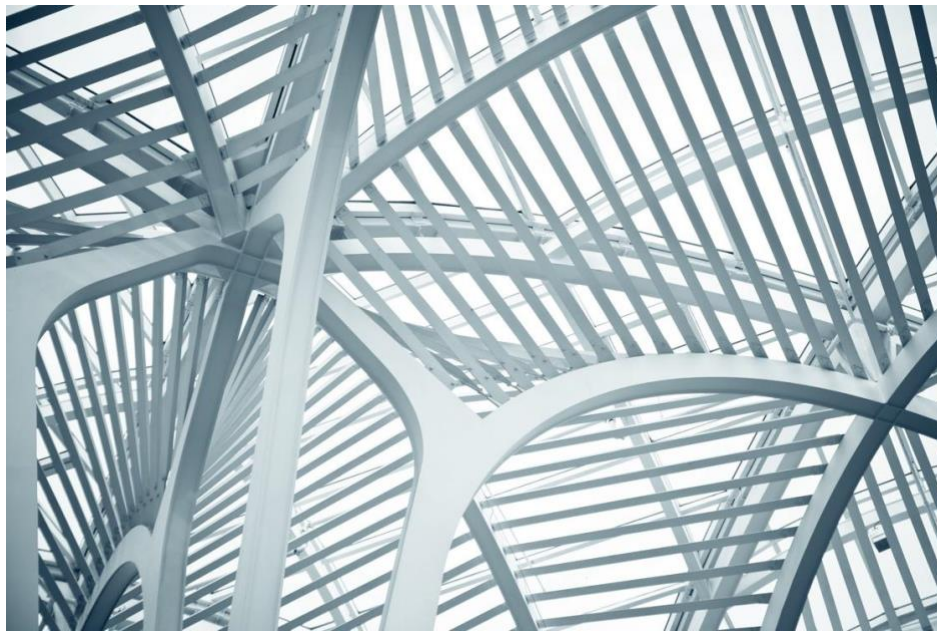


# Steering The Trojan Horse of Social Media:

**THE USE OF HUMAN-CENTERED AI IN MITIGATING HUMAN TRAFFICKING ONLINE**

By Délon Omrow and Vanesa Singh



**DIGITAL LIFE INSTITUTE | VERSION 1.0, 12 MARCH 2025**

© 2025 Délon Omrow and Vanesa Singh, All rights reserved

**Contents**

Introduction 2  
Challenges 4  
Humanity and AI can steer this Trojan Horse 4  
References 6

**Introduction**

The digital age features both the virtues and vices of connectivity, especially when analyzing how prominent human trafficking has grown as a result of the internet and social media. Human trafficking is often described as a modern-day form of slavery, involving the recruitment, transportation, harbouring and/or exercising control, direction or influence over the movements of a person in order to exploit that person (Public Safety Canada, 2025). Operations involving human trafficking often target people of all ages, but there is a disproportionate number of young women and girls impacted by sexual exploitation. For example, Statistics Canada reports that 25 percent of victims of human trafficking are under the ages of 17, and the form of exploitation encompasses violent threats to the victims or their families; sextortion; intimidation by use of force; and sexual assault (Lopez-Martinez, 2023). Considered by some to be the proverbial Trojan horse (Armato, 2023), social media filled a void for youth during the COVID-19 pandemic who were impacted by limited in-person socialization, and forced (and encouraged) them to leverage social networks to connect with friends. Predators, however, exploit this sense of vulnerability, using various social media applications to lure and groom young children and youth into human trafficking rings - among these new applications is Snapchat and its popular disappearing message feature. Unfortunately, Canada has been identified as a source, destination and transit country for victims of human trafficking. Take, for example, the city of London in Ontario: this region has garnered the attention of law enforcement due to its strategic location between Toronto and Windsor on Hwy. 401. Considered a “hub” for trafficking operations, at least 51 percent of victims are Indigenous women and children, most of whom are lured by social media via "Romeo pimp" interactions - that is, traffickers exploiting these women and children’s insecurities and vulnerabilities to develop a romantic relationship they can exploit (Lopez-Martinez, 2023).

The healthy balance between freedom of speech and the protection of the most vulnerable online has created a somewhat circuitous discourse on the role governments should play in reducing the risk of young women and girls being exposed to trafficking and sex exploitation. There have been some developments in the area - most notably, the United Nations calling upon big tech social media platforms to add features that would protect all users in 2020. Specifically, the UN is calling for these big tech companies to use data and algorithm tools to assist in the detection of human trafficking patterns so that law enforcement can identify suspicious and illicit activity (“Using AI to Fight Trafficking Is Dangerous,” 2024). As such, an adversarial relationship has developed between the technology industry and anti-human trafficking advocates, with the latter arguing that social media and Web 3.0 technologies have caused a significant rise in human trafficking for sexual exploitation (Drake, 2025). This serves

as a clarion call because those involved in trafficking operations are using new and evolving technology to gain efficiencies of scale, expanding the online commercial sex market. The question at this juncture, then, is what role AI is poised to play in mitigating human trafficking, opening new vistas in inquiry and activism. In other words, how do we steer this Trojan horse? Perhaps the answer can be found in the Canadian government's "National Strategy to Combat Human Trafficking 2019-2024", which highlights the need for more research on technological advancements and the use of AI. One example of the strategy is Canada's International CyberCrime Research Centre, located out of Simon Fraser University, which promotes education and conducts research in cybercrime prevention, detection, and response, working alongside public and private sectors at the regional, national, and international levels. One such collaboration occurred in 2022 when the Centre teamed up with Sexual Exploitation Education for a study supported by the Ministry of Public Safety and law enforcement. Using AI "web-crawlers", researchers scoured the internet in search of red flags: phrases such as "available 24-7," signs of control, or emojis like cherries, growing hearts or planes. The study concluded that of the 6,000 sex work advertisements scanned in British Columbia, approximately 40 percent of the advertisements were indicative of human trafficking (Radio-Canada.ca, 2022). In a related vein, researchers at McGill University and Carnegie Mellon University (CMU) have developed an algorithm that identifies human trafficking operations online by reviewing similar phrasing and duplication via advertisements. Dubbed "InfoShield", the algorithm collects millions of advertisements, highlighting common phrases and the use of emojis so that law enforcement can direct their investigations more efficiently. During one exercise, "InfoShield" was applied to a set of escort listings and outperformed other algorithms at identifying trafficking ads, flagging these ads with 85% precision (Carnegie Mellon University, 2021).

South of the border, The Office on Trafficking in Persons (OTIP) in the United States has promoted risk management frameworks which use new and emerging technology to prevent, detect and respond to human trafficking. AI-powered monitoring systems have been leveraged to identify suspicious patterns in supply chains and labor practices, along with digital identity verification strategies to protect the most vulnerable (Pimentel, 2025). AI-powered surveillance, also, has used deep learning in mitigation efforts, incorporating algorithms to detect distinctive patterns of trafficking activities (Ijiga et al., 2024). While traffickers elude detection by using social media to enlist victims, AI and deep learning draw upon complex neural networks to analyze and interpret large sets of data, recognizing indicators of suspicious financial activity, unusual travel patterns, and codewords in online advertisements. The automation of this type of work enhances both the speed and accuracy of uncovering potential cases of human trafficking, facilitating timely interventions by law enforcement (MDiv, 2024). The U.S. government has attempted to employ AI to assist law enforcement in prosecuting human traffickers; however, in 2024, the NGO Tech Against Trafficking revealed that of the 300 different technical tools to fight human trafficking listed in their database, just over 10% could be considered AI technologies, and the majority of these AI tools were no longer available as a result of "digital attrition" (Drake, 2024).

## Challenges

While the use of AI in combating human trafficking holds tremendous promise, certain activist groups lament the possibility of AI drawing upon gendered, racial, and socioeconomic biases in some algorithmic platforms. Algorithmic bias can create challenges for law enforcement; for example, language models are often built on discriminatory stereotypes. Shih (2021) reveals that anti-trafficking organizations and law enforcement training draw upon racist optics when searching for potential victims: markers such as poverty, sexuality, and race result in false identification of human trafficking victims, calling for more human-centered research on the best approaches for mitigation when using AI. In fact, the best applications of AI include a focus on activities, infrastructure, and what is referred to as the “digital dust” of the traffickers. Drake (2024) explains that criminals want to be found, just not by law enforcement; as such, social media can be steered by new and improved intelligence-gathering opportunities via AI whereby law enforcement can begin by examining blockchain transactions from cryptocurrencies associated with digital trafficking advertisements on the open or the dark web.

## Humanity and AI can steer this Trojan Horse

The global supply/demand marketplace for trafficking presents many challenges for law enforcement, but Big Data, AI and human ingenuity can adopt a different approach to tackling this scourge. Nosta (2025) suggests that AI and large language models (LLMs), while impressive, serve as massive libraries - that is, they are rich with knowledge but frozen in time once released. The author poses the question: what if AI could learn and adapt in real time, like human beings? This is precisely why human ingenuity cannot be eclipsed by techno-solutionism when addressing human trafficking: we must work with emerging tech to foster real-time adaptability and advanced memory systems, so that machine learning can help us understand patterns drawn from human inferences. Wu (2020) explains that the Defense Advanced Research Projects Agency (DARPA), in recent years, has developed technologies now used by companies such as Marinus Analytics and Giant Oak to assist law enforcement. Marinus Analytics uses AI to uncover and detect the retail components of human trafficking rings by leveraging machine learning in their software *Traffic Jam* to cull data from the Internet, identifying patterns of human trafficking operations.

Marinus Analytics has helped law enforcement in the U.S. recover many victims of human trafficking, highlighting the need to build more communities of practice and develop more innovative software, tools, and resources in the fight against human trafficking. The Anti-Human Trafficking Intelligence Initiative, a non-profit focused on the sharing of intelligence, has developed an app that can be integrated onto cell phones, allowing victims to scan QR codes that are put up in hotel bathrooms and other public places. Law enforcement, in turn, can use this data to follow up on leads by way of requesting a subpoena to gain access to cell phone records, verifying potential human trafficking operations. Despite such technological innovations, it will take humanity and AI to steer the trojan horse of social media. Wu (2020)

astutely comments on the responsibility that humans must exercise when making judgments on when to use the technology and how to use the technology. Technology, AI and machine learning, when reduced to its core, is pattern recognition and humans must develop an ethical and sustainable training set to identify the patterns we wish to identify. To quote Dr. Gary M. Shiffman, Founder of Giant Oak:

“You need clean data that you can run your algorithm through to identify the pattern. There could be problems in your training data. There could be problems in the data that you are running your algorithm against. You need a human in the loop to ensure that someone’s applying human reasoning, common sense, and judgement to the inputs and output” (quoted in Wu, 2020, para. 19).

Clearly, the need for human intuition in the use of technology impacts the outcome of human trafficking investigations, but so too does the voice of victims. Best practices for supporting victims have included reflection delay - that is, allowing victims of human trafficking to remain legally in the country while they seek support by way of access to shelters, legal advice, counselling and medical care (Jorge-Birol, 2011). New paradigms including survivor-centric approaches to understanding victimization rooted in principles of co-design, empathy, and human dignity are being promoted, especially by Polaris, a nonprofit working to end sex and labor trafficking in North America. New machine learning and AI-driven pilots are being explored by Polaris, but in conjunction with what is referred to as “survivor-centered AI”. Data solutions are leveraged as tools to understand trends in human trafficking by Polaris but critical gaps in data about the needs of trafficking survivors remained so Polaris spearheaded the creation of the National Survivor Study (NSS), a survivor-centered, justice-driven data initiative used for gathering data, while safeguarding survivors’ privacy and security. The importance of this data initiative is that it was co-designed with survivors, drawing upon their perspectives to shape the research, survey tools and advance data analysis capabilities (Botti-Lodovico, 2024).

As the fight against human trafficking continues on various fronts, human-in-the-loop design of AI is the only hope for steering the Trojan horse of social media and its role in human trafficking. Transnational cooperation is needed to strengthen governance initiatives with advanced AI technologies for many forms of international crime because of social media’s reach (Omrow et al. 2024). Welle (2024) reminds us that platforms like TikTok have become an even more popular marketplace for traffickers and people smugglers. Whether it is the “coyotes” offering their services on social media platforms to smuggle unauthorized migrants into the USA, or the flurry of advertisements on Meta enticing people with the promise of lucrative jobs in Europe, there is no shortage of human trafficking identifiers or patterns online, and the right balance of human-centered AI must be struck to continue this noble fight.

## References

- Armato, L. (2023, April 1). Is TikTok your best friend Or a Trojan horse? Forbes. <https://www.forbes.com/sites/leonardarmato/2023/04/01/is-tiktok-your-best-friend-or-a-trojan-horse/>
- Botti-Lodovico, Y. (2024, December 20). Catherine Chen & Sara Woldehanna: Ending Human Trafficking with Survivor-Centered AI. Medium. <https://medium.com/patrick-j-mcgovern-foundation/catherine-chen-sara-woldehanna-ending-human-trafficking-with-survivor-centered-ai-9aea555226cf>
- Carnegie Mellon University. (2021, April 23). Algorithm uses online ads to identify human traffickers. Machine Learning | Carnegie Mellon University. <https://www.ml.cmu.edu/news/news-archive/2021-2025/2021/april/machine-learning-ai-algorithm-uses-online-ads-identify-human-traffickers.html>
- Drake, B. (2025). Backwards thinking on artificial intelligence (AI) and human trafficking. Stimson Center. <https://www.stimson.org/2025/backwards-thinking-on-artificial-intelligence-ai-and-human-trafficking/>
- Ijiga, N. a. C., Olola, N. T. M., Enyejo, N. L. A., Akpa, N. F. A., Olatunde, N. T. I., & Olajide, N. F. I. (2024). Advanced surveillance and detection systems using deep learning to combat human trafficking. *Magna Scientia Advanced Research and Reviews*, 11(1), 267–286. <https://doi.org/10.30574/msarr.2024.11.1.0091>
- Jorge-Birol, A. P. (2011). Empowering Victims of Human Trafficking: the Role of Assistance, Protection and Re-Integration Programs. *HUMSEC Journal*. Issue 2. <https://documentation.lastradainternational.org/doc-center/2844/empowering-victims-of-human-trafficking-the-role-of-support-assistance-and-protection-policies>
- Lopez-Martinez, M. (2024, November 26). Human trafficking tactics increasing online: Why advocates are calling for crackdown in the cyberspace. CTVNews. <https://www.ctvnews.ca/canada/article/human-trafficking-tactics-increasing-online-why-advocates-are-calling-for-crackdown-in-the-cyberspace/>
- MDiv, W. L. P. J., PhD. (2024, December 27). Can deep learning rescue victims through recognition and prevention? *Psychology Today*. <https://www.psychologytoday.com/ca/blog/why-bad-looks-good/202412/using-artificial-intelligence-to-detect-human-trafficking>
- Nosta, J. (2025, January 16). How real-time learning is transforming AI into adaptive, thinking partners. *Psychology Today*. <https://www.psychologytoday.com/ca/blog/the-digital-self/202501/teaching-artificial-intelligence-to-think-on-its-feet>



Omrow, D., Anagnostou, M., Cassey, P., Cooke, S. J., Jordan, S., Kirkwood, A. E., MacNeill, T., Mirrlees, T., Pedersen, I., Stoett, P., & Tlusty, M. F. (2024). Compliance and enforcement in a brave new (green) world: Best practices and technologies for green governance. *FACETS*, 9, 1–8. <https://doi.org/10.1139/facets-2023-0105>

Pimentel, B. (2025b, February 25). Technology and human trafficking: Fighting the good fight. Thomson Reuters Law Blog. <https://legal.thomsonreuters.com/blog/technology-and-human-trafficking/>

Public Safety Canada. (2025, January 9). About human trafficking. <https://www.publicsafety.gc.ca/cnt/cntrng-crm/hmn-trffckng/abt-hmn-trffckng-en.aspx>

Radio-Canada.ca. (2022, July 18). AI helps researchers identify victims of human trafficking | RCI. <https://ici.radio-canada.ca/rci/en/news/1899011/ai-helps-researchers-identify-victims-of-human-trafficking>

Shih, E. (2021). The Fantasy of Human Trafficking: Training Spectacles in Racial Surveillance. *Journal of Transnational Women's and Gender Studies*, 22, 105-137.

Using AI to fight trafficking is dangerous. (2024, July 1). Human Rights Watch. <https://www.hrw.org/news/2024/07/01/using-ai-fight-trafficking-dangerous>

Welle, D. (2024, January 26). How social media aids human trafficking and smuggling. *dw.com*. <https://www.dw.com/en/how-social-media-aids-human-trafficking-and-smuggling/a-67958534>

Wu, J. (2020, April 14). AI is helping us combat the economic problem of human trafficking. *Forbes*. <https://www.forbes.com/sites/cognitiveworld/2020/04/14/ai-is-helping-us-combat-the-economic-problem-of-human-trafficking/>