interprets that [pictures in panties] as little less human").

Fear of violating platform norms can lead to creators proactively silencing themselves or leaving content creation [91]. P39 explained how the disproportionate impacts of platform censorship can end up silencing some more than others:

> *It's frustrating to be constantly living censored both on Twitter and Instagram. It's a systematic way of silencing people of color and queer folks so we just give up and stop posting which essentially is what happens. It's effective, because it's exactly what happened. I gave up.*

To make platform moderation more equitable for marginalized social media users, Haimson et al. suggest platforms collaborate with marginalized communities and develop censorship policies and enforcement mechanisms that better reflect the voices and needs of the community [43].

5.1.1 Defenses

To defend against platform precarity and censorship, participants diversified the platforms they used to sell content and/or engaged in self-censorship on the platforms they already used.

Platform diversification. 12 participants diversified the platforms on which they created and sold content. While for some, like P10, diversification was purely to defend against precarity ("the reason I'm building out on another platform is just in case something like that [sexual content ban] happens again, I definitely want to be diversified even more"), others diversified to avoid violating ToS. P12 publishes "prohibited content to another [sex work] platform [that allows it] instead." Participants chose platforms for diversification based on their longevity (P18) or recommendations from other creators (P3, P4, P15). While some participants actively used the accounts they created on other platforms, others username squatted—created accounts with their preferred username to reserve it in case they were deplatformed or OnlyFans ceased to exist in the future—but did not actively use their other accounts.

Self-censorship and Algospeak. Participants self-censored on social media platforms to reduce platform harms ($n = 24$,

§ 5.1) and manage their identity ($n = 22$, § 5.3.2).

P29 created a cheat-sheet to help them self-censor on Only-Fans: "when I started I wrote everything on paper that we can't use, so I'm always looking at that sheet for like okay, this I can, this I can['t]." On advertising platforms (e.g., social media), participants also invested significant effort to ensure that they comply with all non-adult-content related regulations and ToS (e.g., P33 avoids posting with music that "could potentially be flagged for copyright").

Algorithm Speak (Algospeak) encompasses strategies of abbreviations, deliberate misspelling, or substitution of words to avoid automated content moderation systems [2]. This approach is commonly employed by social media users to circumvent platform censorship [55]. To circumvent censorship, our participants "try to be vague as possible to make account[s] last as long as possible" (P25) or used acronyms: "if I tweet about OnlyFans, I'll use the acronym *OF*..." (P8). P41 told us: "there are many words that are forbidden [on OnlyFans]... obviously the words can be taken out of context and not always mean that thing." Therefore, our participants use different spellings (P41, P39, P28) or emojis, symbols, or synonyms for restricted words (P23, P26) to not get *flagged or blocked* on-platform when they are posting or chatting. In some cases, using Algospeak was too cumbersome and participants instead engaged with clients off-platform:

> Once, I got a request to write a dirty story for a client and I had to drop it in the chat and every fourth word was a word that could not go through on OnlyFans... I ended up emailing the story. (P42)

Finally, some tweaked their profile attributes to avoid bias in the censorship process. P34, who intentionally changed their gender (to male) on Instagram explains how it reduced their experiences of censorship: "I've been kind of pushing the envelope with Instagram but I haven't had anything taken down in a while... it's really sad but [changing my gender] made a huge difference."

5.2 Toxic Content

Like at-risk users [90] and other professional content creators [81, 91], 26 of our participants described dealing with different forms of toxic content such as harassment and online hate from fans directly on OnlyFans as well as from other platforms that caused varying degrees of emotional harm. Echoing professional non-sexual content creators' expressions that such hate and harassment is unavoidable [91], our participants noted that toxic content is "just part of being on the platform" (P20) and there are "just clients that have no idea how to message people" (P39).

Harassment and bullying. Participants commonly ($n = 19$) experienced *boundary pushing* where fans become "overly persistent"—exceeding or violating personal or professional limits set by the creators—and P17 indicates that "it's a fine line between being overly insistent and obsession." P1, who

provides messaging service on OnlyFans, was tired of people "pushing to learn more to get a little more familiar." P13 experiences mental distress when fans repeatedly push boundaries and "be disrespectful," making it "complicated" (P25) for creators to effectively manage fan interactions.

Boundary pushing can extend beyond requesting additional information or content, reaching into demands for more of the creator's time. For example, P39 said "people were expecting instant replies all the time from me" and P40 added "you had to be online 24-7, every time you're not answering you're missing out on money and I did not like the stress of that."

8 participants mentioned experiencing other forms of harassment and general bullying on the internet, on OnlyFans, and in-person as a result of their work on OnlyFans. P11 described one of the rude comments: "some people just say that... you're just a cheap whore... why would anyone want to date you if they can just see you for free." These negative experiences take a mental toll (P20) on creators and can do psychological harm [10, 64, 83]: "online harassment is going to affect [me] more mentally than anything else" (P11).

Online hate. Six participants faced online hate based on their appearance, race, or gender. P12 observed "people being transphobic or racist" and like P31, P21 said "a lot of women of color and trans people who are not thin and white, experience the brunt of hate." P31 emphasized the mental toll of increased hate online relating to more visibility:

> With online work, because of how many more people you're seeing, you're exposed to a lot more trolls and time wasters than you would in person because it's just the volume and number of people, especially if you're not white or thin or cis.

P21 noted varying types of hate on different social media platforms, attributing it to the demographics of both users and creators. For example, they perceived Reddit to have an audience "dominated primarily by white American men" and who preferred content from "thin white women" whereas "Twitter is much more diverse with its user base."

5.2.1 Defenses

To address hate and harassment, participants established clear rules ($n = 18$) or blocked ($n = 13$) attackers. Additionally, some participants call on community support (P22: "if someone posted something *mean* I would just retweet it and then the entire internet would ask to delete their account") or respond directly to their harasser themselves (P22: "I'm very skilled at handling [hate and harassment] because I'm sassy"). Others simply ignore the harassment: "I have a really really thick skin from working in this industry for a long time" (P11). This strategy is similar to how other populations manage harassment [81, 90, 91].

Boundary setting with fans. To address the most common form of harassment, boundary pushing, participants made significant efforts to establish *strong* boundaries with fans. P10

received uncomfortable messages and pricing negotiations on their free page, but when they observed that individuals on their paid account exhibited comparatively more polite communication, they deleted the free page. To avoid repeated requests, P4 has added a statement to their bio: "you are welcome to send requests but... I have every right to deny them." During the interview, P2 said, "I mean people are rude for sure sometimes but there isn't a rating system like in so many gig work apps." P38 increased their prices to "weed out assholes" and P36 shared that "they [fans] always wanted the cheaper, the best, even free... I don't have to give them what the broke boys want." P30 clearly communicates their boundaries to fans, but they acknowledge that "expressing boundaries can be hard work and really difficult." For P18, not having to negotiate their established boundaries and saying that "these are boundaries for a reason" gives them a sense of privacy.

Blocking. 13 participants indicated that they employ the platform's block feature when encountering spamming or harassing clients. Upon receiving uncomfortable requests, some participants first tried to warn the fans about their boundaries but would also "restrict them or block them depending on the severity... how aggressive they're being with it" (P9). Similarly, P4 does not react instantly to repeated requests but "ends up blocking people who just kind of go out of hand."

P43 found it easier to handle fan issues online as compared to in-person sex work: "[there] was nothing I couldn't solve by blocking." P11 emphasized the significance of blocking, stating, "if someone is being really crappy towards you, just block them... OnlyFans work is also the biggest privilege within sex work." According to P19, "the only safe thing OnlyFans offers is blocking." P10 further elaborated on the convenience and practicality of the blocking feature:

> *I've blocked way more accounts on these two Only-Fans accounts I've had for a year and a half or two years... once they [fans] start pushing I'm just like no, blocked. I'm not even going to deal with you. It's like no is a full sentence.*

When engaging in off-platform advertising, creators, like P18, "made heavy use of block buttons" to manage heightened harassment due to sex work stigma.

5.3 Content Leakage

Content leaks [90] were the next most common ($n = 24$) risk our participants reported. Although creator content on Only-Fans is paywalled (§ 2), it can be illicitly captured in violation of the ToS and the creators' consent by screenshotting and downloading. Additionally, OnlyFans itself can be hacked: in February 2020, an estimated 1.4 to 4 terabytes of OnlyFans videos and photos were shared across social media [28].

Multiple participants viewed content leaks as inevitable: "there's always a possibility for that [content leak]" (P4), "no matter what you do leaks are inevitable" (P21), "it's [content is] going to eventually get stolen"(P27), and "it's just

something that unfortunately is part of the job" (P31).

Prior work on image-based sexual abuse finds that perpetrators may leak content in order to raise their own social status, using the content they non-consensually obtain and distribute as a form of social currency [31, 49]. In line with this, P36 describes their perception that:

> *The usual porn user is kind of like a hoarder of porn, they save all the porn and try to put them again on platforms to share them around like if they were kind of football cards.*

In the transactional context of OnlyFans, content theft deprives creators of income. Attackers may thus be motivated to engage in content leaks to obtain content for free or to leverage fraudulently obtained content to profit financially, as detailed next.

5.3.1 Harms of Content Leakage

Content leakage can cause harm in a variety of ways, and enable additional attacks.

Privacy violation & image-based sexual abuse. While Only-Fans creators consent to post and share their intimate content within the platform's transactional context, they do not consent to it being further shared nor viewed outside the context of this platform and transaction. Thus, leaks of creators' intimate content are a form of image-based sexual abuse [49, 90]. Such abuse is a violation in itself and has significant emotional and psychological consequences [49, 80, 100]. The permanence of intimate content leaks (e.g., P41 describes fearing their leaked content "being on the internet forever") and severity of consequences from such leaks make them a "severe" form of hate and harassment, as defined by prior work [90].

Outing. Attacker motivations for content leaks may vary. Attackers may also non-consensually distribute others' intimate content as a *targeted* form of abuse by doxxing creators and/or outing them as members of stigmatized communities (e.g., the sex worker and/or LGBTQ+ community). Attackers may do so by leaking creators' content and linking it to their personal information [22], which they can discover through a variety of means including using "facial recognition software... [to] find your old [pictures] from high school and... your name and where you grew up" (P9). Such attacks can cause relationship, reputational, and physical harms [83]. Similar to 13 other participants, safety to P11 is:

> *...being able to do what I do and not worry about being doxxed and having stalkers and people showing up at my house to potentially do harm to me.*

P9 mentioned the potential "safety risk" of "families kicking [creators] out of the house" if they found out about their work, while P22 linked "getting outed as a sex worker" to "receiving professional retribution." P1 fears how permanent content leaks could impact their ability to get other jobs in the future after they finish their schooling, because the "engineering world does not look too highly upon OnlyFans."

Chargebacks. Chargebacks happen when a purchaser falsely disputes a product charge, often by reporting credit card fraud or claiming they do not recognize the payments on a bank statement. The sex industry has particularly high chargeback rates due to the discreet and stigmatized nature of purchasing sexual content [21, 86, 94, 105]. For OnlyFans creators, chargebacks are passed on directly to creators, meaning a client gets access to their content without ultimately paying for it, violating their consent and impacting their revenue. Four participants in our work described being victimized by chargebacks. P23 explains:

> So, if a guy gives you loads and loads of money and then decides to report that he got his credit card stolen, then you don't get that money back.

P7 similarly described their experiences with "the people who come in and drop a whole bunch of money and then back track it." However, they proactively started charging extra for content they particularly wanted to protect and told us: "it would be nice if OnlyFans fixed that loophole."

Catfishing. Catfishing[7] is an impersonation attack [90] in which an attacker creates a fake online profile to trick people, usually for financial gain. In the context of OnlyFans, catfishing means someone copying creators' videos or posts to impersonate their OnlyFans identify elsewhere and engage in fraud. P9 described that scammers create fake Instagram accounts using OnlyFans creators' content to deceive fans, which damages creators' reputations and harms their followers. P12 and P22 shared instances where they found catfish profiles made using their images on Grindr. Additionally, P11 encountered their stolen images on Reddit, where the account attempted to defraud their fans. Fortunately, they were able to have these accounts taken down as they had watermarked their images (see below).

5.3.2 Defenses

In response to the threat of content leakage and its cascading attacks, our participants employed a variety of protections. P9 mentioned how an OnlyFans creator has to engage in adversarial thinking and "get in the minds of people who could potentially want to harm you." They summarized some of the technical protection strategies they were aware of:

> There's a lot of services out there that will get your name scrubbed off of voter records or at least your address so no one can show up at your house... be very careful with not having your Snapchat location [on]... photos taken on iPhone have your information encrypted inside of them.

Limiting prominence (and profit). Seven of our less prominent participants intentionally aimed to limit their popularity,

motivated by the fear of having their location discovered or having their non-OnlyFans professional progress hindered. P4 and P7 expressed contentment with not having a significantly large following as the "stress from that is not great." Given their limited prominence, P7 was confident that their content would not be widely shared, while bigger accounts have to "deal with large chargebacks and take the heat." P23 was worried that increased visibility could impact their interpersonal relationships and privacy:

> It exposes me to much more... there's much more exposure, there's not much privacy, any ex-boyfriend or stalker can look at my OnlyFans and they're seeing my sex work in a way they couldn't before.

On the other hand, because OnlyFans lacks internal discovery and creators engage in cross-platform advertising, managing off-platform prominence becomes challenging: "my advertising is everywhere, and I know TikTok advertises more wherever your SIM card is based and I'm just using my regular phone so they're pushing my videos out more locally" (P24).

Internet presence monitoring. 17 participants actively implemented internet presence monitoring schemes to check for leaked content. P40 said: "the likelihood of it [content] getting leaked was giving me anxiety, OnlyFans leaks are very common and I was constantly googling if my shit got leaked." P2 searches variations of their name on Google to see "what's out there and delete-able." P35 does the same, and also checks PornHub and XVideos for stolen content to "not have videos out there for free." P9 uses image search to look for their leaked content on Reddit and Google. They also leverage a subscription service that searches for their name and scrubs identifiable information from the internet.

11 participants reported using DMCA takedown services[8] to remove stolen content from the internet. However, some participants hesitated to use these services either because of skepticism (P27: "it's not even like a guarantee"), the service's convoluted process and potential non-responsiveness (P33: "this might be really complicated, what if they don't respond?"), or the financial burden (P40: "this would take a lot of cash").

Two participants (P12, P22) who experienced catfishing received help from others in reporting catfishers. For example, when P22's photos were used to create a fake Grindr account, they made their fans aware of the catfish by tweeting and asking people to report the account.

During the interview, P8 and P40 mentioned the feature on OnlyFans that supposedly prevents users from taking screenshots or recording content. P8 described the similarity to platforms like Netflix, where attempting to screenshot results in a black screen, adding "I never actually tried it out to see if it works, I sort of just trusted it." However, OnlyFans does not currently have any features in place to prevent screen

---

[7]Prior literature has sometimes referred to similar attacks using an attacker term of art—"eWhoring"—to specify catfishing fraud conducted using "sexualized images" [51]. The use of the term "whore" is however disrespectful and therefore we opt not to use it [97].

[8]A DMCA takedown happens when the owner of content claims a copyright violation and requests a platform remove the content, https://www.dmca.com/

recording. This highlights the misconceptions that can exist in participants' mental models of platform-provided security tools.

Watermarking. Six participants relied on OnlyFans' in-platform content watermarking to deter content theft (§ 5.3). However, 14 participants practiced proactive watermarking to protect their content from theft and unauthorized use, as well as to trace specific people who stole the content via personalized watermarks. OnlyFans does not challenge intellectual property infringement on behalf of its creators, and so the responsibility for recovery is left to individuals. Prominence also influences creators' decisions about how to protect their content. P43 deems the extensive efforts of watermarking "probably not worth it at my level of visibility." P11, P20, and P37 recognize that leaked watermarked content also serve as a kind of promotion, as people can identify the creators through the username on the content, which potentially helps recover lost income.

Keeping online identities separate. Similar to online account management strategies of in-person sex workers [67], many participants ($n = 22$) kept their personal social media accounts separate from their OnlyFans accounts or the social media accounts used to promote OnlyFans. P25 keeps their social media accounts separate to prevent platforms from suggesting their profile to unwanted individuals. P33 reduced their posts on social media after an incident:

> While I was on holiday I posted a picture of my-self on the beach on my main Instagram which my friends and family see and on my porn Instagram, and I tagged my location, and then someone found the main, and I was like that was so dumb, I can't believe I did that.

Participants' efforts to separate their online identities exist on a spectrum. Seven participants who were already open about their sex work in their personal lives did not worry about the risk of getting outed. For example, P32 said "I was already face out, so I wasn't necessarily concerned about clients taking screenshots of private photos that were revealing my face and sharing them and outing me." One participant (P34) even switched their personal Instagram account to a sex work account and deliberately outed themselves.

Saving username on different platforms. To minimize impersonation risk, five participants strategically saved their usernames on other platforms. To that end, P33 said: "I did worry that someone would want to impersonate me on there which is why I thought it'd make sense to get the username [on another site]." Creators sometimes encounter challenges in securing their usernames, as highlighted by P16: "someone took my username, some cis guy from UK who doesn't even do anything, he just watches people, he doesn't post anything, he doesn't stream. It's just really annoying."

Limiting identifying information on OnlyFans. Our study participants employed various strategies to hide any identify-ing information online, as P7 mentioned: "Sex workers are more at risk as soon as they put nudity online... I do more to keep my identity hidden on my OnlyFans account rather than my personal accounts."

To that end, 10 participants concealed their faces in their posts to maintain anonymity. For instance, P4 angles their camera to exclude their face in nude content, but reveals their face when portraying a cosplay character. P1 hesitates to use blurring software to conceal their face due to concerns about potential de-blurring, and instead opts to physically cover and crop their face before posting for "double security." P33 shared an incident where they did not hide their face in the "most clever way":

> Sometimes I hide the top part of my face and some-times I hide the bottom half of my face, and I had a guy who had stitched together the top half and the bottom half and he sent it to me like look, I've made a picture of your face, this is what you look like!

P3 avoids posting identical content on their OnlyFans and NSFW Twitter accounts to mitigate the risk of reverse image searches linking pictures of their face across platforms.

16 participants proactively used aliases to protect their legal name. P13, a cosplayer, mentioned: "...with cosplayers, it's so easy to get people who are just creepy and trying to find my Facebook account. Then, they were finding out where I worked, where I lived... going by a different name adds a layer of security to anything that I do."

18 participants concealed their location on OnlyFans using security-focused tools such as VPNs. P17 felt "super secure" using VPNs because "just in case somebody was to get a weird idea to try to figure out my location it wouldn't even ping off to what state I am in." P18, on using a VPN, said "I'm assuming I'm not the only one who's heard horror stories. So like just making sure to have like a level of semi-anonymity in there." Some participants further hid their location by, e.g., avoiding taking photos near identifiable landmarks (P9, P10), turning off geo-location from images and videos (P10, P17, P19), and removing image meta-data (P16).

Managing outness offline. 16 participants proactively avoided disclosing ("a polite omission" - P12) their Only-Fans work with certain people in their life, primarily because of the stigma around sex work (P1: "this person has no dig-nity, no morals") or to reduce relationship harm (P11: "I don't think that they [parents] would even understand what is going on") [83]. Maintaining secrecy posed challenges for some; P10 consistently faced stress while upholding a "cover story." Similar to 13 participants, P8 opted to identify as a *content creator* to navigate interpersonal and professional relation-ships without revealing their OnlyFans association.

5.4 Adopting Digital Safety Practices

In contrast to prior work on non-sexual professional content creators [81], where two-thirds of participants adopted protec-

tive practices only *reactively*, we observe that before joining OnlyFans, all of our participants adopted one or more protective practices to avoid content leakage (§ 5.3.2). While most participants focused on protective practices to preemptively keep their online identities separate, some also preemptively limited their prominence, set up internet presence monitoring tools, watermarked content, limited identifying information on OnlyFans, and managed outness about OnlyFans work offline. Additionally, some participants became early adopters of strong boundaries to defend against toxic content (§ 5.2.1).

Learning sources. Participants learned these initial practices from fragmented sources of digital safety advice [76] and from their own "common sense" (P10) and technical skills, like "tech nerd" P16 and "Google queen" P20. For example, P8 engaged in preliminary research on OnlyFans' safety mechanisms to assess the platform's security features:

> *So the types of like encryption that they use or how they watermark photos... definitely didn't research in depth but like very slightly like legality and making a claim in your account about like if people were to redistribute it [content].*

Friends also served as sources of knowledge. For example, P13 learned their safety practices through a friend's doxxing experience in addition to their own stalking experience as a younger cosplayer, and P22 received information from friends about general internet safety. P4 on the other hand learned to watermark any content they share from observing their friend's experience of content leaks, and seeing their friend use watermarks to help trace who had leaked their content.

As reported in prior work [7, 87, 88, 95], our participants also learned certain safety mechanisms from OnlyFans and online sex-work-specific communities. 16 participants derived reassurance and support from being an active member of a strong and closely-knit online community of content creators, where they learned about safety and privacy, content creation and marketing strategies, and other OnlyFans-related experiences. P13 felt from the beginning that they could establish boundaries about the type of content they create when they saw others opting out of creating very explicit content. From being part of the OnlyFans community, P36 learned tips to grow as an online sex worker, perseverance, as well as boundary setting ("they taught me that I don't have to cater to what the boys want"). Likewise, P21 learned about "basic things such as knowing your boundaries, to not feel uncomfortable, how to protect yourself online, and a little bit of anonymity" through a sex work industry video.

Preparedness. Participants mentioned that they learned to be proactive in their defensive approach because of the societal stigma around sex work: "there's definitely a little bit of a stigma... I can't exactly go on my Instagram and start broadcasting or I'm gonna get a label [for] myself" (P1). Recognizing potential risks, P9 advocated for proactive preparation, stating, "ignorance is not bliss in this situation."

While participants were very proactive in their security posture, they also engaged in ongoing learning. To that end, 13 participants actively monitor Twitter and Reddit to learn about safety tips on OnlyFans: "I follow like keywords for safety and security on OnlyFans" (P13), "there's a subreddit and there's a lot of really knowledgeable creators on there that can tell you what to do and how to stay safe" (P9), and "I check reddit, check top content creators, see how they feel about safety... they will post scammer IDs and stuff like that so you can go ahead and pre-block them" (P7).

6 Discussion

Through 43 semi-structured interviews with sexual content creators on OnlyFans, we documented creators' online threats and safety strategies. In this section, we synthesize our results using our research questions as a guide. We conclude with an exploration of the opportunities for better security affordances.

6.1 Creators' Threat Landscape: A Hate and Harassment Framework Perspective

We used Thomas et al.'s online hate and harassment framework [90] as a lens to understand the threat landscape of OnlyFans creators (RQ1). Our analysis reveals that OnlyFans creators experience digital risks common to other "at-risk" users—toxic content and content leakage—as well as more unique risks from platform censorship and precarity.

Similar to non-sexual content creators [81, 91], our participants experienced a high volume and expressed a high degree of concern about hate and harassment attacks. In contrast with non-sexual content creators who prior work finds predominantly face attacks related to toxic content [91], our participants experienced the most frequent and concerning threats from platform censorship and precarity.

Prior work finds that some non-sexual content creators worry about violating platform ToS and resulting platform censorship [81]. Our participants expressed such concerns in greater volume and with greater concern, and additionally raised concern about platform censorship as a result of biased moderation of ToS-compliant content due to the explicit and stigmatized nature of the content they share. Further, our participants expressed significant concern about the precarity of OnlyFans as a platform. Combined with threats creators face from people—of toxic content and content leakage—the threats of platform censorship and precarity foster a strong sense of insecurity among creators (§ 5.1).

6.2 Role of Stigma and Digital Prominence in Creators' Threat Landscape

Our results show that awareness of sex work stigma prompts OnlyFans creators to proactively anticipate and prepare for

online risks (§ 5.4), in contrast to non-sexual content creators who tend to approach digital safety in a more reactive fashion [81]. Participants often turned to their community, either passively through following community-members on social media, or actively by seeking out support from friends and community members. Our findings support prior work illustrating the influence of stigma on digital communities, where shared experiences of stigma foster a strong sense of solidarity and collective support [7, 33, 63]. Together this highlights the role of stigma in digital security education and offers evidence for its influence on differential threat models and protective behaviors between otherwise similar groups of end users (e.g., between digital content creators who do and do not do stigmatized work).

Stigma also intersects with prominence, influencing creators' online risks. Prior research revealed that social media content creators, even without the added layer of sex work stigma, faced increased risks with growing popularity [81]. Experts advise general Internet users to approach online participation with a minimalist mindset [100]. However, the content creation industry inherently requires information sharing for monetary gain, rendering a minimalist mindset impractical. Particularly on OnlyFans, which lacks internal searchability, off-platform information sharing becomes a necessity for creators to promote their content, attract subscribers, and generate revenue.

OnlyFans creators actively strive to maintain separate online identities to overcome challenges posed by stigma and balance the necessity of increased prominence with preservation of personal privacy. However, effectively maintaining privacy while prominently engaging online remains a challenge. Future work may seek to address this challenge by developing privacy-preserving technology to better support persona management [46]. Such work can help non-sexual professional content creators [81] who must manage their identities due to prominence, recreational content creators who are especially at-risk because of a stigmatized identity [61], and professional sexual content creators such as the OnlyFans creators we study who are at risk due to both prominence and stigmatized labor.

6.3 Platforms' Impacts on Creator Risks

Experts recommend using platform-provided tools to moderate abusive messages online [100]. However, both sexual and non-sexual content creators are skeptical about the support offered by platform tools. Although non-sexual content creators used keyword blocklists and manual reviewing to address hate and harassment, they expressed skepticism about the effectiveness of reporting users [91].

Similarly, while OnlyFans offers reporting options for addressing hate and harassment, our participants predominantly opted for *blocking* fans in response to such incidents, which

provided them a sense of empowerment[9] and avoided reliance on the platform for justice. We hypothesize that our participants prefer not to rely on the platform due to experiences of platform precarity. For example, when OnlyFans suddenly banned explicit content (this ban was later reverted, see § 2), creators were left with a sense of helplessness (§ 5.1).

Additionally, some platform policies increased participants' apprehension that the platform would adequately safeguard their interests. For example, chargebacks pose a significant challenge in the sex industry. OnlyFans passes such risk on to creators directly: the community guidelines on chargebacks state that "any amounts that users seek to refund or chargeback will be reported to the creator and will be removed from the creator's income" [24]. This places a considerable financial burden on the creators, treating them as de-facto e-commerce merchants without affording them control over payments—typical to an e-commerce merchant—or the ability to contest chargebacks effectively.

Beyond OnlyFans itself, our participants experienced and/or perceived that mainstream platforms may be biased toward shadowbanning their content or deplatforming them due to the stigmatized nature of their labor and the content they share (§ 5.1). Consequently, they might be reluctant to depend on reporting tools provided by platforms they perceive as indifferent to their safety.

6.4 Suggestions for Solutions

Affordances for diversification. Our participants practiced platform diversification strategies (e.g., saving usernames on other platforms) to avoid sole dependence on a single source, such as OnlyFans, for income (§ 5.1.1). The precarious nature of sex work platforms also means that an account and all data within—photos, videos, messages, connections with fans—could be deleted arbitrarily and without notice [4]. However, diversification can be extremely challenging. It requires a significant time commitment and effort to maintain multiple-platform account activities effectively [44]. For instance, during OnlyFans' ban on sexual content (§ 2), participants expressed a desire for sufficient time to retrieve pending deposits and "safely delete content or transfer it [content] to other platform" (P4). Moving to alternative platforms requires critical evaluation based on factors such as popularity (in comparison to OnlyFans) and fan engagement. Only half of P11's fans were willing to move to other platform and P12 said "creators did not experience the same success that they had on OnlyFans" when creating new accounts, attributing it to the differing audience presence.

There have been various attempts to increase consumer control over personal data. For example, the General Data Protection Regulation (GDPR) Article 20 [37] asserts individuals' right to "receive their personal data in a machine-readable format and transfer that directly from one data controller to

---

[9] "If someone is being really crappy towards you, just block them" - P11

another (wherever possible)." This *right to data portability*, in theory, offers consumers significantly reduced costs for switching between digital service providers [57], and could play a crucial role in enabling OnlyFans creators to diversify their online presence and mitigate risks associated with platform dependence.

Despite the fact that the GDPR has been in force for several years, the right to data portability is still, in practice, largely unusable [89]. Requests can be cumbersome, data can be in unusual formats, and new platforms rarely have a method for ingesting data from another. With further advancements in enforcement and usability, data portability can offer creators stability and facilitate seamless content and data transfer.

Internet monitoring. OnlyFans' ToS recognize redistribution of content by fans on other platforms as a violation, but it neither prevents users from recording their screen nor enforces intellectual property rights on behalf of creators [34].

To search for and take down leaked content, participants practiced extensive "internet monitoring" (§ 5.3.2). This practice is time-consuming and emotionally draining, taking creators' time away from running a profitable business and limiting freedom of self-expression. Although the platform provides a free watermarking service to creators and actively scans the internet to identify external websites hosting large amounts of stolen content from its creators [74], there remain opportunities for OnlyFans to proactively prevent unauthorized content capture. For instance, OnlyFans could implement features that mimic the behavior of streaming websites like Netflix, which detect screen recording and cause the screen to turn black. Additionally, leveraging existing digital rights management approaches, such as content fingerprinting [59,98], could help automate the detection of unauthorized content use. Such advancements would also streamline the process of issuing automated takedown notices [41].

Catfish prevention. Prior work has described various types of catfishing attacks. For example, romance scams involve an attacker using stolen photographs and false identities to develop a fake romantic relationship with their victim to defraud them of large sums of money [50,102]. Other attacks involve selling photos and videos with sexual content of another person to third parties, sometimes while pretending to be the subject [51].

Preventative mechanisms from this area of research could be translated to help OnlyFans creators cope with catfishing (§ 5.3): 1) automatic detection and blocking of impersonators through a hashed image dataset shared across social media and dating sites; 2) verification processes where creators could apply their verified status on OnlyFans to other platforms, or adding a layer of security through physical assets (one-time-passwords on cellular devices); 3) active and regular reporting of catfishers (by verified accounts and OnlyFans automatically generating reports to platforms like Twitter, Facebook, or Grindr, requesting investigation into reported catfish or scam accounts); and 4) asking creators to link their genuine social media accounts and report fraudulent accounts. These mechanisms align with established phishing detection and mitigation techniques: content classification and blocklisting [62, 101, 103, 104], user training and education [29], content take-down [3], and direct reports from users [40].

Chargebacks and financial harm reduction. Although only four participants discussed the unique risk of chargebacks in our study (§ 5.3), chargebacks on OnlyFans are common largely because of sex work stigma [11]. For example, a fan may request a refund for their OnlyFans transactions to hide it from their spouse on their bank statement. To free creators from bearing the brunt of chargeback liability, OnlyFans could challenge chargebacks on behalf of their creators or implement a platform currency: fans purchase tokens from OnlyFans, and then use the tokens to pay creators (who are then, in turn, paid proportional to the tokens they earn). The streaming platform Twitch has already implemented a similar system to challenge chargebacks [32].

7 Conclusion

We interviewed 43 OnlyFans creators with and without prior sex work experience to understand their online risks on- and off-platform. Creators were most concerned about platform risks of shadowbanning and deplatforming from platform censorship and precarity which affected their creativity and emotional and financial safety. Despite facing a substantial amount of online hate and harassment, they demonstrated preparedness to handle toxic content by proactively setting strong boundaries with fans and blocking abusive fans. Creators found intimate content leaks prevalent and unavoidable and employed a range of protective measures from limiting prominence to keeping their online identities separate. Our study emphasizes the significance of platform risks which are concerning for sexual, non-sexual, and marginalized creators. By incorporating platform-specific risks, we can better understand the challenges faced by at-risk populations and enhance security measures to promote safer online environments.

Acknowledgements

Author Contributions

VH, AM, and EMR developed the interview protocol and collected the interview data. AS and VH analyzed the interview data. AS, VH, AD, AM, and EMR wrote the paper.

References

[1] Kendra Albert, Emily Armbruster, Elizabeth Brundige, Elizabeth Denning, Kimberly Kim, Lorelei Lee, Lindsey Ruff, Korica Simon, and Yueyu Yang. Fosta in legal context. *CoLUM. Hum. RTs. L. REV.*, 52:1084, 2020.

[2] What Does ALGOSPEAK Mean? https://www.cyberdefinitions.com/definitions/ALGOSPEAK. (Accessed on 06/05/2023).

[3] Eihal Alowaisheq, Peng Wang, Sumayah Alrwais, Xiaojing Liao, XiaoFeng Wang, Tasneem Alowaisheq, Xianghang Mi, Siyuan Tang, and Baojun Liu. Cracking the wall of confinement: Understanding and analyzing malicious domain. In *Proceedings of the 28th Network and Distributed System Security Symposium (NDSS)*, 2019.

[4] Carolina Are and Pam Briggs. The emotional and financial impact of de-platforming on creators at the margins. *Social Media+ Society*, 9(1):20563051231155103, 2023.

[5] Carolina Are and Susanna Paasonen. Sex in the shadows of celebrity. *Porn Studies*, 8(4):411–419, 2021.

[6] Y. Aydinalp. Struggle for Power and Representation: The Impact of Digital Platform Structures on Sex Workers' Agency. Master's thesis, Frei University Brussels, 2022.

[7] Hanna Barakat and Elissa M Redmiles. Community under surveillance: Impacts of marginalization on an online labor forum. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 16, pages 12–21, 2022.

[8] Eloise Barry. Why OnlyFans Suddenly Reversed its Decision to Ban Sexual Content. https://time.com/6092947/onlyfans-sexual-content-ban/, August 2021. (Accessed on 05/16/2023).

[9] Catherine Barwulor, Allison McDonald, Eszter Hargittai, and Elissa M Redmiles. "Disadvantaged in the American-dominated internet": Sex, Work, and Technology. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–16, 2021.

[10] Corey H Basch, Lorie Donelle, Joseph Fera, and Christie Jaime. Deconstructing TikTok videos on mental health: cross-sectional, descriptive content analysis. *JMIR formative research*, 6(5):e38340, 2022.

[11] Bianca Beebe. "shut up and take my money!": Revenue chokepoints, platform governance, and sex workers' financial exclusion. *International Journal of Gender, Sexuality and Law*, 2(1):140–170, 2022.

[12] Elizabeth Bernstein. *Temporarily yours: Intimacy, authenticity, and the commerce of sex*. University of Chicago Press, 2010.

[13] Rasika Bhalerao and Damon McCoy. An Analysis of Terms of Service and Official Policies with Respect to Sex Work. In *2022 IEEE International Symposium on Technology and Society (ISTAS)*, volume 1, pages 1–14. IEEE, 2022.

[14] Marta Biino. How to make money on OnlyFans, according to creators. https://www.businessinsider.in/tech/news/how-to-make-money-on-onlyfans-according-to-creators/articleshow/89727848.cms, March 2023. (Accessed on 05/30/2023).

[15] Danielle Blunt, Stefanie Duguay, Tarleton Gillespie, Sinnamon Love, and Clarissa Smith. Deplatforming sex: A roundtable conversation. *Porn Studies*, 8(4):420–438, 2021.

[16] Danielle Blunt and Zahra Stardust. Automating Whorephobia: sex, technology and the violence of deplatforming: An interview with Hacking//Hustling. *Porn Studies*, 8(4):350–366, 2021.

[17] Danielle Blunt and Ariel Wolf. Erased: The impact of FOSTA-SESTA and the removal of Backpage on sex workers. *Anti-trafficking review*, (14):117–121, 2020.

[18] Ross Bonifacio and Donghee Yvette Wohn. Digital patronage platforms. In *Conference companion publication of the 2020 on computer supported cooperative work and social computing*, pages 221–226, 2020.

[19] David R Brake. Are we all online content creators now? Web 2.0 and digital divides. *Journal of Computer-Mediated Communication*, 19(3):591–609, 2014.

[20] Virginia Braun and Victoria Clarke. Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2):77–101, 2006.

[21] Nadeem Brown. *Dirty Little Secret: What No One Ever Tells You About Internet Adult Entertainment Industry*. Iuniverse, 2005.

[22] Vicky Bungay and Adrian Guta. Strategies and challenges in preventing violence against Canadian indoor sex workers. *American journal of public health*, 108(3):393–398, 2018.

[23] Rosie Campbell, Teela Sanders, Jane Scoular, Jane Pitcher, and Stewart Cunningham. Risking safety and rights: online sex work, crimes and 'blended safety repertoires'. *The British journal of sociology*, 70(4):1539–1560, 2019.

[24] OnlyFans Help Center. Chargebacks / Credit Card Disputes. https://onlyfans.com/help/2/12/45. (Accessed on 06/05/2023).

[25] Juliet Corbin and Anselm L. Strauss. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. SAGE Publications, USA, fourth edition, 2014.

[26] David Craig. 27. Creator management in the social media entertainment industry. In *Making Media*, pages 363–374. Amsterdam University Press, Amsterdam, Netherlands, 2019.

[27] Scott Cunningham and Todd D Kendall. Prostitution 2.0: The changing face of sex work. *Journal of Urban Economics*, 69(3):273–287, 2011.

[28] Anthony Cuthbertson. OnlyFans leak: Huge file of stolen porn dumped online. https://www.independent.co.uk/tech/onlyfans-leak-porn-hack-videos-photos-file-a9365406.html, February 2020. (Accessed on 06/05/2023).

[29] Ronald C Dodge Jr, Curtis Carver, and Aaron J Ferguson. Phishing for user security awareness. *computers & security*, 26(1):73–80, 2007.

[30] Gwyn Easterbrook-Smith. OnlyFans as gig-economy work: A nexus of precarity and stigma. *Porn studies*, 10(3):252–267, 2023.

[31] Asia A Eaton and Clare McGlynn. The psychology of nonconsensual porn: Understanding and addressing a growing form of sexual violence. *Policy Insights from the Behavioral and Brain Sciences*, 7(2):190–197, 2020.

[32] Monica Eaton-Cardone. Chargeback Challenges On The Rise For Twitch Streamers. https://www.forbes.com/sites/forbestechcouncil/2020/02/28/chargeback-challenges-on-the-rise-for-twitch-streamers/, February 2020. (Accessed on 06/05/2023).

[33] Casey Fiesler and Amy S Bruckman. Creativity, copyright, and close-knit communities: a case study of social norm formation and enforcement. *Proceedings of the ACM on Human-Computer Interaction*, 3(GROUP):1–24, 2019.

[34] Amy Ford. "Id be Mortified if Someone Thought I was Putting my Bum Online for Some Fucking Gucci Sandals": Liminality, Ownership and Identity–A Case-study of Sex Work on OnlyFans. *DEARCADH*, page 27, 2022.

[35] X (formerly Twitter). Non-consensual nudity policy. https://help.twitter.com/en/rules-and-policies/intimate-media, December 2021. (Accessed on 12/12/2023).

[36] Gabriella Garcia. *[SHIFT+CTRL+END+DELETE] Encoded Stigma*. PhD thesis, New York University, 2020.

[37] Art. 20 GDPR – Right to data portability. https://gdpr-info.eu/art-20-gdpr/. (Accessed on 06/05/2023).

[38] Deyan Georgiev. 18 Shocking OnlyFans Statistics to Show How Big It Is [2023]. https://techjury.net/blog/onlyfans-statistics/#gref, March 2023. (Accessed on 04/19/2023).

[39] Steven M Gillespie, Andrew Jones, Kasia Uzieblo, Carlo Garofalo, and Eric Robinson. Coping using sex during the coronavirus disease 2019 (covid-19) outbreak in the united kingdom. *The Journal of Sexual Medicine*, 18(1):50–62, 2021.

[40] Google. Report Phishing Page. https://safebrowsing.google.com/safebrowsing/report_phish/. (Accessed on 05/30/2023).

[41] Joanne E Gray and Nicolas P Suzor. Playing with machines: Using machine learning to understand automated copyright enforcement at scale. *Big Data & Society*, 7(1):2053951720919963, 2020.

[42] Andrew Griffin. As OnlyFans bans 'sexually-explicit content', what else is actually on there? https://www.independent.co.uk/tech/onlyfans-porn-ban-other-content-b1905849.html, August 2021. (Accessed on 05/16/2023).

[43] Oliver L Haimson, Daniel Delmonaco, Peipei Nie, and Andrea Wegner. Disproportionate removals and differing content moderation experiences for conservative, transgender, and black social media users: Marginalization and moderation gray areas. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2):1–35, 2021.

[44] Lee Hair, Ross Bonifacio, and Donghee Yvette Wohn. Multi-platform practices among digital patronage creators. *Convergence*, 28(5):1438–1456, 2022.

[45] Vaughn Hamilton, Hanna Barakat, and Elissa M Redmiles. Risk, resilience and reward: Impacts of shifting to digital sex work. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW2):1–37, 2022.

[46] Vaughn Hamilton, Gabriel Kaptchuk, Allison McDonald, and Elissa M Redmiles. Safer digital intimacy for sex workers and beyond: A technical research agenda. *IEEE Security & Privacy*, 2023.

[47] Vaughn Hamilton, Ananta Soneji, Allison Mcdonald, and Elissa M Redmiles. "Nudes? Shouldn't I charge for these?": Motivations of New Sexual Content Creators on OnlyFans. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2023.

[48] Eszter Hargittai and Yuli Patrick Hsieh. Succinct survey measures of web-use skills. *Social Science Computer Review*, 30(1):95–107, 2012.

[49] Nicola Henry, Clare McGlynn, Asher Flynn, Kelly Johnson, Anastasia Powell, and Adrian J Scott. *Image-based sexual abuse: A study on the causes and consequences of non-consensual nude or sexual imagery*. Routledge, 2020.

[50] JingMin Huang, Gianluca Stringhini, and Peng Yong. Quit playing games with my heart: Understanding online dating scams. In *Detection of Intrusions and Malware, and Vulnerability Assessment: 12th International Conference, DIMVA 2015, Milan, Italy, July 9-10, 2015, Proceedings 12*, pages 216–236. Springer, 2015.

[51] Alice Hutchings and Sergio Pastrana. Understanding ewhoring. In *2019 IEEE European Symposium on Security and Privacy (EuroS&P)*, pages 201–214. IEEE, 2019.

[52] Shagun Jhaver, Quan Ze Chen, Detlef Knauss, and Amy X Zhang. Designing word filter tools for creator-led comment moderation. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pages 1–21, 2022.

[53] Angela Jones. Sex work in a digital era. *Sociology Compass*, 9(7):558–570, 2015.

[54] Angela Jones. *Camming: Money, power, and pleasure in the sex work industry*. NYU Press, 2020.

[55] Daniel Klug, Ella Steen, and Kathryn Yurechko. How Algorithm Awareness Impacts Algospeak Use on TikTok. In *Companion Proceedings of the ACM Web Conference 2023*, pages 234–237, 2023.

[56] Susanne Kopf. "Rewarding good creators": Corporate social media discourse on monetization schemes for content creators. *Social Media+ Society*, 6(4):2056305120969877, 2020.

[57] Sophie Kuebler-Wachendorff, Robert Luzsa, Johann Kranz, Stefan Mager, Emmanuel Syrmoudis, Susanne Mayr, and Jens Grossklags. The right to data portability: Conception, status quo, and future directions. *Informatik Spektrum*, 44:264–272, 2021.

[58] J Richard Landis and Gary G Koch. The measurement of observer agreement for categorical data. *biometrics*, 33(1):159–174, 1977.

[59] Sunil Lee and Chang D Yoo. Robust video fingerprinting for content-based video identification. *IEEE Transactions on Circuits and Systems for Video Technology*, 18(7):983–988, 2008.

[60] Carol Leigh. Inventing sex work. In *Whores and other feminists*, pages 225–231. Routledge, 2013.

[61] Ada Lerner, Helen Yuxun He, Anna Kawakami, Silvia Catherine Zeamer, and Roberto Hoyle. Privacy and activism in the transgender community. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2020.

[62] Bin Liang, Miaoqiang Su, Wei You, Wenchang Shi, and Gang Yang. Cracking classifiers for evasion: A case study on the google's phishing pages filter. In *Proceedings of the 25th International Conference on World Wide Web*, pages 345–356, 2016.

[63] Jessa Lingel and Danah Boyd. "Keep it secret, keep it safe": Information poverty, information norms, and stigma. *Journal of the American Society for Information Science and Technology*, 64(5):981–991, 2013.

[64] Sonia Livingstone. Taking risky opportunities in youthful content creation: teenagers' use of social networking sites for intimacy, privacy and self-expression. *New media & society*, 10(3):393–411, 2008.

[65] Janne Tapani Matikainen. Motivations for content generation in social media. Participations: Journal of Audience and Reception Studies, 2015.

[66] Samuel Mayworm, Michael Ann DeVito, Dan Delmonaco, Hibby Thach, and Oliver L Haimson. Content moderation folk theories and perceptions of platform spirit among marginalized social media users. *ACM Transactions on Social Computing*, 2023.

[67] Allison McDonald, Catherine Barwulor, Michelle L Mazurek, Florian Schaub, and Elissa M Redmiles. "It's stressful having all these phones": Investigating Sex

Workers' Safety Goals, Risks, and Practices Online. In *30th USENIX Security Symposium*, pages 375–392. USENIX, 2021.

[68] Clare McGlynn and Erika Rackley. Image-based sexual abuse. *Oxford Journal of Legal Studies*, 37(3):534–561, 2017.

[69] Jill Nagle. *Whores and Other Feminists*. Routledge, New York, 2013.

[70] Bonnie A Nardi, Diane J Schiano, Michelle Gumbrecht, and Luke Swartz. Why we blog. *Communications of the ACM*, 47(12):41–46, 2004.

[71] OnlyFans About. https://onlyfans.com/about. (Accessed on 06/05/2023).

[72] OnlyFans Terms of Service. https://onlyfans.com/terms. (Accessed on 11/30/2023).

[73] Bahiyah Omar and Wang Dequan. Watch, share or create: The influence of personality traits and user motivation on TikTok mobile video usage. *International Journal of Interactive Mobile Technologies (iJIM)*, 14(04):121–137, 2020.

[74] OnlyFans. Helping Creators Protect Their Copyright. https://onlyfans.com/transparency-center/helping. (Accessed on 01/12/2024).

[75] Helen M Rand. Challenging the invisibility of sex work in digital labour politics. *Feminist Review*, 123(1):40–55, 2019.

[76] Elissa M Redmiles, Sean Kross, and Michelle L Mazurek. Where is the digital divide? a survey of security, privacy, and socioeconomics. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 931–936, 2017.

[77] Adi Robertson. OnlyFans CEO on why it banned adult content: 'the short answer is banks'. https://www.theverge.com/2021/8/24/22639356/onlyfans-ceo-tim-stokely-sexually-explicit-content-ban-banks, August 2021. (Accessed on 12/12/2023).

[78] Valeria Rubattu, Alicja Perdion, and Belinda Brooks-Gordon. 'Cam Girls and Adult Performers Are Enjoying a Boom in Business': The Reportage on the Pandemic Impact on Virtual Sex Work. *Social Sciences*, 12(2):62, 2023.

[79] Jorunn Birgitte Sætre. Performing sexuality, monetizing intimacy and branding the self–an analysis of onlyfans content creators and their interactions with subscribers. Master's thesis, 2023.

[80] Nithya Sambasivan, Amna Batool, Nova Ahmed, Tara Matthews, Kurt Thomas, Laura Sanely Gaytán-Lugo, David Nemer, Elie Bursztein, Elizabeth Churchill, and Sunny Consolvo. "They Don't Leave Us Alone Anywhere We Go" Gender and Digital Abuse in South Asia. In *proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2019.

[81] Patrawat Samermit, Anna Turner, Patrick Gage Kelley, Tara Matthews, Vanessia Wu, Sunny Consolvo, and Kurt Thomas. "Millions of people are watching you": Understanding the digital-safety needs and practices of creators. In *32nd USENIX Security Symposium (USENIX Security 23)*, pages 5629–5645, 2023.

[82] Teela Sanders, Jane Scoular, Rosie Campbell, Jane Pitcher, and Stewart Cunningham. *Internet sex work: Beyond the gaze*. Springer, 2018.

[83] Morgan Klaus Scheuerman, Jialun Aaron Jiang, Casey Fiesler, and Jed R Brubaker. A framework of severity for harmful content online. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2):1–33, 2021.

[84] Sexual Rights Database. https://sexualrightsdatabase.org/page/welcome. (Accessed on 06/05/2023).

[85] Aria Slippert. Platform Pimps: The Puppet Masters of the Porn-opticon. *The iJournal: Student Journal of the Faculty of Information*, 7(2):43–46, 2022.

[86] Zahra Stardust, Danielle Blunt, Gabriella Garcia, Lorelei Lee, Kate D'Adamo, and Rachel Kuo. High Risk Hustling: Payment Processors Sexual Proxies and Discrimination by Design. *City University of New York Law Review*, 26(1):57–138, 2023.

[87] Angelika Strohmayer, Jenn Clamen, and Mary Laing. Technologies for social justice: Lessons from sex workers on the front lines. In *Proceedings of the 2019 CHI conference on human factors in computing systems*, pages 1–14, 2019.

[88] Angelika Strohmayer, Janis Lena Meissner, Alexander Wilson, Sarah Charlton, and Laura McIntyre. " we come together as one... and hope for solidarity to live on" on designing technologies for activism and the commemoration of lost lives. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference*, pages 87–100, 2020.

[89] Emmanuel Syrmoudis, Stefan Mager, Sophie Kuebler-Wachendorff, Paul Pizzinini, Jens Grossklags, and Johann Kranz. Data portability between Online services: an empirical analysis on the effectiveness of GDPR Art. 20. *Proceedings on Privacy Enhancing Technologies*, 2021(3):351–372, 2021.

[90] Kurt Thomas, Devdatta Akhawe, Michael Bailey, Dan Boneh, Elie Bursztein, Sunny Consolvo, Nicola Dell, Zakir Durumeric, Patrick Gage Kelley, Deepak Kumar, et al. SoK: Hate, Harassment, and the Changing Landscape of Online Abuse. In *2021 IEEE Symposium on Security and Privacy (SP)*, pages 247–267. IEEE, 2021.

[91] Kurt Thomas, Patrick Gage Kelley, Sunny Consolvo, Patrawat Samermit, and Elie Bursztein. "It's common and a part of being a content creator": Understanding How Creators Experience and Cope with Hate and Harassment Online. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pages 1–15, 2022.

[92] Katrin Tiidenberg. Sex, power and platform governance. *Porn Studies*, 8(4):381–393, 2021.

[93] Maria Törhönen, Max Sjöblom, Lobna Hassan, and Juho Hamari. Fame and fortune, or just fun? A study on why people create content on video platforms. *Internet Research*, 30(1):165–190, 2019.

[94] Natasha Tusikov. Censoring sex: Payment platforms' regulation of sexual expression. In *Media and law: Between free speech and censorship*, volume 26, pages 63–79. Emerald Publishing Limited, 2021.

[95] Jirassaya Uttarapong, Ross Bonifacio, Rae Jereza, and Donghee Yvette Wohn. Social support in digital patronage: OnlyFans adult content creators as an online community. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*, pages 1–7, 2022.

[96] Jirassaya Uttarapong, Jie Cai, and Donghee Yvette Wohn. Harassment experiences of women and LGBTQ live streamers and how they handled negativity. In *ACM international conference on interactive media experiences*, pages 7–19, 2021.

[97] Jan Pieter Van Oudenhoven, Boele de Raad, Francoise Askevis-Leherpeux, Pawel Boski, Geir Scott Brunborg, Carmen Carmona, Dick Barelds, Charles T Hill, Boris Mlačić, Frosso Motti, et al. Terms of abuse as expression and reinforcement of cultures. *International Journal of Intercultural Relations*, 32(2):174–185, 2008.

[98] Sviatoslav Voloshynovskiy, Farzad Farhadzadeh, Oleksiy Koval, and Taras Holotyak. Active content fingerprinting: a marriage of digital watermarking and content fingerprinting. In *2012 IEEE International Workshop on Information Forensics and Security (WIFS)*, pages 175–180. IEEE, 2012.

[99] Noel Warford, Tara Matthews, Kaitlyn Yang, Omer Akgul, Sunny Consolvo, Patrick Gage Kelley, Nathan Malkin, Michelle L Mazurek, Manya Sleeper, and Kurt Thomas. Sok: A framework for unifying at-risk user research. In *2022 IEEE Symposium on Security and Privacy (SP)*, pages 2344–2360. IEEE, 2022.

[100] Miranda Wei, Sunny Consolvo, Patrick Gage Kelley, Tadayoshi Kohno, Franziska Roesner, and Kurt Thomas. "There's so much responsibility on users right now:" Expert Advice for Staying Safer From Hate and Harassment. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pages 1–17, 2023.

[101] Colin Whittaker, Brian Ryner, and Marria Nazif. Large-scale automatic classification of phishing pages. *Google Research*, 2010.

[102] Monica T Whitty and Tom Buchanan. The online romance scam: A serious cybercrime. *CyberPsychology, Behavior, and Social Networking*, 15(3):181–183, 2012.

[103] Guang Xiang, Jason Hong, Carolyn P Rose, and Lorrie Cranor. Cantina+ a feature-rich machine learning framework for detecting phishing web sites. *ACM Transactions on Information and System Security (TISSEC)*, 14(2):1–28, 2011.

[104] Yue Zhang, Jason I Hong, and Lorrie F Cranor. Cantina: a content-based approach to detecting phishing web sites. In *Proceedings of the 16th international conference on World Wide Web*, pages 639–648, 2007.

[105] Matthew A Zook. Underground globalization: Mapping the space of flows of the internet adult industry. *Environment and Planning A*, 35(7):1261–1286, 2003.

Appendix

A Study Materials:

The interview protocol for this study and the relevant recruitment materials can be accessed from https://osf.io/zbh7c/?view_only=657f69cccc17447f9be89cacaadd73f0.