

Committee: ECOFIN

Topic: The Question of Combatting AI Driven Fraud

Chair: Zoë Spellings

School: Royal Russell School

Summary

The Question of Combatting AI Driven Fraud is a growing concern globally. As Artificial Intelligence becomes more advanced and accessible, as do the methods used to deceive individuals, corporations and even governments. Deepfakes, AI-powered bots, phishing, and voice or identity cloning are all methods used to commit fraud on a spectrum of scales. Some argue that the solution is to continue to use AI, but for sophisticated detection systems that recognise such fraud. Others are focused on regulation and transparency to keep up with the increase in threats. However, member states aren't equally equipped, educated or willing to implement technologies or legal frameworks to combat AI fraud, creating a key imbalance for the issue. This has led to growing pressure for cooperation on shared frameworks to prevent the misuse of AI.

Definition of Key Terms

Artificial Intelligence (AI) – an umbrella term for a range of technologies and approaches that often attempt to mimic human thought to solve complex tasks.

Fraud - the intentional deception or trickery to deprive someone of a legal right or to gain something of value unfairly, such as money or private information.

Deepfake – a video, photo, or audio recording that seems real but has been manipulated with AI.

Phishing - a type of cyberattack and social engineering where criminals impersonate a trusted entity, such as a bank or company, to trick individuals into revealing sensitive information like passwords, bank details, or credit card numbers, or into installing malware.

Authentication - the process of verifying that an individual or entity is who or what they claim to be, often using credentials like passwords, security tokens, or biometric data.

Cybercrime - any criminal activity conducted using a computer, computer network, or internet-enabled device, aiming to steal information, cause financial or reputational harm, or disrupt services.

Background Information

(Overview)

Over time, fraud has evolved alongside technology and commerce, and the 21st century has seen the rise of such activity through the use of artificial intelligence. Deepfake technology, synthetic identity generation and phishing systems are the face of the new generation of fraud tactics. Unlike traditional fraudulent methods, AI allows criminals to act fast and, in some cases, globally.

The response to AI-driven fraud is divided. Developed nations, those housing advanced digital infrastructures, have started deep investments into AI-powered tools across a multitude of industries ranging from e-commerce to agriculture. With this, comes large scale security measures, to protect citizens as well as keeping trust in their heavily automated environment. However, developing countries lack key financial, technological and legal frameworks to counter increasingly sophisticated attacks, leaving smaller business vulnerable. The large disparity in the accessibility of the necessary security measures which come with the growth of AI, highlights a growing divide where the costs of such fraud cases are targeted to those who are least equipped to tackle them.

The nature of artificial intelligence raises significant challenges for international co-operation. Initiatives led by INTERPOL and the Financial Action Task Force (FATF) have made key steps in encouraging positive collaboration and solutions to this ever-growing issue, but international laws remain slow in comparison to the rapid innovation of these malicious AI systems. Moreover, ethical concerns surrounding AI developments further complicate this issue, as the tools which are advancing society, are also being used to exploit it.

Combatting AI-driven fraud requires both internal strategies and the creation of comprehensive international frameworks to enhance the security elements of our advancing systems, ensuring that the technology that serves us, also protects us.

Major Countries and Organizations Involved

China

China has introduced some of the most direct regulations on AI generated content like deepfakes and synthetic media being required to be labelled and watermarked; this reduces opportunities for AI to be used in a fraudulent manner. They also actively participate in UN cybercrime negotiations and pushing for global rules that align with their domestic approach, highlighting their recognition of AI-driven fraud as both a national and global issue.

United States of America

The US actively addresses AI driven fraud through federal action and innovation. The FTC and DOJ both use AI tools to flag scams involving deepfakes as well as The AI Safety Institute being launched to monitor missed AI systems. Despite the US not having a national AI law, regulations and state laws allow them to stay a key leader in AI safety and enforcement.

International Telecommunication Unit (ITU)

They play a major role in coordinating international standards for AI, and overall digital technologies. Through the annual AI for Good Summit and promotion of technical solutions for this issue such as watermarking and verification, the ITU is keen on the creation of a global framework for protection against AI- driven fraud.

UNICRI Centre for AI & Robotics

This centre was established in 2017 to provide support for member states in both understanding and governing AI threats such as fraud. The UNICRI conducts research, provides training programs and develops AI tools for law enforcement. Some of their focus includes the prevention of financial corruption, deepfake fraud operations, building international policy consensus and designing AI based tool to identify and report content in support of fraud investigations.

European Union

The EU has taken a strong stance on this issue, particularly with their adoption on the EU AI Act in 2024. The EU also heavily invests in research on digital identity and AI detection systems, aiming to safeguard both citizens and businesses. By mixing regulation with technical innovation, the EU has become a model in managing AI risks.

Timeline of Events

Date	Description
2001	Budapest Convention on Cybercrime-The council of Europe adopts the first international treaty on cybercrime, laying the foundations for global cooperation on digital crimes
2017	The emergency of deepfake technology- the first convincing deepfakes are shared online
2018	The first documented AI voice fraud- A UK based energy firm loses £220,000 after scammers use AI voice-cloning to mimic the company's CEO and order a fraudulent transfer
2019	UNICRI publishes research on AI and crime- the UNICRI begins warnings about the misuse of AI in financial fraud
2020	AI powered phishing scams emerge-cybersecurity firms report of AI used to automate phishing attacks
2022	INTERPOL issues warnings of rising cases in AI fraud-they call for stronger cross border collaboration
2023	Global reports of voice cloning scams- criminals impersonate relatives or executives to demand urgent payments, losses rise into the hundreds of millions
2024	The UN adopts its first cybercrime treaty-this specifically covers digital fraud and gives legal framework for dealing with Ai enhanced scams
2025	ITU reports on AI fraud risks-at the 'AI for good' summit, the ITU highlights the increasing threat of AI-driven fraud

Relevant UN Treaties and Events

UN Cybercrime Convention - 2024

This was adopted in December 2024 after 5 years of negotiations, becoming the first global and legally binding treaty that directly addresses cybercrime, such as online fraud. It lays out frameworks for international cooperation, evidence sharing and capacity building. This allows for a common legal ground for member states to prosecute AI-driven scams such as identity theft.

Framework Convention on Artificial Intelligence and Human Rights. Democracy and Rule of Law - 2024

Ensuring that AI systems respect human rights, democracy and the rule of law, this framework aims to prevent the abuse of readily available AI systems, including situations of fraud. It sets standards of transparency and accountability; it focuses on the governance and ethical side of AI-driven fraud.

AI for Good Global Summit

This UN-led event organised by the ITU focuses on the promotion of the responsible, and positive use of artificial intelligence. As of 2025, the ITU released a report urging for stronger measures to detect AI generated deepfakes, directly contributing to efforts against AI-driven fraud.

UN Secretary-General's Advisory Panel on Global AI Governance

The panel was established to guide the international regulations of AI. Global experts shape recommendations of AI risks, such as fraud. It continues to contribute towards global awareness and framework direction for preventing AI abuse in fraudulent schemes.

Previous Attempts to solve the Issue

Budapest Convention on Cybercrime – 2001

This was the first international treaty addressing crimes committed via computer networks. It harmonised laws, created investigation standards and promoted international co-operation against offences such as fraud. Although this treaty predates modern AI technology, it laid the

foundation for international responses to digital fraud and established principles which remain relevant in AI-driven crimes.

Establishment of National AI Strategies – 2017 onwards

From 2017, many member nations developed national AI strategies. These often focused on the ethical use of upcoming AI systems, and the risks of their misuse such as fraud. Although their impact and severity vary widely, they represent early and continual efforts to incorporate fraud prevention into AI governance.

Creation of the UNICRI Centre for AI and Robotics – 2019

Opened in the Hague, this centre focuses on the study of the relationship between AI, robotics and crime prevention. This centre highlighted the risks of malicious AI use and provided member states guidance on prevention and security strategies. Its work shows a massive attempt to prepare for the wave of AI-driven fraud that has since emerged.

Possible Solutions

International Legal Frameworks

By expanding treaties to include direct and specific provisions against AI-driven fraud (deepfakes, identity theft, automated phishing), this will allow for a clear and harmonised legal standard across member nations, enabling law enforcement to efficiently co-operate when such fraud takes place across borders.

AI detection tools

By utilising current AI systems or introducing the development of new systems to specifically tackle fraud by using multi-factor authentication, biometric verification or more real time monitoring, this will restrict opportunities for users to commit fraud, even with sophisticated AI tools.

Stronger Financial Security Systems

The finance sector is one of the largest targets for fraud, especially with the increasing use of AI tools. By introducing stronger measures such as blockchains and authentication, the systems

become more resilient. These tools strengthen overall financial security by protecting assets, safeguarding transactions and reinforcing trust in these systems.

Public Awareness and Education Campaigns

The launch of UN led or national public campaigns through schools or social media ads can educate people about the danger of AI tools and how they can be used in fraud. These could highlight common tactics, give advice on how to report incidents and how to keep your data safe. By raising literacy on this, the public will become more resilient against fraud attempts and in turn reducing the success rate of such scams.

Bibliography

https://unicri.org/in_focus/on/unicri_centre_artificial_robotics

<https://www.reuters.com/business/un-report-urges-stronger-measures-detect-ai-driven-deepfakes-2025-07-11/>

<https://www.unesco.org/en/artificial-intelligence/recommendation-ethics>

<https://www.weforum.org/stories/2025/01/how-ai-driven-fraud-challenges-the-global-economy-and-ways-to-combat-it/>

<https://www.sciencedirect.com/science/article/pii/S2949791424000435>