

October 2025

# Limiting Nefarious Exploitation: How to Mitigate Criminal and Terrorist Exploitation of Artificial Intelligence

Dr. Inga Trauthig

---

## Executive Summary

Earlier this year, Eric Schmidt, former CEO of Google, [contemplated](#) that he dreads the “Osama bin Laden of AI.” He joined a list of dignitaries expressing similar fears of groundbreaking technologies exacerbating security threats. This angst is not unfounded. Terrorists are under constant [pressure to innovate](#) given the nature of their activity and goals. They have aptly [exploited](#) emerging technologies from the beginnings of social media in the early 2000s to relying on decentralized technologies like cryptocurrencies for financing or decentralized hosting services to avoid content removals. Similar concerns apply to criminals. This paper argues that limiting malevolent technological exploitation is paramount for safeguarding modern societies. Six avenues should be addressed when aiming to reign in these threats to national security.

- a. Utilization of existing efforts in cybersecurity,
- b. Collaboration and information sharing,
- c. Streamlined regulation,
- d. Continuous trainings of specialists to harness AI to combat AI threats,

- e. Awareness campaigns and training of the public,
- f. Disengagement and prevention programs.

## 1. BACKGROUND

Incorporating Artificial Intelligence (AI), including generative AI, into business, [government](#) and research practices is considered paramount in 2025. This is because AI has the potential to facilitate, automate and scale – overall increasing efficiency significantly. Further, AI can free researchers from miniscule tasks and allow for more time on challenging projects. So far, so utopian. [Advocates](#) of this ideal, such as the [Department of Government Efficiency](#) (DOGE) or consultants advising business on AI integration, are stumbling over roadblocks that have accompanied previous fundamental technological shifts. Amongst them are societal resistance to change, ill-preparedness for complications of still emerging technologies, ongoing questions around effectiveness or hallucinations and the consistent challenge of proofing systems against nefarious exploitation. The latter includes the adaptation of emerging technologies, such as

AI, by malevolent actors threatening societal peace, wealth and national security.

This policy paper outlines policy solutions to malevolent AI exploitation. This is salient as bad actors can jeopardize the positive societal developments, such as the proper embrace of AI as well as directly threaten citizens' safety when employing AI to harm.

For terrorism, AI acts as an [exacerbator](#) of the exploitation of several emerging technologies – most pertinently (1) Social media and virtual reality, (2) Unmanned aerial vehicles (UAVs or drones), and (3) Autonomous cars. For (1) AI can [contribute](#) to creating *more* terrorist propaganda *faster* in a variety of languages and dialects. Further, this propaganda and correlating recruitment tactics could be more [engaging](#) when [incorporating generative AI](#). By now, transnational Salafi-jihadis (like the Islamic State or ISIS) have [released](#) guidelines on how to securely embed AI to develop propaganda and right wing extremists have [underlined](#) the importance to include AI into their “memetic warfare.” For (2) UAVs have been [described](#) as one of the [primary terrorist threats](#) by the United Nations Security Council Counter-Terrorism Committee. The dangers are directly associated with accessibility reductions – not only resource-rich terrorist organizations can afford to [innovate](#) anymore. For (3) Hacking into autonomous vehicles as well as turning those vehicles into weapons (self-driven vehicular bombs) can be supported by some [elements of AI](#).

For transnational crime, AI acts as a key battleground for the continuous attempts of criminal groups to outmaneuver law

enforcement authorities – especially with regard to (1) Economic cybercrime, such as online financial crimes, (2) virtual reality and generative AI, and (3) future border control. For (1) AI is being [fought over](#) as panacea or accelerator of financial crimes. While [utilizing](#) AI as antidotes for cybercrime, public and private actors need to keep in mind potentially unintended, harmful spillovers to society inherent when employing emergent technologies that are not fully tested (yet). For (2) generative AI specifically allows for microtargeting of victims. Materials, such as deepfakes are [utilized](#) to [defraud](#) individuals and companies or harass politicians and civil society who try to push back against organized. For (3) future border control and hence the fight against illegal migration and trafficking, AI is [considered](#) a potentially groundbreaking technology. Within this quest for advancements private actors have been taking on state tasks, which [carries](#) ethical implications.

## 2. MAIN ARGUMENT

The exploitation of AI by criminal or terrorist actors can overlap just as terrorism and crime often overlap in practice. Overall, criminal organizations are free of ideological guardrails that can [inhibit](#) technological adaptations by terrorist groups. This brief calls for a holistic approach to mitigate malevolent exploitation. As these groups often rely on easily available commercial technologies, corporate buy-in to safe design and public-private partnerships are a necessity to alleviate societal harms. Overall, crime and terrorism are social problems for which the mitigation of technological exploitation is just one

avenue to reign in threats that have deep-rooted, underlying drivers. Still, limiting technological exploitation is a main avenue for safeguarding modern societies.

### 3. POLICY CONTEXT

US policy makers are concerned about existing and future threats related to criminal or terrorist exploitation of AI. They are joined in these worries by policy makers around the world who have witnessed local and international groups adopting emerging technologies and threatening their societies. In March 2025, the executive director of Europol, the European Union’s law enforcement agency, [claimed](#) that “Cybercrime is evolving into a digital arms race targeting governments, businesses and individuals. AI-driven attacks are becoming more precise and devastating.” Europol’s annual threat assessment further [added](#) that “Hybrid and traditional cybercrime actors will increasingly be intertwined, with state-sponsored actors masking themselves as cybercriminals to conceal their origin and real disruption motives.”

Just one month before, August Pfluger (R-TX), chairman of the Subcommittee on Counterterrorism and Intelligence, introduced legislation (the “[Generative AI Terrorism Risk Assessment Act](#)”) to examine how foreign terrorist organizations use the internet and emerging technology to recruit, radicalize, and inspire attacks in the United States. And shortly after, in April, the European Union introduced a comprehensive legal framework on AI (the “[Artificial Intelligence Act](#)”) which [addresses](#) threats related to several AI systems while also aiming to promote

innovation. In sum, these threat assessments and legislative efforts on both sides of the Atlantic underscore the urgency and complexity of mitigating malevolent AI exploitation.

In general, policy dialogues focus more on reigning in the technology, rather than access or punishing malevolent users. [This is similar](#) to earlier (and existing) legislation that regulates social media. While some regulators increasingly try to hold platform owners or developers to account, malevolent users are less in focus. This is underlined in a way as the onus on reigning in malevolent users has been delegated mainly to platforms.

In the United States no comprehensive, national AI law [exists](#) but several states have been passing their own AI laws, leading to a fractured regulatory landscape. On the federal level, President Donald Trump revoked an existing executive order on AI by the Biden legislation and issued his own executive order shortly after taking office in January 2025 instead. The spirit of Executive [Order 14179](#) is captured in its title: “Removing Barriers to American Leadership in Artificial Intelligence” – signaling a shift toward AI deregulation and industry-led innovation and self-imposed AI governance as the order outlines. In contrast to the EU, the Trump administration is not interested in potentially limiting American AI companies’ growth through regulation – this was underlined in the [AI Action Plan](#) of the White House from July 2025. However, some federal AI-related frameworks continue to exist – as voluntary efforts, such as the Department of Homeland Security’s “[Roles and Responsibilities Framework for](#)

[Artificial Intelligence in Critical Infrastructure](#)” which provides guidance for AI deployment in key sectors.

#### 4. POLICY RECOMMENDATIONS

Malevolent exploitation of AI is a social problem at heart accelerated by technological affordances. As such, six avenues should be addressed when aiming to reign in these threats to national security.

- a. Utilization of existing efforts in cybersecurity,
- b. Collaboration and information sharing,
- c. Streamlined regulation,
- d. Continuous trainings of specialists to harness AI to combat AI threats,
- e. Awareness campaigns and training of the public,
- f. Disengagement and prevention programs.

To utilize existing efforts in cybersecurity AI-specific trainings, measurements and operations should capitalize on programs that were developed over the last decades. Sometimes this can be as straightforward as translating existing principles into specific, operationalized standards for AI products and procurement that can be used by agency leaders, policymakers, and community groups alike. Further, relevant high-quality research should guide the field at large to ensure that current efforts do not just replicate but enhance existing achievements.

For better collaboration and information sharing, sensible safeguarding protocols for said (potentially critical national security) information and standardized procedures are paramount. As the DHS [explains](#) “While no

one disputes the importance of collaboration, the ‘how to’ is a perennial hurdle when the government and the private sector are concerned.” To mitigate these pitfalls and expand public-private partnerships, simplification, standardization and longevity (versus ad-hoc attempts) of collaborations are critical. The Joint Cyber Defense Collaborative (JCDC) that is part of the Cybersecurity and Infrastructure Security Agency (CISA) is crucial in this regard and its role should be strengthened. Next to public-private partnerships, industry-internal (private-private partnerships), however, should be pursued with added emphasis to allow for individual company threat intelligence insights to be shared – ideally before similar exploitations occur at a different company. These efforts regarding AI can be incorporated into existing frameworks, such as the Christchurch Call Foundation or the Global Internet Forum to Counter Terrorism (GIFCT).

The call for streamlined regulation mainly stems from the reality that a patchwork of regulations currently characterizes the United States. By the end of last year, over a hundred bills [had been attempted](#) at state legislatures leading to around forty new laws. Lawmakers were mainly concerned with criminalizing AI-generated explicit content and protecting individuals (especially minors) from digital exploitation. In general, for beneficial, trustworthy AI to thrive, it [requires](#) expanding legal frameworks to address the specific features of AI-facilitated manipulation. These steps also facilitate the development of consensus principles to guide the safe, ethical, and effective use of AI.

In the face of continuous AI developments, continuous trainings of specialists to harness AI to combat AI threats is paramount. In short, continued federal and state investments in cybersecurity workforce development at public institutions (including law enforcement agencies) and academic institutions that address effective and ethical uses of AI technologies and prop up defenses. These trainings are paramount because of the likelihood that threat actors will not just go for an AI system but more likely the hardware, software, networks, and people. In terms of future threat prevention, some evidence exists that highlights the use of machine learning to predict mafia infiltration. This type of AI counter measures highlights the prevention rather than countering aspect which AI can nurture.

Widespread awareness campaigns and AI-focused trainings of the public are a vital part in developing resilience for low-key attacks and cybercrime targeting. While Americans are widely aware of AI at least to some degree, specific use cases and potential threats are not commonly understood. Further, in September this year Pew Research found out that half of the American public is “much more concerned than excited” about the increased use of AI in daily life. With regard to one of the main tactics of AI deception—deepfakes—76% of Americans feel strongly that it’s important to be able to tell if pictures, videos or text were made by AI or by humans. However, more than half do not trust their own competence to spot AI-generated content.

Turning the spotlight to the perpetrators of nefarious AI exploitation in focus for this paper, disengagement and prevention

programs can address the root causes of terrorism and organized crime. Exploring how AI can potentially be used in these prevention and countering violence extremism programs should be explored in parallel to AI’s potential in counterterrorism. Keeping these structural societal dynamics in mind—next to focusing on AI-specific increases in defense against nefarious exploitation—allows for a more long-term perspective of countering AI exploitation by criminals and terrorists.

To sum up, these six steps are considered the best strategy to safeguard the public and ensure that new developments in the field of AI provide a net good to society. They are led by practical needs while also factoring in human rights and liberal principles guiding any policy in the United States.

## **Author**

Dr. Inga Trauthig is a research professor at the Jack D. Gordon Institute for Public Policy at Florida International University. She is a security studies scholar and received her Ph.D. from the Department of War Studies at King's College London. She has authored over 80 publications, regularly consults with policy makers and strives to contribute to public awareness on the malevolent exploitation of emerging technologies.

## **Disclaimer**

This publication is part of the Jack D. Gordon Institute for Public Policy at FIU's Policy Innovation Series edited by Dr. Inga Trauthig. The views expressed in this publication are solely those of the author/s and do not necessarily reflect the official policies or positions of the Florida International University, the Jack D. Gordon Institute for Public Policy, or any affiliated institutions.