
MULTI-STAKEHOLDER
DIALOGUE ON

**INTERNET PLATFORMS
SEXUAL CONTENT AND
CHILD PROTECTION**

 at
RightsCon

BACKGROUND PAPER
RIGHTSCON EDITION



TABLE OF CONTENTS



INTRODUCTION	1
BACKGROUND INFORMATION	5
About child sexual exploitation material	
About pedophilia and child sexual abuse	
About sex offense registries	
About the adult entertainment industry and payment processing	
About the law	
About the sex industry	
About child sexual abuse prevention and stigma	
CASE STUDIES	16
Lolicon, shotacon, and furry fandom	
Nudist websites that cross the line	
MAPs and registered citizens on social media	
Child modeling	
DD/Ig, ageplay, and cosplay	
BEST PRACTICE PRINCIPLES	27
BIOGRAPHIES OF SPEAKERS	34

INTRODUCTION

MULTI-STAKEHOLDER
DIALOGUE ON

INTERNET PLATFORMS
SEXUAL CONTENT AND
CHILD PROTECTION



The responsibility of Internet platforms when it comes to child protection is often understood to be about removing child abuse exploitation material (CSEM), and suspending users who engage in the sexual grooming of minors. But this is a far too narrow conception of their responsibility, and misconceiving it in this way is one of the reasons why those problems are so intractable.

There are good reasons why Internet platforms ought to see their responsibility in broader terms, because protecting minors from sexual harm also means equipping them to be sexually healthy adults. This may seem like a tall order, but it needn't be. It actually may require less of Internet platforms than the fight against CSEM and grooming—and over the longer term it will make that fight increasingly easy.

Helping children to grow into sexually healthy adults is an integral part of what we call *primary prevention*. So, why aren't other child protection organizations talking about this, and why aren't governments demanding it of Internet platforms? Because it's both not as easily understood, while also being more easily *misunderstood*.

Last month we brought Internet companies, experts, and stakeholder representatives to San Francisco to talk about this. Yes, we talked about removal of content. But just as importantly, we talked about when not to remove content. Because sometimes deciding not to remove content does more to protect minors from sexual harm than removing it.

Make no mistake, the censorship of CSEM and the elimination of online grooming is essential. But too often companies cave to demands that they go further than this, and default to censoring anything that combines the topics of "children" and "sex." This has prevented minors from accessing accurate information about sex, from exploring their emerging sexuality in age-appropriate ways, and from accessing support if they find themselves drawn towards inappropriate behaviors.

A common justification for such broader censorship is that unless we sufficiently suppress child sexuality and how it is represented in the media, we are "sexualizing" children, and creating the conditions under which child sexual abuse can become "normalized." But this runs counter to what we know about how about both child development and adult sexuality. No matter how much we censor, there will always be representations of children that some people will be able to view in a sexual light, and there will always be children beginning to explore and express their own emerging sexual identities online.





Rather than these facts being a justification for ever-increasing censorship, by viewing them through a primary prevention lens, we discover a strong argument *against* censorship. In the face of the inevitable existence of a broad range of lawful sexual content online, we need to be able to contextualize it with the very clear message that children cannot meaningfully consent to sex with adults, and that the distribution of illegal sexual images of children is wrong.

You'll read more about this challenge in this background paper for our follow-up meeting at RightsCon. Before we can engage in dialogue about what it means for Internet platforms to support a primary prevention agenda, we need to establish a shared understanding of some background facts that will ensure that that dialogue is well-informed and inclusive. Since we don't have time for extensive presentations in our workshop, that's what this background paper is for.

This is an updated version of the background paper that we prepared for our San Francisco meeting, with some updates to the presentation summaries, along with new summaries of group discussions on five case studies that raise particularly difficult questions for content moderators (see pages 16 to 26). You won't find names of the discussants, because the meeting was held under the Chatham House rule—as our RightsCon meeting is also.

Also brand new in this revision of the background paper (at page 33) are our draft *Best Practice Principles for Sexual Content Moderation and Child Protection*, which are based on discussions at the San Francisco meeting, and have been developed online by a working group that included volunteers from that meeting, along with some of the panelists that you'll be hearing from today.

We'll be presenting these draft principles for the first time here at RightsCon. In a nutshell, they represent a call for Internet companies to take a different approach towards sexual content and child protection: an evidence-based, human rights focused, and sex positive approach, that prioritizes primary prevention over censorship.

We look forward to discussing these principles with you at our workshop in Tunisia, and we invite you to collaborate with us as we continue this work following the meeting, leading up to their finalization and launch this November at the ATSA Conference in Atlanta and at the Internet Governance Forum in Berlin.



BACKGROUND INFORMATION



These summaries have been prepared by Prostasia Foundation to summarize some of the key points that our presenters imparted in the morning session of our San Francisco meeting. They do not capture all of the points that were delivered in those presentations, and may contain errors or omissions. If so, these are the responsibility of Prostasia Foundation, and not of the presenters.

ABOUT CHILD SEXUAL EXPLOITATION MATERIAL

Based on presentations from Andrew Puddephatt and Jeremy Malcolm

In 2018 analysts at the Internet Watch Foundation (IWF) found over 100,000 URLs of children being sexually abused or exploited. This is the highest number ever recorded; a 25% increase over the figures from 2017, based on a combination of public reports and proactive research by IWF analysts. This increase partly reflects the success of the IWF in locating this content so that it can be removed or blocked, but also reflecting the daunting scale of the problem.

The images that are found to be potentially illegal by the IWF are classed into three categories, based on UK sentencing guidelines. In 2018 23% were placed within Category A which shows sexual activity between adults and minors, 21% in Category B which shows non-penetrative sexual activity involving minors, and 56% in Category C which covers all other indecent images. Contained within Category C are some images that may not be illegal under U.S. law, and some IWF members choose not to include this category in their automated blocking workflow.

Not all of the images reported to the IWF are ultimately identified as falling into any of the three illegal categories. In 2018, the proportion of reports from the public that were ultimately considered to correctly identify illegality under UK law was 45%. The remaining 55% were legal content such as adult pornography, non-obscene images of minors, or other lawful content. No action is taken to restrict content that isn't identified as potentially illegal.





Once new potentially illegal images are identified, they are added to a hash list enabling them to be automatically filtered out by Internet platforms who are the IWF's members. In the United States, NCMEC provides a similar hash list to Internet platforms, through its CyberTipline reporting system. Due to the sensitivity of the content they contain, neither list is publicly available, and websites added to the list are not automatically notified of this (except through subsequent action that may be taken by the police).

Due to the cascading effect of the IWF's determinations, which are not reviewed by a court, it is important to ensure that content is not wrongly flagged as being illegal, in order to preserve public confidence that the process is not being misused to censor lawful speech. For example, in 2008, a webpage from Wikipedia was added to the IWF blacklist because it depicted child nudity on a famous album cover; a decision that was subsequently reversed.

This risk of over-blocking was highlighted in a human rights audit of the IWF conducted in 2013 by the UK's former public prosecution director Lord Ken Macdonald. In response to his recommendations, procedures were reformed to improve the IWF's accountability and to ensure that fewer lawful images were wrongly removed or blocked. For example, IWF members are now encouraged to use a splash page that identifies why a page has been blocked, and allows an appeal to be lodged. The IWF also appointed human rights and digital policy expert Andrew Puddephatt as its Chair in 2017.

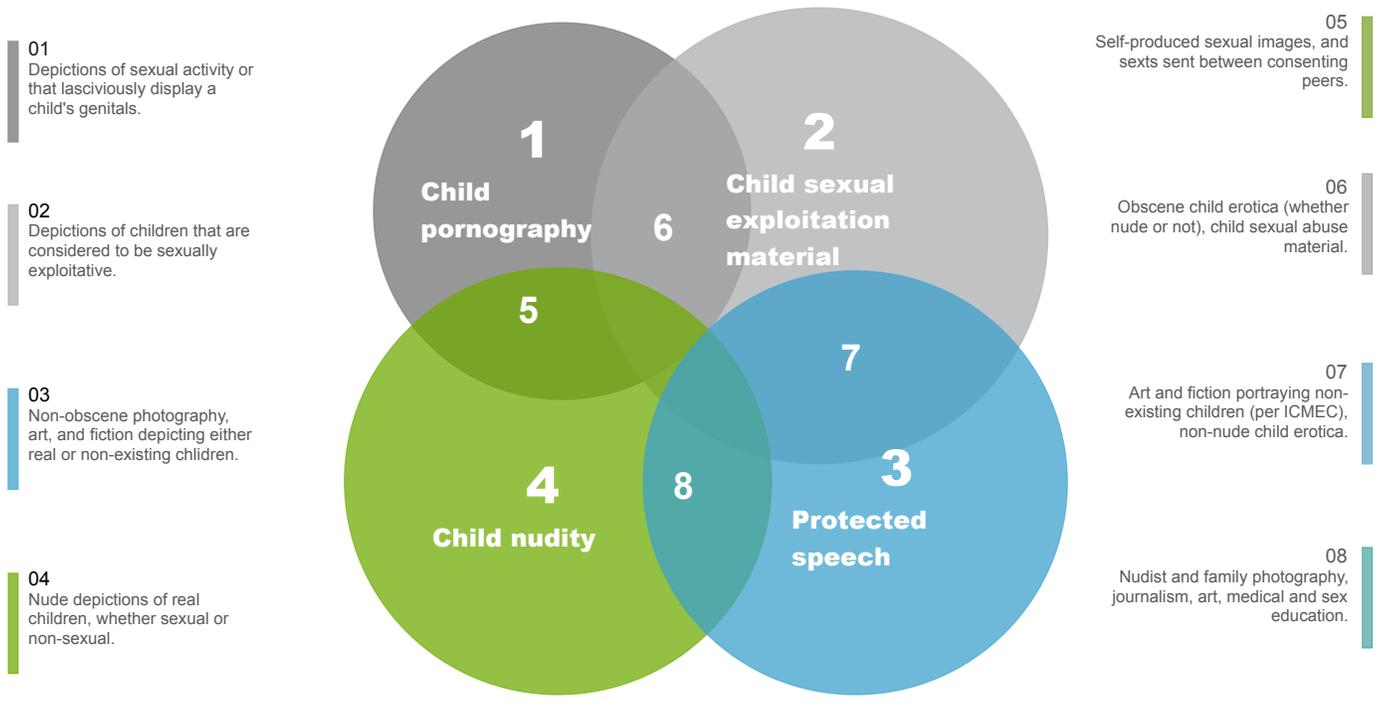
On private platforms, the problem of lack of transparency and accountability tends to be compounded. For example, another blacklist on which Wikipedia has found itself is the Google search index, which fails to return a search result to users searching for the topic "lolicon," referring to the Japanese manga art style (see page 17 below).

Unlike the IWF, there is no public process for the review of private platform policies, although large platforms do informally consult with stakeholder and experts. Facebook is currently developing a more methodical and open process for the review of content moderation decisions, in which ProStasia Foundation is participating. Smaller platforms, however, do not have the resources for accessing expert and stakeholder input—which is one reason why this meeting has been convened.

A key consideration for smaller platforms to understand is captured in the diagram overleaf. Child pornography as defined in U.S. law is a category of content that is not constitutionally protected. Hence in the representation, below, this legal category of content is shown as adjacent to protected speech, but not overlapping with it.

Child sexual exploitation material is shown as a broader category of content that overlaps with the legal definition of child pornography—but only in part, because there is some content that is considered to be exploitative but which doesn't meet the legal definition of child pornography, and is treated as protected speech. For example, when innocent family photographs are shared by strangers, the context in which this occurs may amount to exploitation of the children depicted, even though the images themselves are benign. Platforms are justified in taking action to restrict the availability of such content, even





though it amounts to protected speech. On the other hand there is no such justification for the restriction of protected speech that does not directly harm real children, such as art and fiction—despite the inclusion of such content in the definition of CSEM put forward by some other child protection groups, and included in some other national definitions of child pornography.

Finally the category of child nudity overlaps with all of the other categories: in some cases and contexts it is exploitative, and in others is isn't; likewise, it may fall within the legal definition of child pornography, or it might be protected speech. For Internet companies, breaking down these categories can help them to determine what classes of content they want to allow on their platforms. The only blanket rule should be to prevent their platforms being used to disseminate illegal content; beyond that, a balanced and contextual approach to the removal of lawful content is recommended.

In the long term, we can't continue to just censor and arrest our way out of the problem of image-based child sexual exploitation. This requires a serious and long-term investment in prevention as well as co-operation and resources from government, the charitable sector and the industry itself. As a society we always tend to favor intervention once a crime has been committed (or the illness incurred) rather than in the less glamorous and longer process of prevention. But until we take this approach, we will always be fighting fires.



ABOUT PEDOPHILIA AND CHILD SEXUAL ABUSE

Based on a presentation from Gilian Tenbergen

Child sexual abuse is a huge social problem. Although estimates of the prevalence of child sexual abuse are difficult to generate, research indicates that roughly 1 in 5 girls (20%) and 1 in 20 boys (5%) will be victims of sexual abuse in their childhood (see the work of David Finkelhor and the Crimes Against Children Research Center). In 2016 alone, there were 57,329 reported victims of child sexual abuse in the United States (U.S. Department of Health & Human Services, 2018).

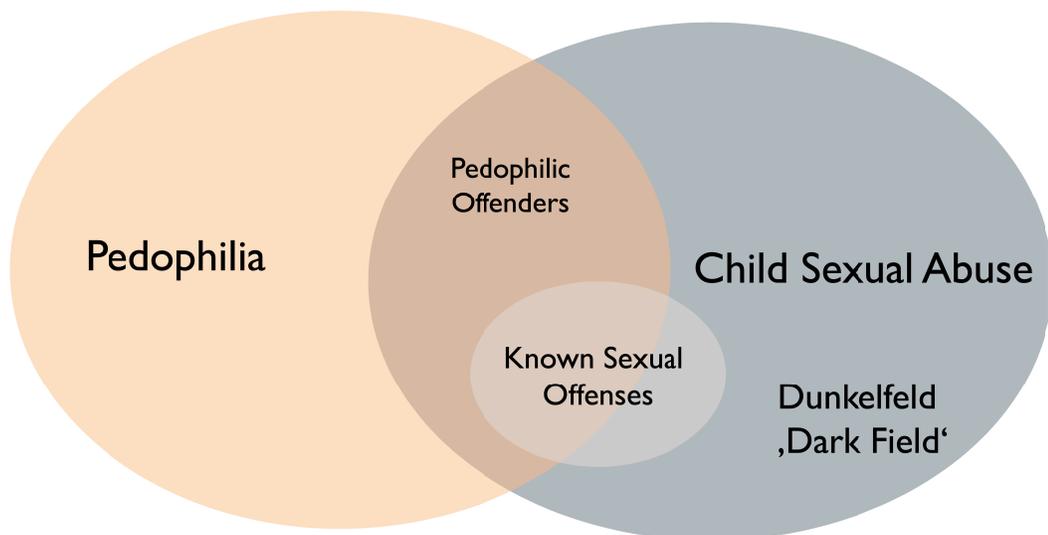
Sexual Preference Disorder

Pedophilia/Hebephilia

exclusive & non-exclusive type

Sexual Behavior Disorder

Child Sexual Abuse



Before meaningful work can start, one must differentiate between pedophilia and child sexual abuse (CSA). These are not the same, although there is some overlap. Pedophilia is the sexual preference for prepubertal children as manifested through persistent and recurrent (distressing) sexual thoughts and fantasies, urges, arousal, /or behavior (APA; DSM-5). 40-50% of sexual offenders against children are pedophilic, the remaining are "replacement offenders" whose sexual preference is for adults, but who may have easier access to a child victim (Seto, 2008; Schaefer et al., 2010).

The types of online CSA behaviors include the consumption and dissemination of child pornography (Child Sexually Exploitative Materials, or CSEM), the sexual solicitation of minors (sometimes enacted through online fantasy chat, but in other cases as a preparatory step towards real world contact) , sex trafficking, and exhibitionism and voyeurism.





Technology enables these offending behaviors through video services (e.g. YouTube), messenger services (e.g. WhatsApp, Threema, Telegram), closed community communication services (Slack, Discord), chat services and message boards (IRC, Reddit), online communication services (Skype, Hangouts), and Dark Web sites.

The perpetrators of online offending are diverse. They include individuals with pedophilia/hebephilia (i.e. Minor-Attracted Persons), but as indicated earlier, they also include many adult-attracted individuals. Typical online offenders are younger than offline offenders, with higher income and education levels. They are also typically male, and tend to score lower in measures of deviancy and relationship difficulties as compared to contact sexual offenders (Seto, 2018; Babchishin, Hanson, VanZuylen, 2015).

Prevention is key when working with this group. CSEM cases are increasing, making it relevant to target this group. However, the law FOSTA (Allow States and Victims to Fight Online Sex Trafficking Act) has made targeting them difficult through forced censoring of *all* CSEM/pedophilia-type information. This includes targets like prevention groups attempting to provide scientific and clinical information and scientists attempting to conduct research. As a result, we are making it *more* difficult for these individuals to seek help, thereby increasing likelihoods to offend.

Internet companies can help by understanding the science of pedophilia and sexual offending against children, interacting with experts and using their knowledge to guide software development/safety protocols/AI use, and implementing the knowledge gained

ABOUT SEX OFFENSE REGISTRIES

Based on a presentation from Guy Hamilton-Smith

The goal of creating safe communities is a valuable one, but well-intentioned efforts often undermine those goals. Many tech platforms, including Facebook and Instagram, have policies that ban anyone convicted of a sex offense from participation. Although superficially attractive, these blanket policies are problematic for a number of reasons, all of which ultimately undermine safety.

First, almost all (>95%) sex offenses, online or otherwise, are committed by people unknown to authorities. Despite the popular perception that those who commit sexual offenses tend to have a string of victims, this is not usually the case. Rates of re-offending by those who have sexually offended are in fact among the lowest for any category of crime. Therefore, to use a sex offense registry as a proxy for ensuring child safety creates a false sense of security.

In February 2019 Michigan Attorney-General Dana Nessel filed a brief in the state Supreme Court case of *State of Michigan v. Snyder* write about the Sex Offender Registration Act (SORA) as follows:





Modern social science research has shown that SORA's extensive burdens are excessive in relation to SORA's purported public safety goals. There are two salient points: 1) research refutes common assumptions about recidivism rates that supposedly justify SORA's extreme burdens; and 2) regardless of what one believes about recidivism rates, registries are not good tools to protect the public.

A further shortcoming of these policies is that they assumes that someone who seeks to take to an online platform to commit a crime will sign up with their real information. Thus, this policy will only affect people who have been held accountable for a crime, and are now trying to lead law-abiding and productive lives.

Social media and technology platforms have a crucial, necessary role to play in helping people reintegrate into their communities. To the extent that people remain isolated and disconnected from their community, they are more – not less – likely to commit more crimes.

In *Packingham v. North Carolina* the Supreme Court ruled it unconstitutional for the state to prohibit a person on a sex offense registry from accessing a social networking website. Justice Kennedy wrote:

Even convicted criminals – and in some instances especially convicted criminals – might receive legitimate benefits from these means for access to the world of ideas, in particular if they seek to reform and to pursue lawful and rewarding lives.

Importantly however, nothing here prevents platforms from policing the behavior of individuals on their platforms. People can still be suspended or banned from platforms on the basis of things that they do now, but should not be banned on the basis of something that they did however many years in the past.

To the extent that people exiting the criminal legal system succeed, everyone succeeds. These bans only impact people trying to lead law-abiding lives, and incentivize deception.

ABOUT THE ADULT ENTERTAINMENT INDUSTRY AND PAYMENT PROCESSING

Based on presentations from Ian O'Brien and Cathy Beardsley

As the national trade association to the adult entertainment and pleasure products industry, the Free Speech Coalition's mission is to lead, protect, and support the growth and well-being of businesses and workers in the adult industry, as well as the communities to which they belong.

FSC's vision is of a world in which the international adult industry, its workers and businesses, have equal rights, protections, and freedoms under the law; and where all people are protected from exploitation, and empowered by age-appropriate sexual health education.





Commonly referred to simply as the “adult industry,” its name alone explicitly identifies the industry by its target market: it is for adults only. In everything that the industry does, care is taken to ensure that the products and services provided are not marketed towards minors.

People in the sex industry are people with families. We are an enormous, historically insular market, and we exist in a changing landscape that is rapidly decentralizing. Although we have solidarity with sex workers whose work is criminalized, we are also largely a legal industry; most types of sex work are lawful and indeed tightly regulated.

The FSC’s advocacy work has largely been about the existential question of whether porn or sexual content should be allowed to exist at all. Very often battles over the legality of adult content have been couched in terms of child protection. For example, the Child Protection and Obscenity Enforcement Act of 1988 places stringent record-keeping requirements on producers of adult films, but also exposes young actors to the risk of having their personal information misused. The FSC has successfully sought to narrow the scope of this law so that records of personal information no longer have to be maintained by every individual website on which an adult film appears.

There has also been an effort by those opposed to the adult industry to have pornography framed as a public health crisis. Resolutions declaring it so have been passed in more than a dozen states. Many of the resolutions are based on a model text written by the anti-porn group National Center on Sexual Exploitation, previously known as Morality in Media. But public health experts say that while there are questions about the effects of pornography use by minors, there is a much more pressing public health justification for the introduction of comprehensive sex education throughout the United States—a measure that the same legislators oppose.

These resolutions have no binding legal effect, but they do reinforce the perception of adult content as being especially abnormal and risky, and this carries over to the way that adult content producers are paid for their content, which involves extraordinary costs and burdens that do not apply to other content producers. Segpay is a global leader in online payments that specializes in serving these producers. It provides technology services that can ensure that a high level of protection for children is maintained, while also minimizing fraud and other illegal practices.

The rules that Segpay and other billers enforce to manage risk come not only from a variety of sources including the federal government, the Federal Trade Commission, the EU Commission, the UK Parliament, card brands, and banks. They include US Federal Regulations such as 18 U.S.C 2257 and FOSTA, privacy regulations such as the EU GDPR and the California Privacy Policy Act, card brand’s Business Risk Assessment and Mitigation (BRAM) policies, age verification rules such as the UK Digital Economy Act, and anti money laundering rules in the US and EU.





Segpay's role is to act as a "content cop" for its customers, ensuring that all site models are of age through a thorough review of websites, ensuring that all adult merchants comply with the 2257 record-keeping requirements, and properly vetting social network and dating sites. Marketing with content that appears underage is also prohibited. We do not process for escort sites, which are a higher risk as they could potentially have ties to sex trafficking.

Segpay also requires its customers that offer adult content to affirm that they will not permit persons who are not of legal age to view or access that content, and requires an 18+ warning page. Customers marketing to the UK are required to comply with new age verification legislation, to further ensure that their materials are not viewed by minors or unwilling adults. We conduct ongoing monitoring to ensure compliance with these standards.

ABOUT THE LAW

Based on a presentation from Cathy Gellis.

The primary law affecting these discussions is the federal law at 47 U.S.C. Section 230 (aka "Section 230"). A few things happened in the mid-1990s that led to its passage. One was the case *Stratton Oakmont v. Prodigy*. Prodigy, like America Online and CompuServe, as well as a lot of smaller "bulletin board" services, was an early online service provider. As an online service provider one of the things it let its users do was post messages that were readable by other users. The problem that came up in this case was that someone tried to hold the service provider the messages were posted on liable for them - and a New York State court let them.

This decision put service providers in an untenable position. If they could be held liable for the things their users said - in all the myriad ways what users said could potentially be legally wrongful - then they could not afford to let people use their systems to express themselves. They would have to try to monitor all the voluminous content that users post on their system, which is an effectively impossible task, which would mean they would inevitably end up overly censoring user expression that might be perfectly lawful. Assuming they could take the risk of being available to users at all.

Meanwhile, in the mid-1990s Congress was also worried about the possible proliferation of pornography on the Internet. And so it passed the Communications Decency Act of 1996 as a way to address this perceived policy problem. Much of that law was challenged, however, in a case called *Reno v. ACLU*. In this case the US Supreme Court struck down the bulk of the law for violating the First Amendment. But one little piece of the law survived: Section 230, which was the part of the law crafted to address the problem caused by the *Stratton Oakmont* case.

In Section 230 Congress is candid about what it is trying to do: get the most good content online, and the least bad. And the way it tries to accomplish both is by turning service providers into partners to help it achieve both these policy goals. In other words,





instead of threatening service providers with punishment in case they let the wrong content stay up or forced the wrong content to come down, it provided them with the immunity they needed to make their best efforts to get both right.

Section 230 makes clear that service providers cannot be held liable for whatever may be legally wrongful in the content supplied by their users. Their *users* certainly can be held accountable, but not the platforms themselves, which is what the court in *Stratton Oakmont* tried to do. Also, because Congress wanted platforms to make their best efforts in getting rid of the worst user content, it additionally immunized them from any liability that might arise from doing what moderation they could.

As a result of this law, the Internet has been able to flourish over the last twenty years or so, along with a plethora of valuable online user expression. Section 230 has, however, always had some limitations. One of them is that Section 230 does not insulate platforms for liability arising from user expression that may violate intellectual property rights. As a result, we've seen more undue censorship around user expression where infringement claims have been able to be made, which stands as a cautionary tale about what happens to speech when service provider protection is weakened.

Section 230 also has never provided any protection to service providers when the thing wrong with the user expression is that it violates federal criminal law. Service providers have always had to proactively ensure that user content does not violate these serious crimes, such as child sexual exploitation. But until the passage of FOSTA, it was a much more straightforward task. FOSTA has now made provider moderation a much more complicated enterprise due to the complicated amendments FOSTA made to that federal criminal law, as well as its direct weakening of Section 230. And the effect has been a complete undermining of everything Congress sought to advance when it originally passed the law. Not only have these changes now caused service providers to remove lawful content, but it has also even effectively *discouraged* them from moderating the expression they do still facilitate, even though it would be better for everyone if they still could help on that front.

Some of those who have been affected by these changes have been challenging the constitutionality of FOSTA in a case called *Woodhull Freedom Foundation v. US*, which continues to pend before the courts. But unless and until FOSTA is deemed unenforceable it will continue to shape the expression that service providers will be able to allow online.



ABOUT THE SEX INDUSTRY

Based on a presentation from Kristin D'Angelo

Sex work is a broad term that encompasses a wide variety of erotic labor, both legal and illegal, including exotic dancing, escorting, phone sex, and porn. Sex trafficking is a narrower term that generally connotes lack of free will or coercion, as it is a term rooted in labor rights and empowerment/sexual freedom (coined by Carol Leigh, aka. Scarlet Harlot).

Human trafficking is a legal term that has distinct state and federal definitions. The Trafficking Victims Protection Act (TVPA) defines human trafficking as engaging in prostitution due to force, fraud, or coercion or being induced into prostitution under the age of 18. It does not need to involve transporting persons across state lines.

Survival sex is prostitution engaged in by a person because of their extreme need. It describes the practice of people who are homeless or otherwise disadvantaged in society, trading sex for food, money, a place to sleep, or other basic needs, including drugs. The criminalization of sex work makes survival sex workers even more vulnerable: there are no services available to them, they are often homeless (80% based on SWOP Sacramento Research), 59% reported being raped, 55% reported beatings, and 27% reported being harmed by an officer.

Censorship of online content harms this large survival sex population. 18% identified migrating after the seizure of the SF Redbook website in 2014. Websites which allowed transition from street work into online sex work are life saving. Criminalization will never stop a person from surviving. It is an unreasonable idea that a person will stop eating if they are arrested.

The premise behind SESTA-FOSTA is a false one: sex work is not inseparable from sex trafficking, and only a tiny percentage of cases of child sexual abuse involve child sex trafficking. Conflating these separate problems is harmful. Laws such as this that are passed in the grip of moral hysteria can literally kill, and minorities such as survival sex workers are the most vulnerable of all.

The Internet intersects with sex work, but is not correlated with the greater supply of sex workers. Rather, it is a support service, that has helped independent sex workers to stay safe. Use of the Craigslist erotic section was associated with a 17% drop in female homicide (Sept 2017, West Virginia University and Baylor University). SF Redbook and Backpage were also a safer way for sex workers to transact business. Today, these have all disappeared, and when services are pulled, sex workers suffer. Support services do not hurt trafficking victims or sex workers; on the contrary, safety benefits us all.

While they have no choice to comply with their legal obligations until SESTA-FOSTA is repealed or annulled, Internet platforms can still defend sex workers from further harm by listening to them, their experiences, and their history, and allowing them to exercise their right of free speech.



ABOUT CHILD SEXUAL ABUSE PREVENTION AND STIGMA

Based on a presentation from Candice Christiansen

One of the biggest problems faced by those who wish to take a fact-based approach to child sexual abuse prevention is that this involves treating those who are at risk of offending as human beings. This in turn results in our society's hatred towards sexual abusers being turned on those who are only trying to help prevent abuse. During 2018, the Global Prevention Project was the target of one such sustained and vicious attack.

Candice Christiansen, founder of the Global Prevention Project, is herself a survivor of child sexual abuse and a licensed clinician, and perhaps the most controversial aspect of her work is to give a voice to non-offending minor-attracted persons (MAPs), as a step towards making it easier for this population to access support and wellness resources and information, with the objective that this will help them avoid offending.

The most important message is that being an anti-contact, non-offending MAP is a choice for those who find themselves attracted to children. Such individuals do exist and they are men, women, non-binary and transgendered individuals with pedophilia, hebephilia, and ephebophilia. There is a broad scientific consensus that these groups' attraction towards children is not a choice—but their behavior in response to that attraction is a choice.

Non-offending, anti-contact MAPs exist throughout society and on social media, whether they are permitted to say so or not. But the stigma that affects clinicians who make a principled choice to support a prevention agenda also affects Internet companies. This has resulted in some platforms making the choice to take prevention resources offline.

There is debate among experts over whether erotic stories and pictures such as cartoons, objects such as sex dolls, and practices such as role play, are helpful or harmful to MAPs. Little to no research exists in this area. For exclusive MAPs, these may be their only way of expressing their sexuality. For those MAPs who endorse addictive symptoms, they report it can be triggering.

Online stories, blogs, podcasts, and posts of MAPs are a simpler case. These are necessary to provide other MAPs support, as there is strong evidence that isolation creates risk. These resources also provide the family members of MAPs with information, educate clinicians on what type of support NOMAPs need, and provide researchers with access to this "underground" community for research.



CASE STUDIES

MULTI-STAKEHOLDER
DIALOGUE ON

INTERNET PLATFORMS
SEXUAL CONTENT AND
CHILD PROTECTION

at
RightsCon

These case studies were presented and discussed during the afternoon session at our meeting in San Francisco, to stimulate thought about how the information presented by our experts during the morning session could be applied to some hard cases—grey areas in which different approaches could lead to different decisions being made by platforms conducting content moderation, implementing automated filtering algorithms, constructing content blacklists, or developing content policies.

Participants were not asked to come up with a single “right” answer about how these types of content should be treated in every context; in many cases the answer to whether such content is acceptable or not might be “it depends.” But that’s not a very useful answer in itself, so we encouraged them to dig a bit deeper. Here are some of the general questions that participants considered in their small group deliberations:

- How would restricting this content protect children from sexual harm?
- Who may be harmed by the restriction of this content, and how?
- If it is proposed to restrict access to this content:
 - In what contexts should it be restricted?
 - Globally
 - Only in particular contexts (which ones and why?)
 - By whom would it be restricted?
 - A social media platform
 - A web host
 - A search engine
 - A payment or advertising intermediary
- If it is allowed, should steps be taken to prevent misuse of the content? If so, what?
- What precedent might the restriction of this content set for other types of content?



LOLICON, SHOTACON, AND FURRY FANDOM

Lolicon and shotacon and furry art are cartoon fantasy art forms with roots in the Japanese manga style of comic book art. Over the past year, platforms such as Discord, Twitter, and Reddit have all placed new restrictions on such art, preventing users from uploading images that appear to represent children—or in the case of furry cub art, childlike animals. For example, Reddit’s current policy states:

Reddit prohibits any sexual or suggestive content involving minors or someone who appears to be a minor. This includes child sexual abuse imagery, child pornography, and any other content, including fantasy content (e.g. stories, “loli”/anime cartoons), that depicts encourages or promotes pedophilia, child sexual exploitation, or otherwise sexualizes minors or someone who appears to be a minor. Depending on the context, this can in some cases include depictions of minors that are fully clothed and not engaged in overtly sexual acts.



In March 2019, ProStasia Foundation submitted comments to the United Nations Committee on the Rights of the Child, which was considering making a recommendation that depictions of non-existing minors in art or fiction should be criminalized. In our submission we pointed out (references omitted, but available in the original):

We agree that there is insufficient research on this question, and we plan to raise funds for more such research. But what research we do have shows exactly the opposite of what ECPAT claims: that access to representations of non-existing children is not associated with greater social acceptability of sexual interaction with children, and that it may actually decrease rates of actual sexual offending against children. This may be because virtual representations such as cartoons and dolls can provide a safe, victimless outlet for some people who are sexually interested in children, but who abhor the idea of harming a real child.

In this case study, we considered such policies against the “sexualization” of non-existing minors in the form of cartoon art. In addition to the general questions applicable to all case studies, groups considered:

- Is the concern about “sexualization” a coherent and valid one?
- How easy is it to determine whether art represents a child, and does so sexually?
- Could there be benefits of allowing minors to be depicted sexually in art?



SUMMARY OF DISCUSSIONS

There was general agreement that the appropriate standard to be used should be “don’t harm other people” rather than “this makes me feel icky.” However, one group said they needed more information about whether sexual depictions of minors in art are a pathway to abusive behavior, an outlet to avoid such behavior, or neither. One participant mentioned research on the effects of violent video games, which had not shown any link to real-world violence.

One group felt that it was important how the images were created, and that while entirely fictional imagery based on imaginary characters should not be restricted, there was a stronger argument against allowing images for which real children have been used as models.

As another participant explained, payment processing rules further complicate the situation, not only due to bans against sexual depictions of minors, but also depictions of bestiality which could impact some furry art, unless the character looks sufficiently humanoid. Bizarrely, this also affects sexual depictions of alien characters.

Ultimately the decision on what to accept lies with private actors, and the context in which the depictions are shared will be relevant to their decision. But although it’s easy to say no, we should also consider that when extreme content is banned from major platforms, it doesn’t disappear altogether but just ends up in darker, less well scrutinized places. Consistent with their obligations under the law and to their business partners, allowing it under clearly defined terms may be the least harmful option for platforms to take.

Best practice ideas

- Industry can help fund research to clarify whether non-realistic representations of fictional minors in art or literature help to prevent child sexual abuse.
- Parties dealing with payment processors can provide them with information to make evidence-based decisions on whether to allow payments for such content.
- Platforms should clarify their terms of service to provide better transparency about what kinds of fictional content are allowed or not.



NUDIST WEBSITES THAT CROSS THE LINE

We can accept that nudism (also called naturism) is a legitimate lifestyle and that families do participate in it responsibly, without exposing their children to harm. However, nudist community organizers need to take responsibility to make sure that children are not being exploited. This applies equally on the Internet as it does in physical venues. There are a number of nudist websites that are legal, but nevertheless are arguably exploitative in their focus on photographs of children.

One such website offers paid memberships which it claims “support the families involved,” and contains the following very assertive disclaimer:

The naturist photo and video documentaries that we offer are legal in every city, every state and every county inside the United States. They are protected by the First Amendment and are not subject to local obscenity laws or ordinances.

On the other hand, the same website also warns:

Our system is specifically designed to meet all legal specifications. Downloading and saving material out of context of PureNudism.com can cause issues. This is for your protection.



Prostasia Foundation reached out formally and extensively to the nudist community, at local, national, and international levels, to invite their participation at this event, but they declined to participate. However the American Association for Nude Recreation did acknowledge the potential for confusion between “nudist and naturist web sites [and] the dating and exhibitionist/porn sites that are all over the Internet.” Since neither the AANR nor the International Naturist Federation (INF) offers any accreditation system to distinguish legitimate websites from illegitimate ones, determining the difference is essentially left as an exercise for law enforcement authorities, and for payment and Internet intermediaries.

In this case study, groups were asked to discuss the following additional questions:

- Is child nudity inherently exploitative?
- Who should make the call about a nudist website that appears to go too far?
- Should Internet intermediaries make it harder to access such (legal) content?



SUMMARY OF DISCUSSIONS

There was broad agreement that there are contexts in which nudist photography is likely to be harmless even if it depicts minors. An example may be the hosting of such images in private groups of the families depicted. It was also noted that while simple nudity is stigmatized in American culture, this is not true to the same extent in some other cultures.

Several participants stressed the importance of ensuring the informed consent by all of the subjects depicted to all of the potential uses of their image. When minors are depicted, the responsibility to consent is shared by their parents or guardians. There was a general view that once nude images of minors are stripped of their original context, the usage can become exploitative, because it negates this consent, or at least renders it less clear.

One group mentioned that privacy principles could also help drive editorial decision making on this issue. Indeed, it was noted that nudity isn't the only reason why someone depicted online might legitimately want their private images to be removed.

Another group concluded that although nudism practiced in the natural environment is not exploitative, it is difficult to avoid the exploitative use of images once they are brought online—and that the nudist community may not have adequately begun to grapple with this problem, by upgrading its privacy and security practices.

One of the suggestions made to deal with this dilemma was that there should be a more active form of moderation of images of children in nudism, to ensure that the context is appropriate and that there is ongoing informed consent. In practice, this would require robust takedown policies on all platforms where these images are allowed.

In cases where legal images are being misused, the question arose as to which participants in the Internet ecosystem should step in. One group felt that action should be taken at the hosting level, but not at a lower infrastructural level such as through the DNS system.

Best practice ideas

- Platforms should consider avoiding blanket “no nudity” policies in favor of a contextual approach based on the likelihood of harm.
- Nudist groups should collaborate with other stakeholders to develop standards to assure that images are being used consensually in a non-exploitative context.
- Platforms that host images of underage persons should have policies that allow their removal of those images if the person depicted withdraws their consent.



MAPS AND REGISTERED CITIZENS ON SOCIAL MEDIA

Several platforms disallow those who identify as minor-attracted persons (MAPs) to post their thoughts online because it offends their other users, and/or because of a belief that it “normalizes pedophilia.” Other platforms place a blanket ban on users who are on sex offense registries, regardless of what they post.

For example, during 2018, Tumblr reportedly terminated the accounts of all self-identifying MAPs as a prelude to the later revision of its policies on sexual content of all kinds. The reason given for termination of these accounts, most of which were dedicated to child sexual abuse prevention, was that they contained “inappropriate content involving minors”—even if they didn’t.

Similarly, Medium suspended a publication called “Pedophiles about Pedophilia” in August 2018, and Discord repeatedly banned a support server called MAP Support Chat, which is now hosted independently.

But not all platforms take the same approach: Twitter allows non-offending MAPs to maintain Twitter accounts, and only terminates those that promote child sexual abuse or otherwise infringe Twitter’s content rules.

It is unlikely ever to be possible for a platform to ensure that it has no MAPs or registered citizens as users. All that it can do is to terminate their accounts if they ever come out or are “outed” as such. Based on the presentations given in the morning session in San Francisco, the motivation for a blanket prohibition on MAPs or registered citizens to have a presence on social media has also been questioned.

On the other hand, we also have to acknowledge that platforms face significant social and financial pressures not to carry such controversial content. For smaller platforms, these pressures may be insurmountable and make a blanket ban seem like the most attractive option. But is it the option that best protects children?

Additional questions for discussion during this case study included:

- Do these policies make sense, or are there better approaches we could take?
- If MAP blogs are allowed, should they be limited to anti-contact non-offenders?
- How can platforms address the feelings of users who are offended by such content?



SUMMARY OF DISCUSSIONS

There was general agreement that the marginalization of minor-attracted persons and those on sex offense registries would not promote prevention, as it only highlights the “otherness” of these groups and prevents them from accessing community support and resources that could promote healthy behaviors.

Conversely, prohibiting access can increase the risk of problematic behaviors for those who are working towards rehabilitation. Cutting them off from online social spaces could prevent them from obtaining necessary resources and support and obtaining meaningful work, which would create a risk factor for reoffending.

It was also acknowledged that every platform can decide what content it carries. However de-platforming of these groups on large platforms with a wide audience was thought to be the wrong approach, indicative of what one group described as “FOSTA paranoia.” One group suggested that we should instead take a restorative approach, while empowering end users to filter out content from communities that they didn’t want to hear from.

Another group maintained that more data was needed on whether it is healthy for these groups to use online platforms as outlets for their thoughts and frustrations, but acknowledged that due to FOSTA, it has become more difficult to have an open and honest discussion on this question and to obtain the needed data.

One group tackled the difference between pro-contact vs anti-contact MAPs, suggesting that the pro-contact camp was less likely to suffer from censorship, as they could always retreat into darknet forums that the anti-contact camp would avoid. They suggested it was necessary to distinguish between those advocating for harming others or who are actively grooming, as opposed to those who are raising awareness and posting non-explicit educational content.

Best practice ideas

- Companies with large platforms should consider avoiding blanket policies to ban users on sex offense registries or those who admit to being minor-attracted.
- Industry should contribute towards the funding of further research into the positive and negative outcomes of social media use by these users.
- Policies should be based on the behavior of the users concerned, and it must not be permitted for users to incite or to engage in child sexual exploitation.



CHILD MODELING

Some of the legal child modeling content online could be considered ethically dubious. An example that has drawn some public attention is Instagram, where young models post content to accounts run by their parents. Often the content that gets the most page views, clicks and comments are the most revealing photos. This image is of a twelve year old model, taken by a professional photographer and posted to Instagram with the approval of the model's parents. Take note of the comments alongside the photo.



Instagram's community standards state:

People like to share photos or videos of their children. For safety reasons, there are times when we may remove images that show nude or partially-nude children. Even when this content is shared with good intentions, it could be used by others in unanticipated ways. You can learn more on our Tips for Parents page.

Beyond this, there are websites dedicated to child modeling, which feature underwear and swimwear shots from professional and amateur models. These often have .com domains and are indexed on major search engines. There are even websites that feature nude adult models from modeling studios that were shut down by authorities for having produced child sexual exploitation images and videos. Where should the line be drawn?

In this case study, groups were invited to consider answers to these additional questions:

- Whose responsibility is it to ensure a child model's welfare?
- Is the removal of sexual comments and solicitations a sufficient response?
- Is it ethical to link to legal content from a studio that also made illegal content?



SUMMARY OF DISCUSSIONS

There was general agreement that platforms should avoid placing excessive restrictions on the ability for children and teens to publish their own modeling photographs online, because treating their photographs as a kind of forbidden fruit would only attract those with bad intentions towards them.

Posting on social media such as Instagram is a common pathway towards becoming a successful model nowadays, and so we should facilitate this while taking a harm reduction approach. Parents and guardians should continue to have primary responsibility for their child's welfare, and children need to be made aware of the risks and benefits and to be offered an informed choice. Comprehensive sex education is important part of equipping them to make this choice.

Inappropriate sexual comments shouldn't dissuade young models from pursuing their chosen hobby or profession, and it is victim-blaming to ascribe these comments that may be posted on a model's images as being the model's fault for "sexualizing" themselves. Parents and platforms have a shared responsibility to shield young users from inappropriate comments.

One participant who had experience in the modeling industry explained that in professional modeling, agencies act as intermediaries between the model and the brand, and that all images are reviewed before publication. Where agencies don't exist for amateur models, this creates a gap where inappropriate images can slip through.

While it is the wrong approach for platforms to assume heavy editorial obligations, one group suggested that after a model's audience on a platform reaches a certain size, it could be helpful for them to have additional assistance and support to manage their account, in the same way that large YouTube personalities get access to additional platform services.

On the other hand, there was broad agreement on the impropriety of the continued use of images from modeling studios that knowingly published child sexual abuse images in the past. These studios are in a different class to those that may merely have omitted to maintain paperwork to document that their models were of legal age, and it would be appropriate for hosting companies to disallow images from these studios.

Best practice ideas

- Platforms should prohibit the posting of inappropriate sexual comments on photographs of children and make it easy to report comments for deletion.
- For larger accounts, platforms with the resources to do so should offer assistance to parents of the managers of their children's modeling portfolios.
- It is appropriate for platforms to disallow and to refuse to promote websites or images that are directly associated with child exploitation.



DD/LG, AGEPLAY, AND COSPLAY

The sexual fetishes of DD/lg (Daddy Dom/little girl) and ageplay are immensely popular, especially with younger women. Fashion, photography, and art all create an aesthetic that calls back to the idea of youth, and some of these representations are overtly or implicitly sexual.

DD/lg and ageplay fantasies often take place in private. But representations of these fantasies can also be found online when users either post photographs of themselves dressed as a younger persona (“cosplay”) on websites such as Tumblr and FetLife.

They may also take the form of text-based chats and stories posted to platforms such as Reddit and Discord, in which participants role play as being underage but give a clear disclaimer that they are actually just pretending.

In most cases, it is obvious that the ageplayer is role playing: they might be wearing a onesie and sucking a pacifier, but their build and features are clearly those of an adult. But in some cases, these depictions may be literally indistinguishable from actual representations of minors.

An extreme example is shown in the photograph shown here, which comes from a popular model, Marina Nagasawa. Believe it or not, she was 21 years of age when this photo was taken, although she is cosplaying as an elementary schoolgirl. As an adult, can we tell her that she doesn't have the right to pose as a child in cosplay photographs that she posts online?



In addition to the questions raised for consideration across all of the case studies, some additional questions presented for discussion here included:

- Should these representations of child sexuality be allowed? With what limits?
- What steps can be taken, and by whom, to prevent minors participating?
- Does it make any difference when it is depicted in sex work or commercial porn?



SUMMARY OF DISCUSSIONS

Participants found it hard to find any justification for the limitation of online representations of role play between consenting adults in an appropriate context. There was no desire to see this type of content driven underground.

One group acknowledged that there may be some who would use these depictions as a proxy for their sexual interest in children. However this group also said we shouldn't allow a few bad actors to limit the speech rights of other adults. They drew an analogy to National Geographic, which had long contained images of non-sexual nudity—this was never banned despite the fact that some people doubtless used it for sexual gratification.

Another group said that rather than banning depictions of sexual role play, the appropriate response by platforms may be to allow such content but to require that it be kept strictly separated from actual (non-sexual) depictions of minors, to limit the possibility of harm from conflating the two. Ageplay and related content should be labeled to make it clear that the depictions are of adults engaging in consensual conduct.

Participants acknowledged that there could sometimes be a logistical problem in determining the ages of the parties involved. Communities dedicated to this type of content could screen and verify that content. When such content is posted to more general purpose platforms, specialized intermediaries could play a role in helping to verify the ages of those depicted in role play activities and that the scenario depicted is consensual.

One group mentioned that in pornographic videos, intro and exit interviews are sometimes conducted, explicitly spelling out the participants' consent to the fantasies that are being engaged in. Disclaimers can also be used, but if so steps should be in place to ensure that these disclaimers are verifiable.

Best practice ideas

- Representations of adults role-playing as minors should be treated as adult content and should be separated from representations of actual minors.
- Such content should be allowed in appropriate contexts, but should be labeled appropriately and may include a disclaimer about the ages of the models.
- Platforms should make use of expert intermediaries and communities who can help to identify and distinguish between role play scenarios and actual abuse.



BEST PRACTICE PRINCIPLES



Internet content platforms (such as search engines, social networks, chat applications, and cloud storage services) are frequently the first port of call for regulators seeking to find easy solutions to the problem of online child sexual abuse. But although platforms have made a vital contribution towards this effort and will continue to do so, there are at least three limitations of the approach that regulators are pushing platforms to take.

First, it tends to promote a “one size fits all” approach that overlooks the differences between platforms in terms of their financial resources and technical expertise. As the Internet Watch Foundation (IWF) testified to the UK government last year:

There is a myth that the tech industry is a-wash with money and the brightest and the best brains, with the ability to solve all the world’s problems and whilst that may be true of some of the larger players, there is a need to recognise that much of the tech industry in the UK is made up of small start-ups that do not have access to the sorts of resources Government think they do.

Second, when platforms are pushed into over-blocking and over-censoring, this frequently results in infringements of the civil liberties of minorities such as sex workers, the LGBTQ+ community, survivors of child sexual abuse, and sex workers. For example, FOSTA was putatively aimed at making Internet platforms liable for the facilitation of sex trafficking, but has in practice also resulted in the censorship of lawful speech, including sex education resources.

Third, an approach that pushes platforms into censoring any sexual content that they instinctively regard as “questionable” does not actually protect children, and could indeed harm them. Sometimes platforms choosing not to censor content is more likely to protect children from sexual abuse. For example, in response to FOSTA, threats of regulation, and public pressure, platforms have been found censoring child sexual abuse prevention materials and forums.

More broadly, United Nations Special Rapporteur David Kaye found in his 2018 report on the promotion and protection of the right to freedom of opinion and expression that the failure of platforms to apply human rights standards in their policies related to sexual





content has resulted in the removal of resources for members of sexual minorities, and depictions of nudity with historical, cultural or educational value.

Currently, many platforms do already have child protection policies as part of their content policies or community standards, however these can be vague and unpredictable in their application even by a single platform, let alone between platforms. Smaller platforms may not have well-developed policies on this topic at all. Even in mid-size platforms, trust and safety teams are typically composed of members who deal with other forms of abusive content such as spam and fraud, but which lack dedicated expertise in child protection. Often, requests to block or restrict content are received from third parties, but are not adequately reviewed internally before being actioned.

Platforms of all sizes need to be empowered to be made more effective contributors towards child sexual abuse prevention, through a more nuanced and better-informed approach towards content moderation and censorship.

Unfortunately, to date two obstacles have prevented this from becoming a reality. First, many of the largest mainstream child protection organizations that have promoted platform liability rules as a solution to child sexual abuse have a broader agenda to eliminate adult content online, and they exclude perspectives of those who don't share that agenda, such as sex-positive therapists and researchers, LGBTQ+ people, sex workers, and the consensual kink community. As a result, there has been nobody to speak up when these communities become casualties of censorship such as over-blocking.

The second factor that has prevented platforms from taking a more nuanced and better informed approach towards content moderation and censorship as it relates to child protection is the powerful sexual stigma that affects all who work in this area. Although approaches based on the prevention of child sexual abuse are effective, stigma makes it difficult for this approach to make headway against the emotionally more resonant approach of identifying and prosecuting offenders. It also makes it difficult to suggest balances and safeguards for child protection laws and policies that are necessary in a free and democratic society.

The combination of these factors has created a tendency for platforms and policymakers to take an exceptionalist approach to child protection; in other words, to consider that it constitutes a *sui generis* policy area that overrides the way that we do things in other areas. But whenever we succumb to this tendency, it tends to mean doing away with important principles like due process and transparency. This is exactly the kind of thinking that led to FOSTA, and one of the reasons why ProStasia Foundation was formed was that we thought we could do better than that.

That's why we're very deliberately not starting this discussion of principles from scratch. There are already some best practice documents that have emerged from the child protection sector, including one from the UK Council for Child Internet Safety (2011) and another from Thorn (2013). But although they contain some worthwhile practical advice





for Internet companies, they have a more operational focus than the document of guiding principles that we have in mind. Furthermore, these documents were not developed in a human rights framework or with a primary prevention approach—so they have a lot to say about *how* content should be restricted, but very little about *whether* it should be.

So instead, we are using a set of principles that are grounded in international human rights norms, and based on years of work on guiding principles in the area of Internet content moderation, in documents such as the *Manila Principles on Intermediary Liability* (2015) and the *Santa Clara Principles on Transparency and Accountability in Content Moderation* (2018), both of which have been endorsed by hundreds of human rights defenders worldwide.

In May 2019, Access Now developed a set of principles for content moderation that draw upon these previous documents and others, titled *Protecting Free Expression in the Era of Online Content Moderation*. Rather than re-inventing the wheel, we presented this document to our meeting in San Francisco, related it to the presentations that had been given in the morning and to the case studies that had been discussed in the afternoon, and proposed using it as the basis for the development of a set of *Best Practice Principles for Sexual Content Moderation and Child Protection*.

Here is a slightly revised and expanded version of what we presented presented in San Francisco. The headlines of the Access Now principles are shown below in orange, followed by our own brief commentary on each of them:

Prevention of harm. The Internet industry must take human rights into account when making content moderation decisions and developing child protection policies. The rights of a child to bodily autonomy and privacy are critical considerations. And preventing violations of those rights is far preferable to just enforcing policies against those who have already violated them. But freedom of expression is also a right—for children, and for others in society. We do not accept that we have to give up one right for the protection of another. Industry can *and must* protect and uphold *all* human rights, for *all* human beings.

Evaluation of impact. A corollary of this is that content that *directly* harms a child must be restricted (there is no freedom of expression to sexually exploit children), but that restricting other types of content requires an appreciation of the impact that this restriction will have on all stakeholders. The *United Nations Guiding Principles on Business and Human Rights* (2008) require companies to:

Conduct due diligence that identifies, addresses and accounts for actual and potential human rights impacts of their activities, including through regular risk and impact assessments, meaningful consultation with potentially affected groups and other stakeholders, and appropriate follow-up action that mitigates or prevents these impacts.





In this context, the industry could do better at monitoring the human rights impacts of its decisions. Some companies have been censoring a broader range of sexual content than was necessary to comply with the law—even taking FOSTA into account—and this has been harming researchers, artists, sex workers, educators, child sexual abuse survivors, and other vulnerable people seeking support.

Transparency. Despite recent improvements, child protection remains one of the least transparent areas of content moderation. For example, when images are added to shared industry blocklists, how many “virtual” images such as cartoons, or sex education publications, or family photos are being included—and do we want them to be? If we do, are we sure that censoring them won’t cause more harm than it averts? Transparency is necessary to establish exactly where lines are being drawn, by whom, and on what basis. We commit to preparing a transparency report covering all major actors in online child protection, and publishing it this year.

Proportionality. The proportionate response in the case of online sexual grooming of a minor, or sharing unlawful images of minors, will often be account termination—and referral to law enforcement authorities. But this might not be proportionate in all cases—for example, when a 16 year old shares selfies with their 17 year old romantic partner or asks them for sex, many jurisdictions recognize that this is not child sexual exploitation—and platforms should take this into account when enforcing their policies in those jurisdictions. The proportionality of responses to other child protection policies should also be assessed individually; for example, sharing legal cartoon images, even if against a platform policies, is not as serious an infraction as sharing images that cause direct harm.

Context. As a starting point, no Internet platform is compelled to carry any particular legal content. Neither should it be compelled to prohibit any form of legal content. Its decision to carry content (or not) should, however, be human rights informed. This means sometimes making a judgment about the context in which content is published. Whether the content is being used to directly harm a child will often be the most relevant contextual consideration. Whether it is being published in an historical, artistic, or journalistic context will inform that assessment. We believe that the consideration of context should also be sex-positive; in other words, whether the content involves only consenting adults is another vital consideration. Where scale permits, context should be considered by a human before automated content filters are applied—except in the case of confirmed illegal images.

Non-discrimination. Sex workers, the consensual kink community, the LGBTQ+ community, those who have offended or are seeking support to avoid doing so, and the professionals who work with these populations, are often discriminated against on the basis of harmful stigmas. This sometimes perpetuates itself in the written or unwritten policies of Internet companies, seeing content from these groups removed while similar content is not removed when it is posted by other less socially stigmatized groups. This is harmful and must change.



Human decision. We do believe that automated decision-making has an important place in child protection: the use of hash-based databases for the elimination of duplicates of illegal child sexual abuse images and videos are an industry best practice that we support. But even in this case, it should always be a human decision that results in an image of video being added to such a database—and we are calling for better transparency and accountability of this process. Artificial intelligence algorithms should not yet be used to identify new content as being an unlawful sexual image of a child or an unlawful act of child sexual grooming, without individual human review—and any such technologies must respect privacy.

Notice. The Internet industry has been improving the comprehensiveness of its notifications to those whose content is removed by a moderator. It is time for similar improvements to be made so that when websites URLs and content hashes are added to shared industry blocklists, the owner of that website or content is informed of this—so that they can defend it if necessary—and determine that the content is no longer blacklisted if the defense is successful.

Remedy. When a platform slips up in its content moderation decisions, it should provide a remedy to those affected. But FOSTA takes this a step further, allowing platforms to be held liable directly to victims of sex trafficking for the effects of that abuse. This goes a step too far—because it encourages platforms to over-censor and over-enforce, and distracts us from pursuing justice against actual perpetrators of child sexual abuse. This doesn't mean that platforms shouldn't be held responsible when they do too little (or too much) to prevent child sexual abuse. But far better results will be obtained if they provide that remedy through means like those we have outlined above—rather than by paying out to individual survivors of child sexual abuse through lawsuits. For this reason, we reject the idea that content moderation should be expressed in terms of a “duty of care” by Internet platforms, as recently proposed in the United Kingdom.

At the conclusion of our meeting in San Francisco, we called for volunteers to join a multi-stakeholder working group that would take these draft ideas and form them into a set of draft principles for presentation at RightsCon 2019 for further community feedback and discussion. The first fruits of that effort are shown on page 33. Between now and November 2019, we will be continuing to refine this work to produce a set of principles that enjoys a broad consensus from a diverse group of experts and stakeholder groups.

It is further intended that the principles might be used as the basis for the development of more detailed operational policies and standards, such as a set of model terms of service for Internet platforms with respect to child protection that smaller Internet platforms



can easily adapt and use. The principles will also be used as the basis for the development of a “no children harmed” certification mark that ProStasia Foundation is currently developing, that we briefly announced at our San Francisco meeting.

We’ll be explaining how you can participate in the further development and application of the best practice principles at our meeting in Tunis. If you can’t make it to RightsCon or are reading this afterwards, be sure to follow our social media feeds or subscribe to our newsletter, where we will also be posting more information.

The final version of the *Best Practice Principles for Sexual Content Moderation and Child Protection* will be presented at the 38th Annual Research and Treatment Conference of the Association for the Treatment of Sexual Abusers from November 6-9, and again at the 14th Annual Meeting of the United Nations Internet Governance Forum (IGF) in Berlin from November 25-29.

REFERENCES

Access Now (2019). *Protecting Free Expression in the Era of Online Content Moderation*. Available from: <https://www.accessnow.org/free-expression-at-scale-a-human-rights-guide-to-content-moderation/>

Electronic Frontier Foundation (2015). *Manila Principles on Intermediary Liability*. Available from: <https://manilaprinciples.org/>

Electronic Frontier Foundation (2018). *Santa Clara Principles on Transparency and Accountability in Content Moderation*. Available from: <https://santaclaraprinciples.org/>

Kaye, David (2018). *Report of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression*. Available from: http://ap.ohchr.org/documents/dpage_e.aspx?si=A/HRC/38/35

Thorn (2013). *Sound Practices Guide to Stopping Child Sexual Abuse*. Available from: <https://www.thorn.org/sound-practices-guide-stopping-child-abuse/>

UK Council for Child Internet Safety (2011). *Good Practice Guidance for the Providers of Social Networking and Other User-Interactive Services*. Available from: <https://dera.ioe.ac.uk/id/eprint/11099>

UN Human Rights Council (2008). *Protect, Respect and Remedy: A Framework for Business and Human Rights*. Available from: <http://www.refworld.org/docid/484d2d5f2.html>



DRAFT BEST PRACTICE PRINCIPLES FOR SEXUAL CONTENT MODERATION AND CHILD PROTECTION

Version 0.2–8 June 2019

Prevention of harm	Sexual content should be restricted where it causes direct harm to a child. Indirect harms should not be the basis for blanket content restriction policies unless those harms are substantiated by evidence, and adequate measures are taken to avoid human rights infringements.
Evaluation of impact	Companies should evaluate the human rights impacts of their restriction of sexual content, meaningfully consult with potentially affected groups and other stakeholders, and conduct appropriate follow-up action that mitigates or prevents these impacts.
Transparency	Companies and others involved in maintaining sexual content policies, databases or blocklists should describe the criteria for assessing such content in detail, especially when those policies would prohibit content that is lawful in any of the countries where such policies are applied.
Proportionality	Users whose lawful sexual conduct infringes platform policies should not be referred to law enforcement, and their lawful content should not be added to shared industry hash databases, blocklists, or facial recognition databases.
Context	The context in which lawful sexual content is posted, and whether there are reasonable grounds to believe that the persons depicted in it have consented to be depicted in that context, should be considered before making a decision to restrict or to promote it.
Non-discrimination	Content moderation decisions should be applied to users based on what they do, not who they are.
Human decision	Content should not be added to a hash database or blocklist without human review. Automated content restriction should be limited to the case of confirmed illegal images as identified by a content hash.
Notice	Users should be notified when their content is added to a hash database or blocklist, or is subject to context-based restrictions.
Remedy	Internet platforms should give priority to content removal requests made by persons depicted in images that were taken of them as children, and provide users with the means of filtering out unwanted sexual content.



BIOGRAPHIES OF SPEAKERS

MULTI-STAKEHOLDER
DIALOGUE ON

INTERNET PLATFORMS
SEXUAL CONTENT AND
CHILD PROTECTION

at
RightsCon



JEREMY MALCOLM (SAN FRANCISCO AND TUNIS)

is Executive Director of ProStasia Foundation. Prior to his work at ProStasia he held the position of Senior Global Policy Analyst at the Electronic Frontier Foundation, where he led the development of the Manila Principles on Intermediary Liability and worked in numerous other areas defending digital privacy, free speech, and innovation. Before that he worked for Consumers International coordinating its global program Consumers in the Digital Age.

Jeremy graduated with degrees in Law (with Honors) and Commerce in 1995 from Murdoch University, and completed his PhD thesis at the same University in 2008 on the topic of Internet governance. He is the author of Multi-stakeholder Governance and the Internet Governance Forum.

Jeremy's background is as an information technology and intellectual property lawyer and IT consultant. He is admitted to the bars of the Supreme Court of Western Australia (1995), High Court of Australia (1996) and Appellate Division of New York (2009). He is a member of the Multistakeholder Advisory Group of the United Nations Internet Governance Forum.



GILIAN TENBERGEN (SAN FRANCISCO)

is currently a Visiting Assistant Professor of Psychology with SUNY Oswego, specializing in the neurobiology of pedophilia and its translation into the prevention of child sexual abuse. In Hannover, Germany, she spent several years working with the German Prevention Project Dunkelfeld and the German national NeMUP Research Consortium. Her research interests focus primarily on the roles of neuropsychology and neurobiology on the development of paraphilias and risk factors in the commission of sexual offenses.





CATHY GELLIS (SAN FRANCISCO)

is a lawyer in the San Francisco Bay Area with a practice focused on intellectual property, free speech, intermediary liability, privacy, and other innovation policy matters affecting technology use and development. She regularly writes, speaks, counsels, and litigates on these and other related topics, particularly with respect to how these issues relate to Internet platforms. Examples of her work include defending the free speech rights of anonymous bloggers, representing an organization of college webcasters before the Copyright Royalty Board, and authoring numerous amicus briefs, including in litigation challenging NSA Internet surveillance. Her writing on the policy implications of technology regulation has appeared in various widely-read publications, including the Daily Beast, Law.com, and the technology news site Techdirt.com, where she is a regular contributor. Prior to becoming a lawyer she was an aspiring journalist-turned-Internet professional who developed and managed websites for enterprises in Silicon Valley and Europe. She has a B.A. from the University of California at Berkeley in Mass Communications and Sociology, where she studied information technology and user adoption trends, and a J.D. from Boston University.



ANDREW PUDDEPHATT OBE (SAN FRANCISCO)

is Chair of the Internet Watch Foundation. He has worked to promote human rights for twenty years with specific expertise in freedom of expression, transparency, and the role of media and digital communications in promoting human rights. Specific skills include:

- Policy analysis, evaluation and advice on projects concerned with communications, human rights and democracy internationally. Advisor to the Ford Foundation on digital issues. Provided guidance on the internet's implications for human rights to the Open Society Foundation, the UN Special Rapporteur for Freedom of Expression and the OAS Special Rapporteur for Freedom of Expression.
- Expert in international human rights, with a specific focus on freedom of expression and access to information. A specialist in modern digital communications and transparency issues. For UNESCO, has produced the Media Development Indicators, Journalist Safety indicators and a model set of internet indicators.
- Regular public speaker and presenter on internet public policy issues; developed and delivered training programs for the EU on political economy analysis and its application to the field.



- Previous work includes strategic policy analysis, scoping exercises, evaluation of programs, needs assessments, and the development and delivery of capacity-building programs. Has worked with experts and governments in Europe, the United States, the Middle East, sub-Saharan Africa, and Latin America.

Andrew's other roles include chairing the Danish based NGO International Media Support, and serving on the board of both the Sigrid Rausing Trust and the European Council for Foreign Relations. He is an entrepreneur having founded two social enterprises: Adapt in 2017, and Global Partners Digital where he remains the Chair and in which capacity he leads the secretariat for the 30 country Freedom Online Coalition.



CANDICE CHRISTIANSEN (SAN FRANCISCO)

is the founder of several programs including Namasté Center for Healing, The Global Prevention Project, and the Prevention Podcast. Candice and her expert team provide cutting edge, evidence based treatment for men, women, couples and families with trauma, problematic behavior related to porn/sex addiction, risky sexual attractions and behavior, infidelity; as well as a variety of relational/sexual issues, substance abuse, & financial disorders.

Candice is an expert witness who administers sex specific mental health evaluations in state and federal courts, for the Board of Pardons and medical/professional licensing boards, & immigration cases. Candice is a member in good standing with the Association for the Treatment of Sexual Abusers (ATSA) Prevention Committee & Diversity Subcommittee. She serves on the DOPL Substance Abuse Board in Utah, is on the advisory board for Legal Reform For Individuals with Intellectual and Developmental Disabilities(LRIDDD), & is a member of the United Nations Human Trafficking Task Force. She authored *Mastering The Trauma Wound: A Mindful Approach to Healing Trauma and Creating Healthier Relationships* and is published in various online magazines and newspapers nationwide.



GUY HAMILTON-SMITH (SAN FRANCISCO)

is a graduate of the University of Kentucky College of Law and is the Sex Offense Litigation and Policy fellow at the Mitchell Hamline School of Law. He has written about these issues for Harvard Law's Fair Punishment Project, the Texas Journal of Civil Liberties & Civil Rights, the American Bar Association, and other outlets. You can follow him on Twitter @G_Padraic, or read more of his writing at his website, guyhamiltonsmith.com.





CATHY BEARDSLEY (SAN FRANCISCO)

is President and CEO of Segpay, one of four companies approved by Visa USA to operate as a high risk Internet Payment Service Provider (IPSP) in the U.S. With more than 600 merchants, Segpay has become a top brand among IPSP's in the U.S. and Europe. Beardsley has more than a dozen years of experience leading and managing payment processing companies.

Prior to founding Segpay, Beardsley was President for the Internet Billing Company (iBill). Prior to her IPSP work, Beardsley spent seven years with telecom giant MCI WorldCom. Cathy earned a B.A. in Business Administration/ Marketing from California Polytechnic University, San Luis Obispo and a MBA in Business Administration/ Finance from the University of Santa Clara.



KRISTIN DIANGELO (SAN FRANCISCO)

is a community organizer, activist, educator and the Executive Director of The Sex Workers Outreach Project Sacramento. Kristen has worked worked with Amnesty International, The Harm Reduction Coalition, and a Special Rapporteur to the UN and the Geneva Convention to formulate their views on the sex trade and how to best reduce harm. Executive Producer and Interviewee in of the award winning documentary American Courtesans, Kristen co-wrote the needs assessment Sex Work and Human Trafficking in the Sacramento Valley; A Needs Assessment which is currently used as teaching curriculum in Universities across the US.



IAN O'BRIEN (SAN FRANCISCO)

is Senior Director of Programs and Operations at the Free Speech Coalition. Ian works with regulators of both the pleasure products and adult film industries, to compile, analyze and provide accurate, evidence-based research on issues ranging from zoning and lubricants to workplace safety and public health. He also works on research on the industry as a whole, working to combat pseudo-scientific concepts like porn addiction and the porn public health crisis, a crucial focus at a time when over two dozen states are using such studies to ban or censor adult content.

O'Brien attended graduate school at Columbia University's prestigious Mailman School of Public Health.





HEATHER BARR (TUNIS)

is the acting co-director of the Women's Rights Division at Human Rights Watch. She has done research in countries including Afghanistan, Bangladesh, Burma, Nepal, and Papua New Guinea, on issues including child marriage, girls' education, violence against women, refugee and prisoners' rights, and trafficking.

She joined Human Rights Watch in 2011 as the Afghanistan researcher, after working for the United Nations in Afghanistan and Burundi. After law school she litigated for discharge planning for prisoners with psychosocial disabilities in New York City, and founded an alternative-to-incarceration program. Before law school, she worked with homeless women. She is a graduate of London School of Economics, Columbia Law School, and John Jay College of Criminal Justice.



JAC SM KEE (TUNIS)

is Jac sm Kee is the APC Women's Rights Programme manager and works on the issues of sexuality, women's rights, violence against women and internet rights and freedoms. She initiated the Take Back the Tech! global campaign and the EROTICS (exploratory research on sexuality and the internet) project. Jac currently serves as a board member for the Association for Women's Rights in Development (AWID) and the New York chapter of Creating Resources for Empowerment and Action (CREA), and as director of the Centre for Independent Journalism, Malaysia. She has published several papers on the issues of internet governance, censorship, women's rights and sexuality. Jac read law and gender at the University of Nottingham and the London School of Economics and Political Science, respectively.



JILLIAN C YORK (TUNIS)

is a writer and activist whose work examines the impact of technology on our societal and cultural values. Based in Berlin, she is the Director for International Freedom of Expression at the Electronic Frontier Foundation and a fellow at the Center for Internet & Human Rights at the European University Viadrina.

Jillian co-founded Onlinecensorship.org, an award-winning project that seeks to encourage companies to operate with greater transparency and accountability toward their users as they make decisions that regulate speech. She is a frequent public speaker on





topics including censorship, surveillance, and the impact of social media on our lives and our societies. Her writing has been published by the New York Times, Al Jazeera, the Atlantic, the Guardian, Quartz, The Washington Post, and Die Zeit, among others.



MARIEL GARCÍA-MONTES (TUNIS)

is a public interest technology capacity builder and researcher from Mexico. Her main topics of interest are privacy and information security, social exclusions in technology and participatory processes. Most recently, Mariel was a graduate student at the Comparative Media Studies program at MIT, and a research assistant at the Center for Civic Media at the MIT Media Lab. There, she worked at the Codesign Studio and the Design Justice project, and wrote a thesis on tensions in organizational approaches to work on youth and privacy issues in the Americas.

Mariel has worked in communications, instructional design and research around open data, privacy and security, strategic communications and other digital literacies for organizations in Mexico and around the world. She is a philosophy graduate from the National Autonomous University of Mexico.

Mariel loves passionate opinions in the intersections of technology and society, creative communications efforts, random acts of kindness, passport stamps and both eating and dancing salsa.





ABOUT PROSTASIA FOUNDATION

THE STIGMA-BUSTING CHILD PROTECTION ORGANIZATION

Prostasia Foundation is the only s.501(c)(3) nonprofit child protection organization that is progressive and sex-positive, and supports the free and open Internet. Our diverse team includes digital rights activists, lawyers, mental health professionals, and representatives of sex workers and other stakeholders who are normally excluded from discussions about child protection.

<https://twitter.com/ProstasiaInc>



<https://fb.com/ProstasiaInc>

