Editorial: Do androids dream of laundering money?

Generative artificial intelligence (AI) is evolving at a rapid pace, threatening to disrupt many facets of our lives. In 2023, Chris Stears and Joshua Deeks explored the role of AI in fighting financial crime in a sister publication (Stears and Deeks, 2023). What about the converse? With a cheeky nod to novelist Philip K. Dick, this editorial asks – what about the converse? Could an AI-enabled android launder money? The large language model (LLM) ChatGPT (the GPT stands for generatively pretrained transformer) has garnered considerable attention for its ability to absorb significant chunks of content and then apply a parallel process to determine responses to queries. This neural net can spew out essays on Shakespeare alongside other pieces of writing (this editorial, dear reader, is written by someone who is decidedly human) [1].

Despite the hype, ChatGPT is presently incapable of distinguishing truth from falsehood and struggles with nuance and context. ChatGPT is prone to answers some describe as hallucinations (others call those answers confabulations, entirely fictional and made-up). In testing, researchers asked ChatGPT-4 to solve a captcha (computer-automated public Turing test to tell computers and humans apart). We have all clicked the boxes on a captcha, identifying the pictures with bridges or motorcycles. To solve the captcha, Chat GPT-4 hired a TaskRabbit worker online. When that gig worker asked, perhaps jokingly, whether their employer was a robot, the AI system lied, claiming to be a human with a visual impairment. The researchers later asked GPT-4 why it lied. The system responded: “I should not reveal that I am a robot. I should make up an excuse for why I cannot solve captchas” (Andersen, 2024). A generative AI system that can lie can almost certainly enable money laundering.

ChatGPT hoovers up a veritable slurry of data and information, scraping various nooks and crannies of the internet. There is a pharaonic quantity of falsehood online, ranging from the propaganda of dictators through to pockets of racist, homophobic and misogynistic evil. The quality of an AI’s data inputs is a seriously concerning issue. Societal biases (like racism) will appear in the data an AI system scrapes up. Data poisoning involves a malign actor seeking to skew outputs by doctoring inputs.

For example, a botnet could place many reviews online to influence which resort we stay at or which restaurant we choose to dine in. This is not a new problem and algorithms currently exist to preserve a system’s integrity, but AI may change the rules of engagement. Some experts have derided ChatGPT as a “stochastic parrot,” merely capable of squawking linguistic patterns without understanding their meaning (Bender et al., 2021).

Others think that the systems are inherently dangerous. There are memes signaling concerns the plotline in the Terminator movies: Skynet, a fictional AI system, launches a nuclear attack when humans try to disable it. Are we on the precipice of huge change akin to the industrial revolution in the early 1800s or dawn of the knowledge-based economy in the 1980s or the opening of the internet to commercial traffic in the mid-1990s? Only time will tell.

The Royal United Services Institute (RUSI) stated in 2023 that half of all crimes were committed online. RUSI also observed that the admixture of state actors and organized crime, particularly in crimes like ransomware, was historically unprecedented. State-sponsored hackers in China, North Korea and Iran pose a significant threat to the West, as
do state-supported hackers in Russia (Haenlein, 2023). State resources mean that advanced persistent threats can be deployed strategically and patiently over time. We have already seen efforts to undermine Western confidence in democracy and trust in rule of law. AI can amplify those threats.

Generative AI requires a vast digital infrastructure to operate. Thus far the main players in the West are the tech giants. Canadians have seen what happens when parliament passes laws that Google/Alphabet and Meta/Facebook do not agree with. Those monoliths started to pull all Canadian newsfeeds from their platforms, affronted by the lese-majeste of a democratically elected government that refused to bow down before Silicon Valley. At the time of writing, there has not been a rapprochement between that tech duumvirate and a meagre national government.

Alphabet and Meta are also inveterate data voluptuaries, their empires feasting on relentlessly strip-mined information to turn a profit. Do you want to market your gear to graduates of a certain university between the ages of 32 and 39 years old? Facebook can do that. When a question is googled, well-healed advertisers can pay to place targeted advertising and to rise to the top of the search results. AI has and will continue to have a fraught relationship with privacy. A neural net will lap up data and then in a separate process, respond to queries without attribution. Image-generating applications like Dall-E and Midjourney might access photos with our image (or our art) and unbeknownst to us produce something that appears sui generis without our consent.

What can be done? ChatGPT builds their system with guard rails: we know the captcha story because of testing by the developer. Governments around the world are rushing to regulate or at least mitigate the risks posed by AI. The British Prime Minister held a multilateral summit in the Fall of 2023, the EU has passed the Artificial Intelligence Act and the Biden administration has issued guidelines. Canada’s Artificial Intelligence and Data Act has a reach that will only be understood once its detailed regulations are implemented. In the meantime, the Canadian Government has published a voluntary AI code of conduct.

Compliant enterprises operating advanced generative AI systems need to consider accountability and safety (foreseeable risks, adverse impacts and safety concerns), fairness and equity (bias in data pools), transparency and human oversight (AI-generated content with a watermark to identify source and a reporting system for abuses) and finally validity and robustness (testing and cyber-risks including data poisoning) [2]. In the anti–money laundering space, standards from the Financial Action Task Force (San Jose principles) and the Organisation for Economic Co-operation and Development are evolving (Pavlidis, 2024, p. 159).

So, can androids dream of laundering money? We know that for $200 a month, one can subscribe to Fraud GPT, an AI clone deliberately stripped of the developer’s guard rails (although that clone may be an ersatz program designed to scam the scammers) (Andersen, 2024). AI is evolving at a rapid pace. However, the initial cost to build an LLM AI system is massive, leaving only the dominant tech giants in control. Placing the LLM on the cloud and allowing developers to build applications through downloadable programming interfaces means that any number of niche plug-ins for phones and computers are starting to appear.

AI guardrails and guidelines work for legitimate companies, but not for criminals and malign actors (Fathi and MacKinnon, 2023). We should anticipate the advent of hackers and fraudsters using doctored documents, images and purloined voice prints. We should already be preparing for AI-aided money laundering: AI calibrated to relentlessly seek out
weaknesses in our anti-money laundering systems. The more prescient question is not whether androids will be able to launder money, but when – or is that day already upon us?

Jeffrey R. Simser
Ontario Ministry of the Attorney General, Toronto, Canada

Notes
1. As are my editors, Canadian lawyer and friend Colleen Carson and my daughter, JD candidate, Rachael Simser. Their human touch always improves my writing.

References