



PROSTASIA
FOUNDATION

Keeping platforms safe, inclusive, peaceful, and healthy

June 4, 2020



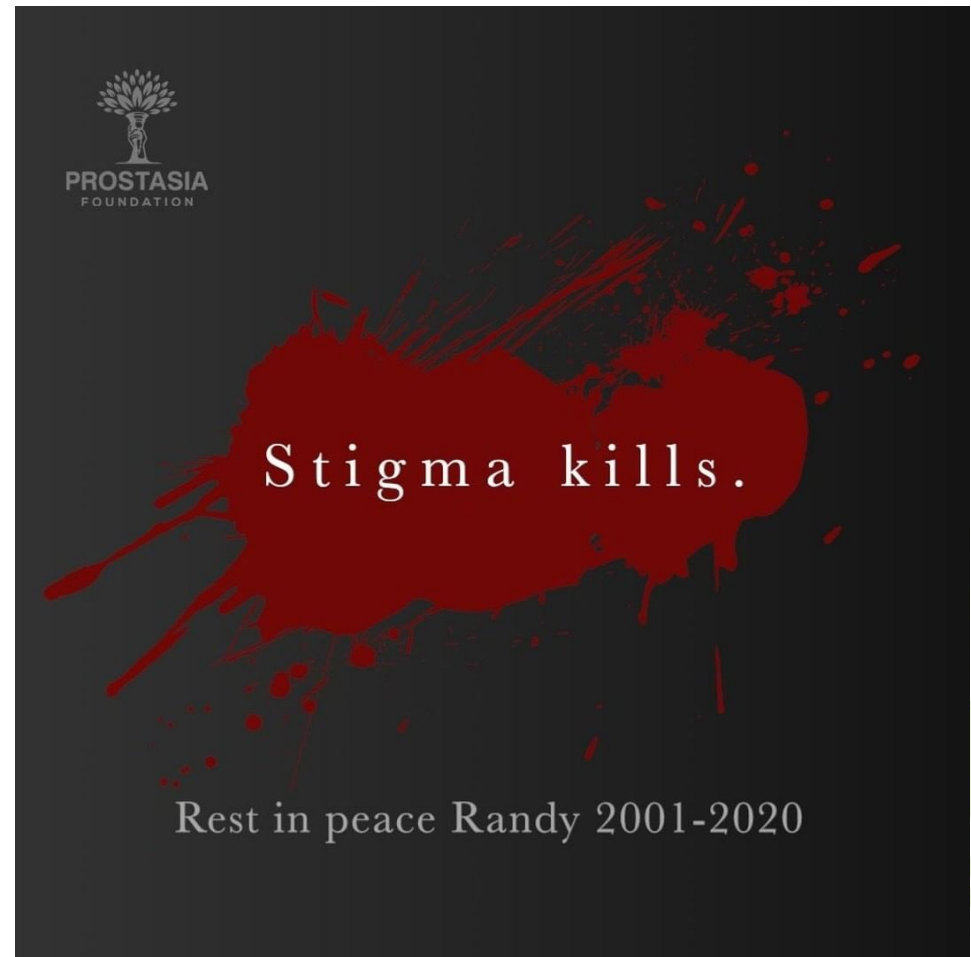
Outline

- Background to the problem
- History of our work with platforms
- Best practice principles for sexual content moderation and child protection
- Two case studies
- A call to industry
- Discussion questions



Background to the problem

- Major platforms are hotbeds for abuse
- Minority content isn't worth protecting
- Poor privacy settings expose minors and survivors to NSFW art
- Engagement model is mentally unhealthy



Our work with Internet platforms

- 2019 Multi-stakeholder Dialogue on Internet Platforms, Sexual Content & Child Protection
- RightsCon 2019 and online working group
- Publication of #SexContentDialogue principles at IGF 2019
- No Children Harmed certification soft launch in Korea in October
- Ongoing consulting work





BEST PRACTICE
PRINCIPLES FOR

**SEXUAL CONTENT
MODERATION AND
CHILD PROTECTION**

VERSION 1.0 — NOVEMBER 2019

- Prevention of harm
- Evaluation of impact
- Transparency
- Proportionality
- Context
- Non-discrimination
 - Human decision
 - Notice
 - Remedy

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, unless such notification would be prohibited by law.



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.



Case study 1: Fanexus

- An excerpt from the draft Fanexus Community Guidelines:
 - All pornographic artworks and animations depicting real minors are prohibited; this includes erotic artworks/animations of fictional characters portrayed by underage actors, if the depictions are undeniable likenesses of the actors.
- Community Guidelines are designed as the starting point for a template for other sites



Case study 2: 8kun

- 8kun is an anonymous chan site
- They came to us because of conflicting advice on the legality of anime pornography
- We helped to clarify the situation based on their risk profile and user preferences



What are we worth?

- The only child protection organization opposing EARN IT and supporting s.230
- A unique sex-positive, human rights focus
- 17,000 signatures to our petition to the UN
- Blog and monthly podcast/vodcast series
- Served anti-CP ads to 250,000 people
- Monthly newsletter with 6000+ subscribers



How you can help

- Follow us on Twitter @ProstasiaInc
- Like us on Facebook
<https://fb.com/ProstasiaInc>
- Visit <https://prostasia.org> to:
 - Subscribe to our newsletter
 - Become a member or sponsor
- Hire our consultancy services
- Include us in industry events



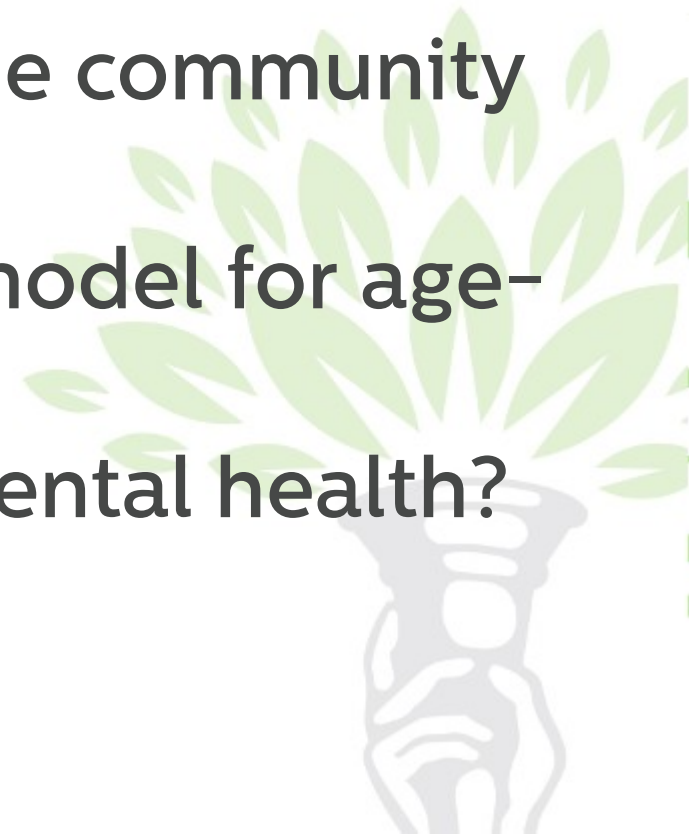
Resources mentioned

- No Children Harmed policy:
<https://prostasia.org/no-children-harmed>
- #SexContentDialogue principles:
<https://prostasia.org/sexual-content-moderation-principles/>



Discussion questions

- What are some of the product design problems that can lead to toxic engagement?
- What role should users have in development of platform or community policies?
- When can gatekeeping of online community membership be justified?
- Is self-identification the right model for age-gating?
- How can platforms promote mental health?





PROSTASIA
FOUNDATION