



NATIONAL  
CENTER ON  
SEXUAL  
EXPLOITATION

Established 1962

KINDSEY PENTECOST CHADWICK | Interim President

DAWN HAWKINS | CEO

BENJAMIN A. BULL, ESQ. | SVP & General Counsel

#### BOARD MEMBERS

RON DEHAAS | Chairman  
Founder & CEO, Covenant Eyes

DAN O'BRYANT, ESQ. | Vice Chairman  
Former Fellow, Weatherhead Center for International  
Affairs, Harvard University

RHONDA GRAFF | Secretary  
Former Aerospace Program Manager/Industrial Engineer

BETH COONS | Treasurer  
Chairman, Farnsworth Companies

NATHAN ANDERSEN, ESQ.  
General Counsel, LeSeuer Investments

KINDSEY PENTECOST CHADWICK | Interim President  
President Emeritus, ItWorks! International

JOHN FOUBERT, PH.D.  
Dean, College of Education, Union University

ROBERT GEORGE, PH.D.  
McCormick Professor of Jurisprudence,  
Princeton University

HON. TANYA GOULD  
Director, Virginia Attorney General's Office on  
Human Trafficking

DONALD L. HILTON, M.D.  
Neurosurgeon, University of Texas - San Antonio

DONNA M. HUGHES, PH.D.  
Retired Professor, Carlson Endowed Chair,  
University of Rhode Island

AMBASSADOR SWANEE HUNT  
Chair and Founder, Swanee Hunt Alternatives;  
Eleanor Roosevelt Lecturer of Public Policy,  
Harvard Kennedy School of Government

MARY ANNE LAYDEN, PH.D.  
Director, Sexual Trauma and Psychopathy Program,  
University of Pennsylvania

MOST REVEREND PAUL S. LOVERDE  
Bishop, Dioceses of Arlington, VA, Retired

SHERRY NILLES  
National Regent, Catholic Daughters of the Americas

PATRICK A. TRUMAN  
President & CEO Emeritus of NCOSE

LINNEA W. SMITH, M.D.  
Retired, Psychiatrist

MELEA STEPHENS, L.P.C., M.M.E.T.  
Marriage and Family Therapist

JASON SYVERSON  
Chair & Chief Investment Officer, S8  
Impact Foundation

CAROLYN WEST, PH.D.  
Professor, Clinical Psychology,  
University of Washington

July 30, 2024

Mr. Sam Altman, CEO  
OpenAI  
3180 18<sup>th</sup> Street, Suite 100  
San Francisco, CA 94110

### RE: National Center on Sexual Exploitation's Rapid Assessment Report *The High Stakes of AI Ethics: Evaluating OpenAI's Potential Shift to "NSFW" Content and Other Concerns*

Dear Mr. Altman,

Given OpenAI's expressed desire for artificial general intelligence (AGI) to benefit humanity and to understand user and societal expectations for its AGI models, your pledges at the AI Seoul Summit, as well as your commitment to safety by design standards for preventing child sexual abuse, **we write to express deep concern and alarm regarding:**

- 1) OpenAI's lack of dataset transparency and the very likely unethical inclusion of sexually explicit material in its pre-training and training datasets**
- 2) the perils of OpenAI's partnership with Reddit due to Reddit's data being riddled with sexually explicit, as well as sexual exploitation material**
- 3) the threats of harm posed by OpenAI's potential "NSFW" rule change pertaining to the generation of sexually explicit material/hardcore pornography.**

The National Center on Sexual Exploitation (NCOSE) is the leading organization preventing sexual abuse and exploitation on a mass scale by eliminating institutional practices and societal norms that perpetuate these harms. As leading experts in combating sexual exploitation, NCOSE is uniquely positioned to assist you in the critical endeavor of building AI that helps the world. Our decades of experience, combined with extensive research, direct engagement with hundreds of survivors, and advocacy with many other tech giants, equips us with a profound understanding of the impacts—both positive and negative—of technology on society.

The National Center on Sexual Exploitation recognizes the incredible promise of AI and the contributions of OpenAI to this rapidly developing field. OpenAI's technology will help solve some of the most pressing issues the world faces today and could help solve problems related to sexual exploitation—if it so chooses.

1201 F St. NW, Suite 200, Washington, DC 20004 | 202.393.7245 | [EndSexualExploitation.org](https://EndSexualExploitation.org) | [public@ncose.com](mailto:public@ncose.com)

BUILDING A WORLD WHERE ALL PEOPLE CAN LIVE AND LOVE FREE FROM SEXUAL ABUSE AND EXPLOITATION

However, ignoring the substantial harm caused by the misuse of AI tools is impossible. We cannot sugarcoat the reality: the misuse of AI tools has unleashed massive sexual harm, impacting thousands who will suffer the consequences for their entire lives. **OpenAI's lack of dataset transparency and its very likely inclusion of sexually explicit material in pre-training and training datasets, its partnership with Reddit, and its proposed "NSFW" rule change are poised to further these harms.**

The impact of AI on the field of sexual abuse, violence, and exploitation prevention has already been catastrophic. Examples of AI and related technologies fueling an explosion in sexual exploitation include:

- **chatbots** fabricating allegations of sexual assault against real persons, disseminating harmful sexual advice, and their potential use to scale child victimization through grooming
- **nudifying apps** spawning a surge of nonconsensual sexually explicit images, affecting thousands of women and children
- **AI-Generated<sup>1</sup> sexualized images of children** flooding social media sites, further normalizing child sexual abuse
- **AI-Generated CSAM** exacerbating the existing crisis of online child sexual exploitation, making the task of identifying real child victims even more daunting.

As the enclosed rapid assessment report documents, **the resulting problems are nothing short of a hellscape—a hellscape created by the AI sector.**

We hope, upon consideration of the information provided, OpenAI will agree that allowing AI to further fuel child sexual abuse, image-based sexual abuse, and other forms of sexual abuse, violence, and exploitation is a grave misuse of the power and promise of AI. NCOSE calls upon OpenAI, as a frontier AI developer, to recognize that its paramount responsibility is to safeguard the well-being of humanity from the unintended consequences of AI and to help rectify the harms AI has already thrust upon us by acknowledging:

- its responsibility as a frontier AI company to maintain the highest ethical standards and to protect human life and well-being
- that failure to prioritize safety within the AI sector has already unleashed a tsunami of sexual harm upon thousands of individuals, the repercussions of which will last their lifetimes
- the necessity of ensuring that AI pre-training and training data sets are thoroughly purged of *all* sexually explicit and exploitative material (i.e., hardcore pornography depicting adults, CSAM, IBSA, and AI-generated versions of such material)
- that failure to remove all sexually explicit and sexual exploitation material results in abuse-trained models that normalize sexual exploitation and further exploit the exploited
- that OpenAI's partnership with Reddit is posed to replicate mistakes akin to those of Stability AI's LAION 5b text-image model unless *very robust* measures are taken to filter out the massive amount of sexually explicit and exploitative material from its dataset
- that broad harms to public health are incurred and intensified by the widespread availability of hardcore pornography (i.e., "NSFW" material) on the Internet, and

---

<sup>1</sup> We use the term "AI-generated" broadly to mean any computer-generated images using any AI-related computer technology including machine learning, deep learning, neural networks, computer vision, and natural language processing.

Mr. Sam Altman  
Page 3  
July 30, 2024

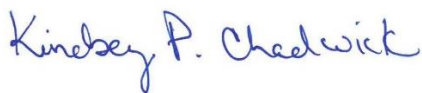
- the unqualified impossibility of the “ethical” generation of “NSFW” sexually explicit material by AI.

Further, we call on OpenAI to take the following immediate actions:

- reestablish and empower its “Superalignment” group to prioritize safety and adherence to ethical practices
- publicly release data regarding the provenance of your pre-training and training datasets and detailing your efforts to remove sexually explicit material
- fix the deficiencies of CLIP and work with users of CLIP to ensure that previous datasets created using CLIP are free from all sexually explicit material
- launch robust efforts to ensure that all its existing pre-training and training datasets are thoroughly purged of hardcore pornography, CSAM, IBSA, as well as AI-generated versions of these materials
- move immediately to ensure that any datasets obtained from Reddit are robustly vetted and purged of all sexually explicit and exploitative material
- review and update its Usage Policy and Approach to Frontier Risk to address the deficiencies as per the Rapid Assessment Report
- quickly announce a decision to **not** allow the use of any OpenAI model or tool for the generation of “NSFW” material.

A range of appalling sexual harms resulting from AI are happening now. Failure to take the steps outlined above will allow these harms to proliferate, posing severe and intolerable risks to humanity. Thus, NCOSE, survivors, our allies, and all who hope for a world free from sexual abuse and exploitation look expectantly for OpenAI to act decisively to implement these and any other measures within its powers, to mitigate future risks, ameliorate current impacts, and stop further sexual abuse and exploitation facilitated by AI. We offer this Rapid Assessment Report as part of our commitment to collaborate with OpenAI to ensure that its technologies are used responsibly and ethically. Please be aware that this report is also being released to the public. We look forward to discussing our findings and recommendations with your team at the soonest opportunity. We can be reached at [kchadwick@ncose.com](mailto:kchadwick@ncose.com) and [dawn@ncose.com](mailto:dawn@ncose.com).

Sincerely,



Kindsey Pentecost Chadwick  
Interim President



Dawn Hawkins  
Chief Executive Officer

Enclosure