



## **Input for report to the United Nations General Assembly: UNICRI Centre for AI and Robotics – Existing and Emerging Sexually Exploitative Practices against Children in the Digital Environment**

1. Reference is made to the United Nations Interregional Crime and Justice Research Institute (UNICRI) and its Centre for Artificial Intelligence and Robotics (the ‘AI Centre’) established in The Hague, The Netherlands, in 2017. The AI Centre’s mission is to advance understanding of artificial intelligence (AI) and related technology. One of its largest initiatives, *AI for Safer Children*, focuses on the crime of child sexual exploitation and abuse, and how technology such as AI can support law enforcement agencies in combatting it. This focus ensures that all the initiative’s activities relate to existing and emerging sexually exploitative practices against children in the digital environment. The following two paragraphs will outline this and another relevant project of the Centre, from which the insights of this report are drawn.
2. The *AI for Safer Children* initiative was launched in 2020 as a joint partnership with the Ministry of Interior of the United Arab Emirates, and aims to build the capacities of law enforcement worldwide to leverage the positive potential of AI to support investigations into the specific crime area of child sexual exploitation and abuse.<sup>1</sup> It does so through an online platform (hereinafter ‘the Global Hub’) with a collection of information on AI tools and their responsible implementation, along with specialized training adapting this information to global law enforcement.
3. A second pertinent initiative concerns guidance for law enforcement on responsible AI innovation by developing practical resources designed to operationalize ethical and human rights principles. Entitled the Toolkit for Responsible AI Innovation in Law Enforcement (hereinafter the ‘AI Toolkit’), these resources were launched in June 2023, in partnership with INTERPOL and with financial support from the European Union. A revised and interactive web version was released in February 2024.<sup>2</sup> The AI Toolkit promotes practical knowledge to foster the responsible development, procurement and use of AI as a crucial component not only to uphold ethical and legal principles, but also for it to have the desired positive impact on the communities whom law enforcement serves.

### **How technologies are used to facilitate the sexual exploitation and abuse of children and resources addressing them**

4. The activities of the *AI for Safer Children* initiative involve identifying emerging trends in this area, some prominent ones of which are listed below:
  - a. **Self-generated child sexual abuse material** is a rapidly growing phenomenon, ranging from consensual sexting among peers to coercive grooming by a stranger online. In 2023, 64 per cent of actioned reports contained self-generated imagery (this

---

<sup>1</sup> *AI for Safer Children* initiative, UNICRI and Ministry of Interior of the UAE.

<sup>2</sup> *Toolkit for Responsible AI Innovation in Law Enforcement* (hereafter the ‘*AI Toolkit*’), UNICRI and INTERPOL.

number rises to 94 per cent for reports involving girls), compared to 40 per cent in 2022 (or 80 per cent for girls, when broken down by gender).<sup>3</sup>

- b. **Live-streamed child sexual abuse material** is another prominent trend made possible by increasing connectivity and access to devices worldwide.<sup>4</sup>
- c. **Sexually coerced extortion**, or “sextortion”, a type of blackmail where someone tries to use intimate, naked or sexual photos or videos of a person to blackmail or extort them, often for either more images or for money, has also risen alarmingly by at least 463 per cent over 2021.<sup>5</sup>
- d. **AI-generated child sexual abuse material** is one of the most worrying recent trends in this field, both given the scale and speed at which child sexual abuse material can be created. Thus, the following section will focus on this trend.

### **AI-generated child sexual abuse material: how the technology is misused, challenges and mitigation measures**

5. Freely available and easy-to-use AI software is being used by offenders to create child sexual abuse material. Open-source AI models can be downloaded from dark web forums to generate sexual abuse images of known victims or to create new content with fake faces of children.

This can be done through several techniques:

- a. Text-to-Image (or “txt2img”), meaning that users simply type in a description of the image they wish the AI to produce. The text provided is referred to as a “prompt” and can include styles, subjects, and other details.
- b. Image-to-image (or “img2img”), where images and text prompts are used as input or a basis to create a new, AI-generated image; or
- c. Inpainting, an image-to-image technique that allows users to dictate the specific areas of a real image they wish to change, down to the finest detail.<sup>6</sup>

These techniques lower the barrier to entry significantly, as increasingly little technical knowledge is required to use these tools.

6. The unlimited creation of AI-generated child abuse material hinders law enforcement investigations as already overburdened officers do not know how to distinguish AI-generated materials from real imagery. Indeed, not only the number of materials, but their advanced creation and detailing is making the distinction nearly impossible. This is an issue for several reasons:
  - a. The sheer volume of child sexual abuse imagery that can now be mass generated almost instantaneously creates a lot of noise and increases the number of files investigators must sift through exponentially. This also drives down the prices of child sexual abuse material, possibly leading offenders to produce more.

---

<sup>3</sup> [IWF Annual Report 2023](#); see also research by [Thorn](#).

<sup>4</sup> See the [Scale of Harm prevalence study](#), p. 15, based on information from the International Justice Mission (IJM), INTERPOL, Europol and WeProtect Global Alliance.

<sup>5</sup> [ActiveFence 2023](#); see also [IWF Annual Report 2023](#).

<sup>6</sup> From UNICRI’s background information for the Hackathon for Good 2023.

- b. Law enforcement is not sure whether they are viewing real victims or fake ones, or whether the image or likeness of a real victim was used and simply tampered with through AI systems.
  - c. If law enforcement investigators do find real victims, they are not sure whether other elements, for instance the places or acts portrayed, are real.
  - d. The workload in turn leads to another overwhelming issue – reviewing all the files in child abuse investigations takes a toll on both the physical and mental capacity of investigators, especially considering that the default approach to verify if an image is fake or not is by reviewing them manually, one file at a time.
  - e. A final issue stemming from AI-generated child abuse material that has arisen is the so-called “liar’s dividend”, where malicious actors benefit from the existence of fake content, such as deepfakes or AI-generated media, as it allows them to discredit genuine evidence.
7. While there is currently no reliable tool to detect deepfakes, innovative approaches and tools are being piloted by technology companies, such as checking file source information (based on the camera model, GPS data, etc.), deep fake detectors for faces or for other identifiers,<sup>7</sup> and one of the latest - a “FakeCatcher” working with blood flow and pixels correlation to distinguish real videos.<sup>8</sup>
8. On the part of law enforcement, there are other software tools that can help alleviate the amount of files needed to review. Contributing to and using hash databases builds a shared collection of known child sexual abuse material – including labels, for instance of the files which have been identified as AI-generated – which investigators can immediately flag or disregard.<sup>9</sup> Using even one hash database can reduce the amount of files to review by 50 per cent. If the victim or suspect is known, identification tools can further classify which images are most likely to contain them.

**Challenges and solutions for enhancing the prevention, detection, reporting and protection of children from sexual exploitation and abuse online**

9. A large challenge for law enforcement is a simple lack of knowledge of tools available to help them deal with child sexual abuse material. The over 80 AI tools displayed on the AI for Safer Children Global Hub have significant potential to help fight the increasing scale of online sexual abuse of children through a wide range of features applicable at different stages of an investigation workflow:
- a. **Preventative technologies:** AI tools can help flag content likely to contain child sexual abuse material, either after it is uploaded to a platform, to facilitate takedown measures, or even before young people or potential abusers share files – along with a link to a helpline.<sup>10</sup> Other possibilities presented by AI tools include, *inter alia*, the ability to identify the age of users on a platform or identify a suspect’s online fingerprint to prevent them from logging on to harmful sites.

---

<sup>7</sup> Sadly, this technique is rarely applicable to child abuse cases (example <https://www.duckduckgoose.ai/>).

<sup>8</sup> Read more here by [INTEL](#).

<sup>9</sup> [Technology Coalition resources on hash databases](#).

<sup>10</sup> [Chatbot Prevents Millions from Searching for CSAM](#).

- b. **Content generation:** AI systems enable the generation of new content such as images or text. As seen, this can be used to generate child sexual abuse materials but it can also be leveraged for good, by supporting law enforcement agencies with undercover operations such as creating fake online personas to infiltrate a criminal network.<sup>11</sup>
  - c. **Text and speech analysis:** AI systems can be used to recognize, process, tag and extract meaningful information from texts and audio recordings. This can be used by law enforcement officers for, for instance, chat categorization to swiftly identify and flag suspicious messages. Other applications that can support investigators consist of proactively monitoring the dark web or social media platforms to flag language suggesting grooming, luring or other possible child abuse activities. It can also automatically fill in reports to expedite legal processes such as subpoenas to internet service providers and speed up the removal of child abuse material from their platforms.
  - d. **Image analysis:** AI systems can be used to analyze photos and videos to automatically recognize, classify, and contextualize an image or elements within that image. This can be used, for instance, to support the identification of suspects, victims and locations of child sexual abuse. AI can also be helpful for triage during search and seizures by cutting through the vast amount of data to determine relevant digital devices, and in more thorough identification in digital forensic labs.<sup>12</sup>
  - e. **Process and workflow optimization:** another benefit of AI systems' ability to analyze large data sets is to identify anomalies and patterns and suggest ways to optimize and automate specific workflows. In the law enforcement context, this can be used to make connections between pieces of evidence by correlating certain events that occurred at the same time or place, activities carried out using the same devices, or other similarities. This can be beneficial, for instance, in grouping images with the same creation date or similar subjects together to identify a series – which furthermore holds potential for identifying a series of AI-generated images.
10. In order to take full advantage of these tools, law enforcement personnel need proper training – firstly, to raise awareness that these solutions exist and how to implement them responsibly, as provided by the *AI for Safer Children* initiative. Law enforcement personnel also require adequate training for specific tools and their specific legal environment.
- a. One solution available to law enforcement is through the *AI for Safer Children* initiative, which provides specialized training free of charge to further advance capacities to combat child sexual exploitation and abuse online through the use of AI tools including the capabilities described in the previous paragraph.
  - b. These trainings are tailored to the specific needs and contexts of participating law enforcement agencies, and necessarily involve technology providers and other stakeholders as a multi-stakeholder approach is needed to address this transnational issue.
  - c. While our initiative works toward this capacity-building goal, more resources are required to ensure these efforts continue, reach the whole world – including less technologically advanced countries – and keep up with tech and regulatory changes.

---

<sup>11</sup> [AI Toolkit](#).

<sup>12</sup> [AI Toolkit](#).

## **Regulation and standard-setting for AI in law enforcement dealing with child abuse cases**

11. The AI Toolkit provides the most concrete guidance related to the legal and ethical use of AI, through a set of unique resources tailored to the global law enforcement community.<sup>13</sup> Collectively, they support agencies aligning their AI innovation with human rights law. The following resources are particularly relevant.
  - a. The *Principles for Responsible AI Innovation* describes the guiding principles derived from international human rights law, AI ethics and policing principles to ensure that AI systems are developed and used for the benefit of society while protecting human rights.
  - b. The *Responsible AI Innovation in Action Workbook* supports agencies in operationalizing responsible AI innovation and documenting decisions taken throughout the AI life cycle, including questionnaires to reflect on the observance of the principles at each stage.
  - c. The *Risk Assessment Questionnaire* helps agencies to identify and evaluate the risks that an AI system may pose to individuals and communities if the principles are not sufficiently fulfilled.
  
12. Since this crime type is transnational, much of the challenges of information-sharing – for instance with the enormously useful hash databases of known illegal material – arise from differences in legislation. States must therefore work on harmonizing terminology and laws in accordance with new technological developments such as AI-generated child sexual abuse material.
  - a. In the meantime, the Global Hub provides resources on worldwide legislation related to child sexual abuse material, such as the Global AI Law and Policy Tracker<sup>14</sup> and INHOPE’s Legislative Overview.<sup>15</sup>
  - b. Furthermore, a classification schema is being developed by INHOPE to build a foundation for more effective inter-jurisdictional data sharing.<sup>16</sup>

## **Practical recommendations for combatting child sexual exploitation and abuse online**

13. In summary, our practical recommendations are as follows:
  - a. Law enforcement needs proper awareness, training and tools to support the investigation of child sexual exploitation and abuse within their communities.
  - b. Industry should continue to develop AI tools for good, with responsible AI principles taken into account throughout their lifecycles.
  - c. Service providers should implement safeguards from the beginning, before photos or videos are uploaded, to ensure that they are not child sexual abuse material, and

---

<sup>13</sup> [AI Toolkit](#).

<sup>14</sup> [IAPP, Global AI Law and Policy Tracker](#).

<sup>15</sup> [INHOPE, legislative overview](#).

<sup>16</sup> [INHOPE classification schema](#).

to label content as AI-generated. Furthermore, they should have a trained team responding to prevent and take down child sexual abuse material, willing to work with law enforcement.

- d. States should update or interpret legislation in accordance with technological developments, and join efforts to harmonize child sexual abuse material classification as much as possible to facilitate multi-stakeholder and cross-border information sharing.