

Response to RFI on NIST's Assignments under Sections 4(1), 4(5), and 1(1) of the National Defense Authorization Act for Fiscal Year 2023

To Whom It May Concern,

I am writing on behalf of the South West Grid for Learning (SWGfL) and its Revenge Porn Helpline in the UK, in response to the Request for Information (RFI) by the National Institute of Standards and Technology (NIST). The Revenge Porn Helpline has been instrumental in supporting adult victims of non-consensual intimate image abuse since 2015, offering significant insights into this evolving and increasingly complex issue.



Updated Evidence and Trends from the UK Revenge Porn Helpline:

1. Surge in Non-consensual Image Abuse Reports:

- In 2023, reports received by the Helpline more than doubled to 18,421, from 8,924 in 2022, indicating a disturbing rise in such abuses.

2. High Content Removal Rate and Gender Disparity:

- With an 89% success rate in content removal, the Helpline reported 11,923 images in 2023, 94.6% of which depicted women. This underscores a significant gender imbalance in victimization.

3. Increased Sextortion Cases:

- Sextortion cases rose to 35% of total reports in 2023, a 54% increase from 2022, predominantly affecting male victims (93%).

4. Perpetrator Dynamics:

- Analysis reveals gender-specific patterns in perpetration: 67% of female victims reported a current/former partner as the perpetrator (where known), whereas 89% of male victims faced sextortion typically by criminal gangs.

5. Rising Incidence of Deepfake Intimate Imagery:

- A concerning trend in 2023 was the doubling of reports from victims of deepfake intimate imagery. The impact of such synthetically created images is as devastating as genuine intimate image abuse, leading to severe psychological and social consequences.
- Recent advancements in technology facilitating the creation of synthetic imagery have predominantly yielded pornographic content. Research (<https://www.homesecurityheroes.com/state-of-deepfakes/#key-findings>) indicates that 98% of deepfakes are pornographic, with 99% targeting women.

- Nudification apps, which disproportionately target women and are readily available in app stores, exemplify this trend. The research also highlights that 94% of targets are in the entertainment industry, differing from other forms of intimate image abuse typically aimed at causing direct harm.
- Deepfake content appears driven more by personal gratification and peer sharing, rather than direct harm to individuals.

Given this alarming situation, it is imperative to develop robust standards for AI technology to prevent non-consensual intimate image abuse or other tech facilitated gender based violence. Facial recognition and other AI technologies themselves should be able to provide opportunities to help identify and prevent synthetic intimate image content from being shared online without consent. Companies should be required to ensure their technologies are not used to abuse others and be accountable if insufficient measures or protections are not deployed. SWGfL supports the development of technical standards relating to AI technologies, for example Content Authenticity Initiative.

Technologies like [StopNCII.org](https://stopncii.org) are pivotal in this fight. StopNCII.org offers a unique solution for victims, allowing them to prevent the further sharing of their intimate images online on participating platforms. Using advanced hashing technology, it creates a digital fingerprint of the images, which can then be used by platforms to block or remove these images, thus preventing their spread.

About StopNCII.org

StopNCII.org is the world's first device-side hashing technology available to anyone in the world. It allows people concerned that their intimate images may be shared without their consent, including those under the imminent threat, to protect them by creating a unique hash (a digital fingerprint that uniquely identifies the image or video) which is then shared with participating platforms to stop the images being shared or reshared. Crucially, the images never leave the victim's device, giving total trust and security and protecting their privacy. Launched in December 2021, StopNCII.org has already supported over 200,000 cases and is protecting over 500,000 images. To date, it has actively enabled its partners to prevent over 10,000 intimate images from being shared online.

[For more information on how StopNCII works.](#)

Moreover, SWGfL has encountered cases where children have used AI to create and share intimate images of their peers. Publishing an article that highlighted [Children must understand risk as UK schools say pupils abusing AI to make sexual imagery of other children - UK Safer Internet Centre](#), this emerging trend underscores the need for stringent standards and ethical guidelines in AI development, particularly concerning the creation and distribution of synthetic content.

Victim/Survivor Voice

SWGfL has a number of impact statements from victims and survivors of NCII, including synthetically created content using deepfake and AI. SWGfL would be happy to confidentially share these impact statements to inform this process.

Recommendations to NIST:

1. Development of AI Technical Standards:

- The escalation of non-consensual intimate image abuse, exacerbated by AI and deepfake technologies, necessitates urgent development of AI technical standards to prevent the creation and distribution of such content.

2. Mandatory Integration of Prevention Measures on Platforms:

- Platforms should be required to integrate existing technologies like hashing to proactively combat the spread of non-consensual intimate images.

3. Addressing Gender Disparities and Sextortion:

- Standards and platform obligations must address the pronounced gender disparities and the unique challenges posed by sextortion.

4. Specific Focus on Deepfake Abuse:

- Given the alarming rise in deepfake abuses, standards should include mechanisms to detect and prevent the spread of synthetically created intimate images, particularly those targeting women.

In conclusion, the evolving nature of online harms, highlighted by SWGfL's data and experience, underscores the need for comprehensive and dynamic AI technical standards and platform obligations. These measures must be inclusive and reflective of the diverse experiences of victims to create a safer online environment.

Thank you for considering our input.

About SWGfL

SWGfL (www.swgfl.org.uk) is a UK charity, working internationally to ensure that everyone can benefit from technology, free from harm. SWGfL is 23 years old and is responsible for award winning services, resources and campaigns as well as operating three Helplines that each support victims of differing online harms ([Revenge Porn Helpline - 0345 6000 459](#) | [Revenge Porn Helpline](#), [Report Harmful Content - We Help You Remove Content](#) and [Professionals Online Safety Helpline - UK Safer Internet Centre](#)). In December 2021, it released [Stop Non-Consensual Intimate Image Abuse | StopNCII.org](#)

It works with various Government Departments both at home and abroad, tech sector and other national and international agencies. In 2020, DCMS recognised SWGfL as one of the [UK Safety Tech sectoral analysis - GOV.UK \(www.gov.uk\)](#).

SWGfL greatly values close collaboration with, and contribution to [the Global Partnership for Action on Gender-Based Online Harassment and Abuse - United States Department of State](#). and participated in a [White House Roundtable on AI Safety](#) in November 2023.

SWGfL, alongside partners Childnet and IWF, are collectively the [UK Safer Internet Centre](#). First created in 2011, the UK Safer Internet Centre is the UK national awareness centre and is responsible for raising the nation's attention to online safety issues as well as managing online criminal content

and supporting professionals via its unique helpline. It is one of 32 European Safer Internet Centres that make up the Insafe network.

About Revenge Porn Helpline - 0345 6000 459

The first of its kind in the world, the Revenge Porn Helpline (RPH) was established by SWGfL in 2015 and supports victims in the UK of non-consensual intimate image abuse. Since its creation, the Revenge Porn Helpline has supported thousands of victims. With an over 90% removal rate, they have successfully removed over 300,000 individual non-consensual intimate images from the internet.

The Revenge Porn Helpline is co funded by the Home Office, Nominet, donations and charitable reserves.

In 2019 the Helpline managed 1,600 cases compared to just under 11,000 in 2022 ([RPH 2022 Report](#)). In 2023, the Helpline continues to see dramatic increases in case load, most recently [Revenge Porn Helpline Sees Record Number of Cases in August | SWGfL](#). (the caseload in August 2023 exceeded the number of cases received during the first 3 years of operations)