CYBER CIVIL RIGHTS INITIATIVE
Coral Gables, FL
CyberCivilRights.org

Preventing Image-Based Sexual Abuse (IBSA): CCRI Best Practice Recommendations for Tech Companies

Cyber Civil Rights Initiative (CCRI) offers the following recommendations for platform providers, content hosting services, and related tech companies to address the immediate harms of Image-Based Sexual Abuse (IBSA), including Nonconsensually Distributed Intimate Imagery (NDII); synthetic NDII (also known as "deepfake" intimate images); and sextortion.

1. **Development of Industry Standards on Consent:** Tech platforms should implement robust consent safeguards for all tools and services, ensuring that no user is allowed to generate or distribute synthetic or actual intimate imagery of any other person without consent.

    a. Proof of Consent: Currently, most consent protocols require that the uploader click a dialog box that confirms that the uploader received consent from the individuals depicted. This consent protocol is inadequate for the distribution of intimate images, which should require a higher level of oversight, including proof of consent.

    b. Automatic Detection: Use hashing technology and/or machine learning/AI to remove or block the upload of images previously identified as nonconsensual.

    c. Opt Out: Tech platforms should consider honoring an individual's withdrawal of consent to the distribution of intimate imagery even if the individual previously consented to it.

2. **Development of Industry Standards on Age Verification:** Tech platforms should implement robust consent safeguards for all tools and services, ensuring that no underage user is able to generate actual or synthetic intimate imagery of themselves.

a. CCRI advises against the use of biometric authentication of minors and discourages the default use of biometric authentication for adults.

b. CCRI notes that significant concerns related to privacy and accuracy will arise with most identity verification methods and therefore encourages research into the strengths and weaknesses regarding the various approaches. These may include document verification (government identification, credit cards, etc.); third-party verification (guardian, public database, mobile phone carrier, or other third-party verification); AI and behavioral analysis; or other age verification measures.

3. **Limiting downstream distribution:** Companies should prohibit both downloads and screen shots of intimate imagery in order to limit the redistribution of NDII and synthetic NDII. Companies should also innovate tools that can detect cropped or filtered images, which can evade detection through current hashing technology, allowing images to "go viral."

4. **AI For Good/Improving Detection Tools or Algorithms**: The tech industry should make use of advanced machine learning and AI techniques to better detect IBSA burner accounts, grooming language, and other potentially harmful behavior.

5. **Responsible Content Moderation:** Tech companies must ensure that their training and employment practices for human content moderators—especially those who will be exposed to violent, sexually explicit, or racist material—include transparency about the requirements of the work and offer long-term mental health support, appropriate compensation, and other health and safety guarantees. Companies should invest in research to determine the effects of extended exposure to disturbing content and ways to mitigate these effects.

6. **Privacy Settings:** All AI and tech platforms should engage the strictest default privacy levels for new users, including, at a minimum, hidden contacts; restricted interactions to user-approved contacts only; filtered search that excludes explicit content; and two-factor authentication.

7. **Synthetic NDII:** In addition to all steps described in this document, tech companies should amend protocols to explicitly prohibit the use of machine

learning or other technologies to create synthetic NDII; utilize AI algorithms to detect synthetic NDII; identify such images through watermarks and warning notifications to users; and immediately request consent verification from the person depicted if known, or from uploaders if the subject's identity is not known.

8. **Proactive User Education:** Tech companies should integrate safety prompts and messaging directly into the user experience. These could include warning messages if a post appears to be a NDII or synthetic NDII; cautions to the user to cease their interaction with material of concern; prompts on how to seek help; and notifications regarding other related suspicious or unlawful activity alerts.

9. **Efficient Reporting Mechanisms for NDII:** Tech companies should implement IBSA reporting mechanisms with the following characteristics:

    a. Easy to locate and easy to use in stressful circumstances. Reporting mechanisms should not be buried in hard-to-access areas of a site, but instead should be clearly marked and available next to the synthetic NDII material and/or offender account.

    b. Allow for batch reporting.

    c. Include a transparent reporting dashboard where users can track the status of their reports in real time.

    d. Result in prompt corrective action. Both burner accounts and repeat offender accounts should be deactivated permanently and, when possible, blocked at the device and/or IP level. Complaints may be addressed through an appeals procedure.

10. **Cross-Industry Collaboration.** Tech companies should collaborate to prevent and respond to IBSA by sharing best practices; denylists of known offenders and/or hashes; technological solutions; and educational strategies and campaigns.

11. **Annual External Audit:** Tech companies should engage reliable, external third parties to conduct and publish annual safety audits, so that the public can make independent and informed choices about which products and platforms to use rather than relying on a company's self-produced reports or assurances.

Federal Oversight of Tech companies

1. **Legislation and Policy:** CCRI encourages the passage of carefully targeted, First Amendment-compliant, comprehensive state and federal laws to regulate NDII, sextortion, and digital forgeries; Section 230 reform to halt the tech industry's use of recklessness as a design concept; and the adoption of global norms and agreements.

2. **Creation of a Global NDII Hash Database:** NDII and synthetic NDII that appear on one tech platform should be hashed and banked in a database to facilitate the efficient removal of the imagery and prevent its proliferation on other sites. The hash database used by National Center for Missing & Exploited Children (NCMEC) for child sexual exploitation material (CSEM) provides a useful model.

3. **Financial Penalties:** Companies who are out of compliance with any of the above should be fined, and those fines should be allocated to victim-survivor compensation funds, direct service providers, researchers, schools, and other civil society organizations.