PERSONAL AI FRAUD PREPAREDNESS / NOVEMBER 2026

AI SAFETY. Everyday confidence.

AI can fake the voice, the face and the story.

A practical companion to the YouAIFool assessment.
For you, your family and the people who trust you.

01-02 Identity & manipulation

03-04 Evidence & private data

05-06 AI access & incident response

PAUSE. CHECK. THEN ACT.

Slow down. Verify through a route you already trust.
Only then decide what to do.

01 When someone sounds familiar

A voice, face or account is a clue. It is not enough to authorise a sensitive request.

THE SITUATION

Your daughter calls from an unfamiliar number. She sounds frightened, says her phone is broken and asks you to transfer money without telling anyone.

WHAT MAKES THIS CONVINCING

A scammer may use a cloned voice, personal details or a compromised account. A real-looking video call or a message in an existing conversation can still be part of a scam.

WHAT TO DO

End the incoming call. Contact the person using a saved number or another established route. If you cannot reach them, check with a second trusted person before acting. Do not use a number supplied in the suspicious exchange.

PUT THIS SAFEGUARD IN PLACE

Agree a private family phrase for unusual requests. Keep it off social media and out of shared chats. Treat it as an extra check: if it is exposed, change it. Never share login codes or approve a login you did not initiate.

WORDS YOU CAN USE

“I want to help. I will call you back on the number I already have.”

02 When the story fits your life

Personal knowledge can make a fraudulent request feel like a natural next step.

THE SITUATION

A recruiter knows your role and recent projects. After a plausible exchange, they ask you to install a special interview app. Or an online acquaintance builds trust before suggesting an investment.

THE PRESSURE TO NOTICE

Urgency, secrecy, flattery and fear can narrow your attention. Accurate personal details may have come from public posts, leaked data or earlier conversations. They do not establish trust.

WHAT TO DO

Verify the organisation and person through contact details you find independently. Pause before installing software, moving money or sharing documents. Discuss unusual requests with someone outside the conversation.

PUT THIS SAFEGUARD IN PLACE

Make a personal rule: no financial decision under pressure. Review what strangers can see about your family, job and routines. Never give an unsolicited contact remote access to your device.

WORDS YOU CAN USE

“Please send the details. I will verify them independently and come back to you.”

03 When seeing feels like proof

A screenshot, document or video can support a story without proving it.

THE SITUATION

A convincing clip shows a public figure promoting an investment. A professional-looking dashboard displays profits. A payment screenshot appears to prove that money is on its way.

WHAT TO CHECK

Look for confirmation at the original source through a route you find yourself. Check your own bank account for funds received. Verify a changed bank account with a known contact before paying.

WHAT TO DO

Separate the claim from the media: who is asking you to do what? A realistic image does not prove a transaction, identity or endorsement. Small successful withdrawals or initial payments do not establish that a scheme is safe.

PUT THIS SAFEGUARD IN PLACE

Use more than one independent source for consequential claims. Do not rely on visual glitches, polished language or an AI detector as a final verdict. If the claim cannot be verified, pause the action.

WORDS YOU CAN USE

“I will check the original source and confirm the payment in my own account.”

04 Before you paste or upload

A useful prompt can reveal more about you or someone else than you intended.

THE SITUATION

You ask an AI tool to summarise a medical letter, contract or customer spreadsheet. You remove the name, but leave a rare job title, exact dates and enough context to identify the person.

WHAT TO REMOVE

Passwords, security codes and recovery keys should never go into a prompt. Remove unnecessary identifiers and sensitive details. Changing a name alone may leave a person identifiable.

WHAT TO DO

Use fictional examples or the smallest relevant extract. Check the tool’s current privacy, retention and sharing settings. A private-looking chat is not a promise that information stays only on your device.

PUT THIS SAFEGUARD IN PLACE

For work data, use approved tools and follow your organisation’s rules. Check attachments and screenshots before uploading. Verify important AI answers against reliable original sources before using them.

A SAFER PROMPT

“Explain this type of contract clause using a fictional example” is often enough. Start there before sharing the real document.

05 When AI can act for you

An assistant that can send, buy or change things needs tighter limits than one that only drafts.

THE SITUATION

You connect an AI assistant to your inbox and files to help plan a trip. It asks for permission to read all messages, send email and manage documents.

WHAT TO CHECK

Read the actual permissions. Does the task need access to everything? Can the assistant send messages, spend money, delete files or share information? Instructions in a webpage or document may try to steer its actions.

WHAT TO DO

Give only the access needed for the task. Start with viewing or drafting where possible. Require your review before payments, messages to others, deletion and sensitive sharing. Review recipients, amounts and attachments yourself.

PUT THIS SAFEGUARD IN PLACE

Use trusted tools and extensions. Review connected apps regularly and remove access you no longer need. If an agent behaves unexpectedly, stop it and revoke its access before investigating.

BEFORE YOU APPROVE

“What can this tool read, what can it change, and which actions will still need my approval?”

06 If something has already happened

Act quickly. You do not need to prove that AI was involved before asking for help.

MONEY OR CARD DETAILS

Contact your bank or payment provider immediately through its official app or a known number. Explain what happened and ask about blocking further payments and attempting recovery. Recovery is not guaranteed.

PASSWORDS, CODES OR ACCOUNT ACCESS

Use a trusted device to secure affected accounts. Change exposed and reused passwords, sign out other sessions and review recovery details. Secure your email first if it can reset other accounts. Enable multi-factor authentication.

REMOTE ACCESS OR UNWANTED SOFTWARE

Disconnect the affected device from the network and stop using it for sensitive activity. Seek trusted technical help. If work systems or work data are involved, notify your IT or security contact immediately.

KEEP EVIDENCE AND REPORT

Keep messages, account names, URLs, times and transaction references. Report through the affected platform and your local police or national cybercrime reporting channel. Tell trusted contacts if your account was used to approach them.

WATCH FOR A SECOND SCAM

Do not pay someone who unexpectedly promises to recover your money. Avoid sharing more information with the original contact.

THE TAKEAWAY

Make safety a shared habit.

Choose three actions today. Add the others this week.

AGREE A VERIFICATION RULE

Choose a private family phrase and a callback rule. Anyone can pause a request without embarrassment.

PROTECT YOUR KEY ACCOUNTS

Use unique passwords or passkeys. Enable multi-factor authentication, especially for email and banking.

REVIEW AI ACCESS

Remove unnecessary app permissions. Require review before payments, sending, deletion and sensitive sharing.

KEEP A RESPONSE PLAN

Save your bank’s official contact route. Know a trusted person to call and how to reach IT at work.

05 / SHARE WITHOUT BLAME

“If something feels wrong, call me. Even if you have already clicked or paid.”

KEEP NEARBY

The 20-second safety check.

01 Was I expecting this?

02 Is it asking for money, access or private information?

03 Have I verified it through a trusted route?

If you are unsure, pause the action.

SOURCES & FURTHER READING

FBI / IC3 / Criminals Use Generative Artificial Intelligence to Facilitate Financial Fraud (3 December 2024)
https://www.ic3.gov/psa/2024/psa241203

Canadian Centre for Cyber Security / Generative artificial intelligence - ITSAP.00.041
https://www.cyber.gc.ca/en/guidance/generative-artificial-intelligence-ai-itsap00041

OWASP / AI Agent Security Cheat Sheet
https://cheatsheetseries.owasp.org/cheatsheets/AI_Agent_Security_Cheat_Sheet.html

GOV.UK / NCSC reporting guidance / Report internet scams and phishing (UK reporting routes)
https://www.gov.uk/report-suspicious-emails-websites-phishing

Examples are illustrative. A high assessment score does not guarantee safety. Follow your bank’s, employer’s and service provider’s incident instructions. Public sources inform this guide; they have not endorsed YouAIFool or Globawise. Sources checked 7 November 2026.

youaifool.com