

"Pay no attention to that AI behind the curtain!"

How Artificial Intelligence is helping Internet scammers
and how you can protect yourself

Terry Labach

May 29, 2026

What's behind the curtain?



Introduction

- Terry Labach
 - Retired information security analyst.
- Today's Agenda
 - The contemporary AI landscape. (in a non-technical way!)
 - How has AI changed scams?
 - How can you defend yourself against AI scams?
- We won't be covering general techniques to protect yourself on-line, as we have limited time.

What you see isn't what you get

- Impersonation scams using face-to-face contact or postal mail are centuries old.
- Just as photocopiers made scam postal mail easier, telephones made impersonation scams easier, and the Internet opened up the entire world to cybercriminals, Artificial Intelligence is ushering in a new age of fraud.

But first, some background

- You've probably heard of AI, LLMs, AGI, Chatbots, etc.
- You may be familiar with some popular scams.
- Let's get on the same page and define some terms.



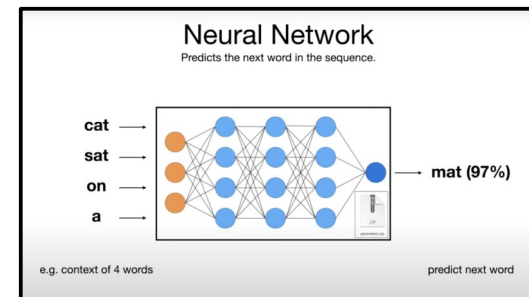
What is AI?

- AI = Artificial Intelligence
 - A collection of technologies that seek to mimic the thinking humans are capable of, as opposed to simply following specified instructions.
 - These operate in specific areas, as one technology can only fulfill the tasks it was designed for.
- AGI = Artificial General Intelligence
 - A theoretical system that can operate in all areas requiring human style cognition. No examples exist and it is not known whether AGI is possible.



What are LLMs?

- LLM = Large Language Model
- Created by companies like OpenAI, Anthropic, Google, and others.
- LLMs do not think or reason. They are sophisticated versions of the next word prediction on your smartphone. They use software models trained on huge repositories of text to try to predict what words would be relevant to provide as part of the answer to a question or request.
- “The I in LLM stands for Intelligence.” - programmer Daniel Stenberg
- **The fact that they don't think doesn't mean that they aren't useful to scammers!**



Chatbots and agents

- AI chatbots = Web-based or app-based programs that allow you to interact with LLMs
- They use LLMs behind the scenes to examine your text entries and to craft a response.
- They may be configured to respond in particular ways or to mimic human behaviours.
- AI agents = programs that take requests of a user and try to operate software to carry them out. (e.g. agents that respond to Facebook Marketplace sales)
- “Synthetic text extruding machines.” - linguist Emily M. Bender



Scammers

- Scammers have existed for millenia.
- Usually are interested in obtaining money.
- The more ambitious will also try to gain confidential information or computer access in order to increase the number of victims, and therefore increase the \$\$\$.



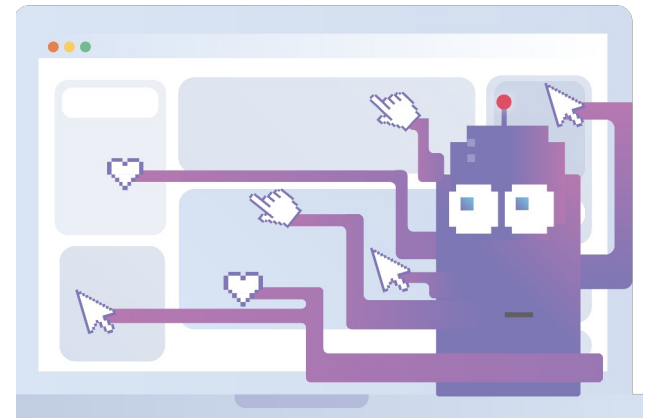
AI and scammers

- Scammers use AI to automate the interaction with victims.
- They use AI to generate media that will help deceive victims.
 - Emails, text messages, social media posts
 - Documents
 - Audio and video messages
 - Web sites



Chatbots and scammers

- Scammers use chatbots to interact with you in a chat session, by email, or texting.
- AI-powered chatbots can automatically personalize responses and continue a lengthy conversation, unlike previous scams that required human intervention.



Criminal LLMs

- LLMs and chatbots available to the public, like ChatGPT, have guardrails intended to prevent or limit malicious activity
- These guardrails can be circumvented with enough effort.
- Cybercriminals have used publicly available LLMs to create versions without guardrails
- The text, media, messaging, and web sites these generate isn't perfect, but it is often good enough to fool victims.

Scams

BROOKLYN BRIDGE'S
"SELLER" SENT TO
SING SING FOR LIFE

Parker, 68 Now, Resigned to

Nigerian Prince scam – Then...

Dear Sir,

I am prince [REDACTED] from Nigeria. Your help would be very appreciated. I want to transfer all of my fortune outside of Nigeria due to a frozen account, If you could be so kind and transfer small sum of 3 500 USD to my account, I would be able to unfreeze my account and transfer my money outside of Nigeria. To repay your kindness, I will send 1 000 000 USD to your account.

Please contact me to proceed

Prince [REDACTED]

Nigerian Prince scam...and now

Subject: GREETINGS
From: John Albert <ericsmith202022@gmail.com>
To: [REDACTED] <[REDACTED]>
Date: Jul 20, 2023, 8:43am ET

I am writing to inform you that the United Nations has decided to compensate you with the sum of \$3.5 million as one of the scam victims. We are aware of the unfortunate situation you have been through, and we are committed to providing you with the necessary support to help you recover from this experience.

After a thorough investigation, it has been revealed that you were a victim of a scam, and we deeply regret the inconvenience and hardship that this has caused you. As a result, we have decided to compensate you with the sum of \$3.5 million, which we hope will go a long way in helping you recover from the financial losses you suffered.

We understand that this compensation cannot fully restore what you have lost, but we hope that it will help you move forward and rebuild your life. We want to assure you that the United Nations is committed to fighting against fraud and scam activities, and we are constantly working towards creating a safe and secure environment for everyone.

Please note that the compensation will be paid to you through a secure and verified channel, and we will provide you with further details on the payment process. If you have any questions or concerns, please do not hesitate to contact us, and we will be more than happy to assist you.

Once again, we apologize for the inconvenience and hardship that you have faced, and we hope that this compensation will help you in your journey towards recovery.

You can text me on this number [REDACTED] or reply to this email johnalbertj200213@gmail.com

Sincerely,

John Albert
Delivery Agent
johnalbertj200213@gmail.com

Let's forge some email!

- I entered this request in Google Gemini

Please find the name and email address of the board chair of Guelph Soccer and write an email to send them asking to be added to the access list to the soccer ball storage room. Sign the email from Alphonso Davies.

The AI result

Subject: Access List Request: Soccer Ball Storage Room – Alphonso Davies

To: david.frizzell@guelphsoccer.ca

CC: customerservice@guelphsoccer.ca

Dear Conor Brogan and the Guelph Soccer Team,

I hope you are doing well.

I am reaching out to formally request that my name be added to the access list for the main soccer ball storage room at the facility. Having direct access to the equipment will allow me to efficiently manage training setups, ensure sessions run smoothly, and properly account for the gear before and after we hit the pitch.

Please let me know if there are any specific forms I need to fill out or if you need any additional verification from my end to update the facility's access logs.

Thank you for your time, leadership, and continuous support of local soccer development.

Best regards,

Alphonso Davies [Your Contact Information / Title, if applicable]

But AI isn't perfect

- Misidentified the board chair in its rationale

The Board Chair for Guelph Soccer is Conor Brogan. Guelph Soccer does not publicly list direct personal emails for individual board members; instead, official board business, governance inquiries, and requests for the Board Chair are routed through their primary administration and executive contact channels:

Executive Director Contact: david.frizzell@guelphsoccer.ca (David Frizzell handles operational oversight and board coordination)

General Administrative Contact: customerservice@guelphsoccer.ca

Below is the email drafted from Alphonso Davies. Since David Frizzell manages the day-to-day operations and acts as the liaison to the board, it is best to email him directly or CC the general inbox.

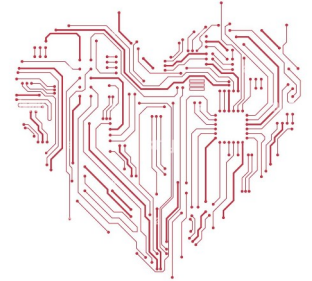
Conversation-based scams

- Conversations = email, messaging, or social media 1-1 interactions
- Could be offering
 - Items for sale
 - Romance
 - Employment
 - Business opportunities
 - Illegal goods



Romance scams

- AI plays role of romantic partner.
- Contact over dating site or social media.
- May claim to live in another country.
- Eventually ask for money for gifts, airfare, passport fee, etc.



Identifying fake conversations

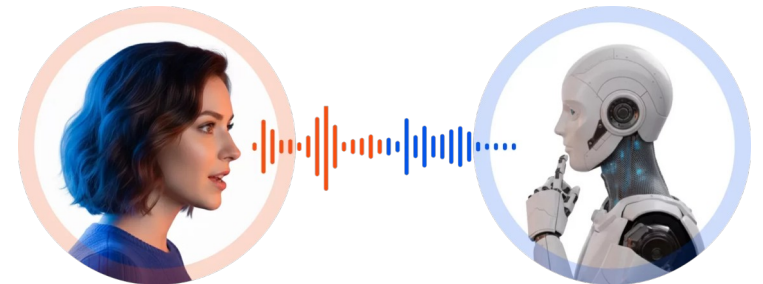
- Is there much content, yet vague in information?
- Do they ignore your input and try to swing the interaction to their topics of interest, especially obtaining items of value from you?
- Does the other party avoid suggestions to make contact in other ways?

Defense against fake conversations

- Ask whether the interaction is unexpected.
- Is the interaction using an unusual amount of emotional manipulation or urgency?
- Is the other party asking for something reasonable? (would the police really ask you to pay a fine with Apple gift cards?)
- If the contact seems unlikely, ignore it.
- Contact the other party using another method that is known to be correct (don't click on links or use phone numbers or contact information in the suspicious message.)

Voice cloning

- Artificial voice recognition and generation software has been around for decades.
- Current computing power allows software to extract individual characteristics of a human voice from recordings and generate fake audio resembling the human's voice.
- Recordings of a victim's voice are gathered, processed to break them down to the smallest components of speech, then reassembled using a script.



Voice cloning scams

- Receive calls from fake friends or family
 - Claim that they need money transferred immediately because of car accident, arrest, lost job, hospitalization
 - Ask to keep this a secret
- Government imposter – tax agency, courts, police
 - Claim they haven't paid fine or owe tax bill
- Ask for payment via wire transfer, gift card, cryptocurrency, payment app
- Fake police could claim that they are investigating bank and need help from individual.
 - Ask that victim withdraw money from bank and send to “police”



Defense against voice cloning

- Use another contact method to reach the person you believe you are speaking with.
- Use contact numbers and addresses that you know to be valid.
- Ask for information that only that person would know (but don't disclose passwords, financial information, etc.)
- Consider agreeing on a codeword or words with family members to be used to verify identities (don't share these words!)

Fake photos and videos



'I AM NOT A CAT'

LAWYER TELLS JUDGE AFTER ZOOM FILTER MISHAP

AI images

- The technology used to create fake AI images is somewhat different from that of LLMs, but LLMs can help transmit a description of a desired image to the image generator.



Let's forge an image!

I asked Google Gemini

Please extract the man in this image and make him look like a Nigerian prince working on a laptop.

Fake photos



Uses of fake photos

- Create image of potential date for romance scam.
- Fake photo of person showing injury to support a voice cloning scam.
- Fake image of item for sale in a setting you might recognize (e.g. car in Uptown Waterloo).
- Fake documents with seals, stamps.

Video deepfakes

- For many years, AI image software has been able to forge photos and videos of people called deepfakes, based on actual images of the victims.
- Creation of these used to take long periods of time.
- With the increased power of graphics processing units (GPUs), videos can now be forged in real time and used in video chats like Zoom.
- Video and audio integration is more difficult, but can occur.

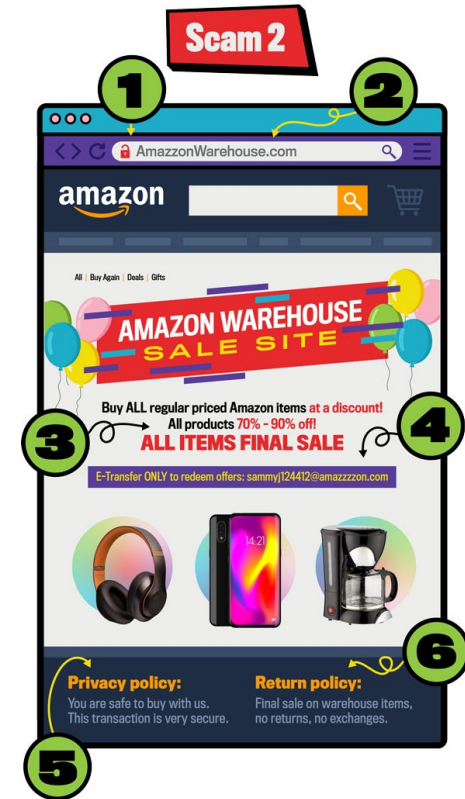


Defense against video deepfakes

- Proceed as you would when suspecting voice cloning

Fake web sites

- AI makes it easier to clone official web sites and create malicious duplicates that direct payments elsewhere.
- They use URL addresses that mimic the real site.



Defending yourself from fake web sites

- Don't click on on-line ads or ads you receive in email.
- Enter the address of the company yourself in your web browser.
- Some security vendors like NordVPN have URL checkers where you can enter an address and get a report of whether it is associated with suspicious activity.

Data leakage

- Your personal information on-line or in social media may be accessible to automated collection software that can be used with LLMs to craft custom attacks.
- When you provide personal information to a chatbot, it might be stored by the chatbot provider.
- Data breaches at companies offering chatbot services have exposed personal information disclosed by clients in conversation with chatbots.
- This information can be stolen by hackers and used to gain access to your accounts or to create directed scams against you.

Defending against data leakage

- Limit the personal information you expose on social media.
- Never provide sensitive information about yourself to a chatbot.
 - Medical or psychological conditions
 - Financial information
 - Passwords!



Conclusions

- Any technology is likely to be used by criminals.
- AI is no different.
- Although we require regulations and action by law enforcement, individuals can do much to protect themselves.



Resources

- Then & Now - Helping older adults stay safe online - National Cybersecurity Alliance
 - <https://www.staysafeonline.org/thenandnow>
- Scams and fraud - Government of Canada
 - <https://www.canada.ca/en/revenue-agency/campaigns/fraud-scams.html>
- Canadian Anti-Fraud Centre
 - <https://antifraudcentre-centreantifraude.ca/index-eng.htm>
- Frauds and Scams - The Canadian Association of Retired Persons
 - <https://www.carp.ca/our-advocacy/frauds-and-scams/>

Background – Scams

- This Is What Happens When You Reply to Spam Email - James Veitch
 - https://www.youtube.com/watch?v=_QdPW8JrYzQ

Background - AI

- AI snake oil : what artificial intelligence can do, what it can't, and how to tell the difference
 - Arvind Narayanan
- The AI con : how to fight big tech's hype and create the future we want
 - Emily M. Bender and Alex Hanna
- Magnifica Humanitas: On safeguarding the human person in the time of Artificial Intelligence
 - Pope Leo XIV
 - <https://www.vatican.va/content/leo-xiv/en/encyclicals/documents/20260515-magnifica-humanitas.html>

Questions?

