**Forum:** United Nations Office on Drugs and Crimes (UNODC)

**Issue**: Combating the Use of Generative AI for Illicit Activities

**Chairs:** Leyan Masri**,** Elana Tayeb, & Nada Himrane

## Introduction

AI has rapidly developed from generating images that people can easily discern from real and fake to indistinguishable from real photos. AI can do many things such as simulate voices, produce essays, and even create realistic images. While it has its applications that doesn't mean it hasn't been used for malicious reasons. AI has negatively impacted many people with crimes such as deep fakes. Given the misuse of AI, the UNODC has prioritized this issue as a global threat.  It has yet to explore the misuse of AI's impact on policy making. Delegates will aim to see how current law attempts to regulate this issue and what action can be taken against it.

## Definition of Key Terms

### Generative AI

A type of artificial intelligence that can produce content such as text, images, audio, or video based on input data. Notable examples include ChatGPT and image generators like DALL-E.

### Deepfake

A synthetic media created using AI, typically involving the manipulation of photographic media to falsely depict someone saying or doing something. Deep Fakes are generally known to be used in cybercrimes, and misinformation campaigns.

### Illicit Activities

Illegal actions which refer to activities such as fraud, identity theft, and child exploitation. In the context of AI, illicit activities involve using generated content for malicious activities.

## Background Information

AI is a versatile tool that has revolutionized global living circumstances. It has heavily impacted multiple sectors, such as entertainment and finance and many more. It has done all of this through its advanced machine learning. Being able to improve upon itself at every opportunity, it is always adapting and picking new information. While this is quite a beneficial tool which aids our society into advancing much faster, it does have its downsides. All of its vast capabilities can be misused for crime and with AI being so accessible it makes it a very dangerous tool. Recent studies have shown multiple concerns towards the possibility of generative AI enabling illicit activities. It reveals how accessible AI-powered tools are and how they can be used to commit cybercrimes. There have been multiple investigations proving how anyone can easily bypass the AI platform's safeguards. Recent studies have shown multiple concerns towards the possibility of generative AI enabling illicit activities. It reveals how accessible AI-powered tools are and how they can be used to commit cyber-crimes. Such tools are appealing to criminals, who are aware of the capabilities of AI tools, as they can use it for generating crimes such as fraudulent content, phishing schemes, deep fakes, etc. There have been multiple investigations proving how anyone can easily bypass the AI platform's safeguards. This is a concerning disadvantage which criminals can exploit to generate further malicious content such as spreading misinformation and hate speech. Government organizations as well as researchers urgently demand stricter regulation for AI tools in order to prevent the inhibited usage that serves to risk their security. In summary AI brings many positives to all aspects of life but it has its dark side with it being so easy to be weaponized by criminals posing significant risk to our society.

## Major Parties Involved

### United States

The United States is actively combatting the use of generative AI for illegal purposes through initiatives led by the Department of Justice (DOJ). The United States is actively combatting the use of generative AI for illegal purposes through initiatives led by the Department of Justice (DOJ). The initiatives include a better accurate detection and prosecution of AI-assisted crimes. Not to mention, harsher sentencing for AI-based crimes became much more serious due to its misuse. The main priorities include protecting election integrity from misinformation spread using generative AI, as well as protecting national security through the "Disruptive Technology Strike Force.";

### European Union

The issue is currently being addressed by the European Union in regards to the misuse of generative AI for malicious activities. They adopted the EU Artificial Intelligence Act (AI Act). This act seeks to establish a comprehensive legislative framework for trustworthy AI usage. The AI Act uses a risk-based approach that categorizes AI systems based on their risk levels so as to implement regulatory restrictions. High-risk AI systems must meet strict requirements such as: transparency, data governance, human monitoring, and risk management. Meanwhile, certain AI applications like social score are explicitly prohibited. This Act requires pre-market conformity evaluations as well as post-market monitoring to reduce risks to fundamental rights. The Act also provides special safeguards for general-purpose AI models, such as GPT, that address dangers specific to their widespread usage;

### China

China is taking a very active approach towards combating the misuse of generative AI for illicit activities, focusing on legislative measures to guarantee ethical deployment. The country has imposed regulations that push AI developers into avoiding the creation of harmful content, and that lines up with the overall state goals of data security and public safety. Specific approaches involve making the AI models integrate safety into their development, then implementing content moderation practices. These regulations are part of China's more broad effort to govern AI technologies and preventing risks connected with misinformation, and other potential misuse;

### Interpol

Interpol collaborates with international organizations like UNICRI to produce publications such as "Artificial Intelligence and Robotics for Law Enforcement" in efforts to counter the misuse of generative AI in illegal operations. This project points to skill gaps within law enforcement agencies and emphasizes the need for enhanced international coordination. Interpol also participates in workshops and produces technical information with a view to making sure law enforcement organizations around the world are able to efficiently predict, prevent, and respond to crimes made possible by AI technologies;

**UNICRI**

UNICRI is an important force in countering the use of generative AI for criminal purposes through programs such as the AI for Safer Children program. This effort, developed in concert with the UAE Ministry of Interior, uses AI and machine learning to identify and block online child sexual abuse material to help law enforcement track down criminals and protect children. UNICRI works to improve trust in AI technologies, combine criminal prevention and respect of individual privacy, and establish international co-action among interested stakeholders. Their work illustrates the potential of applying AI to counter emergent concerns, including the potential misuse of generative AI;

## Timeline of Key Events

| **Date** | **Description of Event** |
|---|---|
| **2020-2023** | Since 2020, the DOJ in the United States has focused on AI tools and strategies to combat crimes involving these AI technologies, such as cybercrimes, fraud, and disinformation campaigns. |
| **April 2021** | The European Union proposes its AI Act, targeting high-risk applications to mitigate criminal misuse. |
| **2018-2021** | Interpol combats the misuse of generative AI through international collaboration with organizations such as UNICRI. Additionally, they prevent it by knowledge-sharing, and |

| | |
|---|---|
| | strategic partnerships in order to equip law enforcement agencies with the necessary tools to respond to AI-driven crimes effectively. |
| **2017-2023** | China put in place comprehensive regulations related to generative AI and implementing rules that require developers to prevent harmful content |
| **November 2020** | UNICRI launched AI for Safer Children in a collaborative partnership with the UAE's Ministry of Interior to curb CSAM online using AI and machine learning. |

## Previous Attempts to Resolve this Issue

1. **Legislative Efforts**

Countries such as the United States and members of the European Union have signed legislation regarding the strict regulation of AI. The EU has proposed an AI Act which classifies generative AI in a high-risk category that requires extensive supervision;

2. **International Cooperation**

UNICRI and Interpol aim for international cooperation. for the purpose of fully tackling the misuse of AI across borders, they placed a strong emphasis on information exchange. Despite previous attempts, this solution appeared to be inadequate. There are multiple limitations preventing its implementation such as the lack of robust laws and the disintegration of regulatory standards;

3. **Awareness and Education Campaigns**

   Government organizations collaborating with tech companies have initiated Public Awareness campaigns so as to inform the public of scams involving artificial intelligence. They aimed to reduce the impact of social engineering tactics that exploit generative AI;

## Possible Solutions to Resolve this Issue

1. **Establish international standards for Generative AI use**

   Attempt to push for UN member states to create global standards for AI use. These standards will ensure the ethical and productive use of AI and deter any misuse;

2. **Enhanced Cross-Border Enforcement Mechanism**

   Create a task force in the UNODC to reduce generative AI crimes and gather data on said crime that can be shared globally to aid in investigation efficiency;

3. **Promote AI Transparency Requirements**

   Push to policies to require AI developers to watermark any AI generated media to ensure misinformation is not spread such as deep fakes;

4. **Develop Specialized AI Crime Unit**

Encourage governments to create teams to deal with generative AI crimes. These teams will consist of highly trained operatives to prevent crimes such as identity theft and exploitation.

## Bibliography

Hacker, P., Engel, A., & Mauer, M. (2023). Regulating ChatGPT and other Large Generative AI Models. *2023 ACM Conference on Fairness, Accountability, and Transparency*, 1112–1123. http://dx.doi.org/10.1145/3593013.3594067

Katieal. (2024). Impact of artificial intelligence on criminal and illicit activities. U.S. Department of Homeland Security. Retrieved from https://www.dhs.gov/sites/default/files/2024-10/24_0927_ia_aep-impact-ai-on-criminal-anD-illicit-activities.pdf

Kumta, A. (n.d.). *DOJ announces initiative to combat ai-assisted crime*. Compliance and Enforcement. https://wp.nyu.edu/compliance_enforcement/2024/02/27/doj-announces-initiative-to-combat-ai-assisted-crime/

National Counterterrorism Center. (2023). First Responders Toolbox: Violent Extremists' Use of Generative Artificial Intelligence. https://www.dni.gov/files/NCTC/documents/jcat/firstresponderstoolbox/151s_First_Respo nders_Toolbox-Violent_Extremists_Use_of_Generative_Artificial_Intelligence.pdf

Research Institute (UNICRI). (2022). Generative AI: A Threat to Child Sexual Exploitation and Abuse. https://unicri.it/Publication-Generative-AI-Threat-Child-Sexual-Exploitation-Abuse

Velasco, C. (2022). Cybercrime and Artificial Intelligence. An overview of the work of international organizations on criminal justice and the international applicable instruments. *ERA Forum*, *23*(1), 109–126. https://doi.org/10.1007/s12027-022-00702-z