Artificial Intelligence (AI) and Child Safety Online Roundtable: Insights

The risks for children online are changing rapidly, both in scale and in form.
In order to prevent those risks from turning into harm, we need the sharpest minds to be on our side.
On the side of children.[1]

HM Queen Silvia

How can we use artificial intelligence as a catalyst for child safety online?

The question was posed to 60 experts from business and academia, law enforcement agencies, regulatory bodies and child protection organisations who gathered for a global roundtable in the Royal Palace in Stockholm on November 20, 2019. The meeting was hosted by Their Majesties King Carl XVI Gustaf and Queen Silvia of Sweden and organized by World Childhood Foundation and Global Child Forum.

The roundtable was held under the Chatham House Rule, whereby participants were able to speak freely without worrying about being quoted. The aim was to equip participants with the opportunity to understand how the different stakeholder groups relate to one another and impact the issue, to learn from one another, to move towards commitments for more collaboration and joint actions, and to push the existing boundaries and develop ideas to accelerate transformative AI-driven solutions for ending online child sexual abuse.

AI techniques are advancing in line with progress in research, computer capacities, and access to large data sets. AI can help prevent online child sexual abuse, and to vastly speed up investigations in such cases. However, AI remains only one piece of the puzzle. Other technical resources – as well as human resources and offline infrastructure – all need to play their part.

Global trends

An update concerning online child sexual abuse was provided by Interpol and Ecpat International. Increases in internet access and speed have led to increases in online abuse. For example, from 2017 to 2018 there was a 541% increase in videos depicting online child sexual abuse. The severity of the abuse is also increasing: 46.6% of the photos in Interpol's database of child sexual abuse images contain an adult committing sexual assault against the child and/or instances of sadism or bestiality. More than 60% of the children depicted are prepubescent, and the younger the victim, the more severe the abuse. Enabling factors for the abusers are anonymity through encryption, and online communities where the behavior is normalized and advice for committing abuse and avoiding detection is shared.

[1] Speech by HM Queen Silvia at the AI Roundtable, The Royal Palace, Wednesday 20 November 2019
https://www.kungahuset.se/kungafamiljen/hmdrottningsilvia/tal/hmdtal/hmdrottningensavslutningstalvidairoundtablekunglslotetonsdagenden20november2019.5.4f00651a16e3a89c9d07c4f.html

Tools already exist which use Artificial Intelligence to combat child abuse. However, the offenders are still one step ahead. New approaches are needed.[2]

HM King Carl XVI Gustaf

Insights

The roundtable discussions brought forth a number of insights:

- **Technology is not the (only) barrier**

  Several AI tools to combat online child sexual abuse already exist. More are being developed.

- **AI solutions are not developed in coordination with each other**

  The solutions are fragmented and sometimes overlap

- **Existing AI tools are not known by or accessible for those who need them the most**

  Law enforcement agencies in low-income countries – where the risks for children online are increasing at an alarming pace lack access to technical tools to investigate online related crimes against children.

- **Big tech companies at the forefront of development – but accountability is lacking**

  A few global tech companies have the resources and know-how to develop AI solutions to combat online child sexual abuse – and they are at the forefront of doing so. However, this is made on a voluntary basis and with the company's best interest put first. There is a regulatory vacuum that leaves global tech companies without clear accountability.

- **The privacy issue – a huge barrier**

  It is now technically possible to identify, block or remove child sexual abuse material, suspected grooming conversations, and live streaming abuse. However, legitimate concerns regarding privacy of users often act as a barrier, obstructing tools and efforts to remove or flag child sexual abuse material, flag and investigate illegal contacts with children etc. There is a need for increased awareness and responsibility from tech companies and users that privacy has to be balanced with measures to protect children online.

[2] Speech by HM King Carl XVI Gustaf at the AI Roundtable, The Royal Palace, Wednesday 20 November 2019 https://www.kungahuset.se/royalcourt/royalfamily/hmkingcarlxvigustaf/speeches/hmspeeches/hmthekingswelcomespeechatairoundtabletheroyalpalacewednesday20november2019.5.4f00651a16e3a89c9d07995.html

- **Increased use of end-to-end encryption will cripple existing tools and efforts to protect children online**

  The vast majority of the millions of cases of suspected child sexual abuse reported to national and international law enforcement are found on non-encrypted platforms and chat services. If those platforms are encrypted, more abusers will easily be able to share material and hide their abuse.

- **There is a huge potential in involving cutting-edge AI expertise in the fight against child sexual abuse online including policy discussions**

  AI experts from various fields expressed interest in involving the latest technical advancements towards combating child sexual abuse. Child safety online needs to be raised as a relevant field for AI research and investment. New technical solutions should be developed with a child safety perspective. Law enforcement and child protection experts should try to involve an even wider group of AI experts in addressing the challenges we are facing. And the use of AI goes beyond image classification. It can and should involve other areas of AI, applied to for example research into how potential perpetrators can be prevented from committing abuse.

- **The "radioactivity" of child sexual abuse material**

  In order to develop and improve AI solutions which identify child sexual abuse images – and thereby victims and perpetrators – tech companies need access to material that AI can learn from. But since child sexual abuse images in most jurisdictions are illegal to handle, there are limited possibilities to develop solutions. It also increases the risk of silos and duplication of efforts. This has also led to isolated national databases and national classifications. Coordinated efforts are needed to solve this obstacle.

- **Inherent bias**

  The AI tools now developed are based on material mainly of Caucasian children and adults. It is unknown how this might affect the AI tools' capabilities of recognizing children and adults who are of a different ethnicity. I.e. there is a risk of an inherent bias which might result in AI tools not detecting and identifying children being abused.

- **Information for children and parents about risks online remains important**

  Evidence-informed risk awareness initiatives among children and parents are still needed both online and offline. Strengthening children´s resilience is still a key strategy in low-income countries where broadband expansion risks vastly increasing the number of child sexual abuse cases.

- **It´s about children**

  Professionals with in-depth knowledge about trauma and children's behavior online need to be involved in the development of new tools.



Tools presented during the roundtable

The Arachnid Crawler
A webcrawler that automatically finds and removes online child sexual abuse material
https://projectarachnid.ca/en/

Microsoft's Artemis
A real time grooming identification tool
https://blogs.microsoft.com/on-the-issues/2020/01/09/artemis-online-grooming-detection/

Griffeye Brain
A child sexual abuse material classifier
www.griffeye.com/multi-object-detection-now-in-griffeye-brain/

Google Content Safety API
A toolkit to increase the capacity to review child sexual abuse content
https://www.blog.google/around-the-globe/google-europe/using-ai-help-organizations-detect-and-report-child-sexual-abuse-material-online/

Thorn's Spotlight
A cloud-based data-collection and analysis service used by law enforcement in child sex trafficking cases https://www.thorn.org/spotlight/ and

Thorn's Safer
a solution for detecting, reviewing, removing and reporting child sexual abuse material
https://getsafer.io/

SomeBuddy
AI-assisted legal service for the victims of online abuse
https://somebuddy.com/

The AviaTor Project
Augmented Visual Intelligence and Targeted Online Research by Web-IQ and Ziuz
http://aviator.isfp.eu/