ARTIFICIAL INTELLIGENCE AND FRAUD:

A Tutorial on Detection, Verification, and Protection

Doctor Vermeille Global EdTech & Research LLC

SIX TYPES OF AI-ENABLED FRAUD

Phishing

An unsolicited email, text, or message designed to trick the recipient into revealing sensitive information (passwords, credit cards, tax identifiers, account details) or clicking a malicious link. Phishing messages imitate trusted organizations—banks, government agencies, companies. They create urgency and claim that immediate action is required. The sender's address may look official but the controlling domain is fake.

Spoofing

Impersonating a trusted person, organization, or device by disguising the sender's true identity. The display name says Bank of America but the actual sender address is different. The phone shows a number that looks like it belongs to the IRS but originated elsewhere. Caller ID can be spoofed. Email addresses can be forged to show a company name while the controlling domain belongs to a criminal. The visual appearance is deceptive.

Deepfake

An audio or video file created or heavily modified using artificial intelligence to show a person saying or doing something they did not actually say or do. AI can now generate realistic face movements, lip sync, and voice characteristics. A deepfake video might show a government official announcing a policy they never announced. A deepfake audio might present a corporate executive ordering an immediate transfer of money. Humans detect deepfakes less than 25 percent of the time.

Voice Cloning

AI-generated audio designed to sound like a real person's voice. A scammer can record a few seconds of a CEO's speech and use AI to generate an entire conversation that sounds authentic. The victim receives a call and hears a familiar voice saying: This is Jim. I need you to transfer $500,000 immediately. Do not tell anyone. The voice sounds right. The urgency feels real. The person believes they are following the CEO's instruction. Voice-cloning fraud increased 1,633 percent in 2026.

Impersonation

A person or AI-generated communication pretending to be someone they are not. The message claims to come from the Federal Trade Commission, the Internal Revenue Service, the Social Security Administration, or a government agency. It uses official terminology, agency names, logos, and legal-sounding language. It may include accurate public information (your name, address, or social security number partial) to build credibility. The goal is to obtain money, personal data, or action from the victim.

Recovery Scams

A fraudster contacts someone who has already lost money to a scam and promises to recover the lost funds for an upfront fee. The recovery artist may claim to be an FBI agent, attorney, or private investigator. They offer to negotiate with the original scammer or freeze stolen accounts. They say the recovery is almost certain and

requires a guarantee deposit. The victim has already experienced loss and is psychologically motivated to recover money. Recovery scams cost victims an additional $500 million annually.

POLISHED GRAMMAR, VIDEO, AND AUDIO: NO LONGER PROOF OF AUTHENTICITY

For decades, a simple rule worked: if someone sent professional writing, clear video, or clear audio, they were probably legitimate. Poor spelling or bad grammar signaled a scam. That rule no longer applies.

Artificial Intelligence can now generate:

Professional emails with perfect grammar, formatting, and institutional terminology

Official-looking documents, letterheads, logos, and trademark notices with authentic design

Realistic video showing a person making statements, signing documents, or confirming instructions

Convincing audio with correct voice tone, accent, cadence, and emotional inflection

Customized messages that accurately reproduce private information and tailor the message to the recipient

The FBI warned in 2026 that sophisticated scammers use AI-generated deepfake videos to impersonate government officials and claim they can help recover stolen money. The videos showed realistic facial expressions and lip-sync. The victims reported feeling confident the person was real.

THREE-LAYER VERIFICATION STANDARD

1. SENDER IDENTITY

Verify the person or organization independently. Do not click links in the message. Instead, go directly to the agency or company's main website by typing the URL yourself. Call the organization using a phone number you obtain independently, not from the message. Ask whether the communication is legitimate.

2. COMMUNICATION AUTHENTICITY

Check whether the message appears in official systems. If it claims to be a government notice, is it in your official account? If it claims to be from your bank, can you confirm it in your bank's secure portal? Official communications are recorded in official systems.

3. REQUEST REASONABLENESS

Ask whether the request makes sense for this organization and stage. If a new trademark applicant receives a demand for a five-year maintenance fee within days of filing, that is wrong timing. If someone claiming to be from the Social Security Administration demands an iTunes gift card, that is implausible procedure. If a CEO sends an urgent wire-transfer instruction via text, that breaks normal protocol.

THE PROTECTIVE WORKFLOW

PAUSE: Do not respond immediately. Take 15 minutes to verify.

VERIFY: Use independent channels. Call, email, or visit official websites directly. Do not use links in the message.

COMPARE: Match the claim against official records, timing, procedure, and fee schedules.

PROTECT: Do not send money, personal data, credentials, or access until verification is complete.

PRESERVE: Screenshot or save the original message, sender address, timestamps, and URLs for potential reporting.

CASE STUDY SUMMARY: TWO FRAUDS ILLUSTRATING THE TUTORIAL

CASE 1: USPTO TRADEMARK IMPERSONATION

Target:

An applicant who filed six federal trademark applications. The marks included: NI + AI, YOUR BRAIN IS THE BOSS, and AI AND ME: MY SMART HELPER.

What Happened:

Within days of filing (July 2-4, 2026), emails arrived claiming to be from the USPTO. The display name said USPTO.GOV but the actual sender domain was trademark-uspto.com. The emails correctly reproduced the mark names, serial numbers, business entity, and address (all public information). The strongest email claimed the applications had completed examination and demanded $2,300 ($575 per application) for a combined Section 8 filing.

Why This Was Sophisticated:

The $575 amount is a real USPTO fee for a Section 8/15 combined filing. However, that filing is only available after registration during the fifth-to-sixth-year period. Newly filed, unexamined applications cannot owe it. The scammer borrowed an authentic fee amount while removing the legal prerequisites.

Red Flags Detected:

1. The sender domain was not @uspto.gov
2. Official TSDR records showed applications were awaiting examiner assignment (contradicted the email)
3. A legitimate fee amount applied to an impossible procedural stage

Protection Used:

The applicant checked TSDR directly, verified the domain was not official, compared the fee claim to official schedules, reported to the USPTO, and did not pay the $2,300 demand.

CASE 2: CRM PUBLISHING DECEPTIVE MARKETING

Target:

A published author who received unsolicited marketing solicitation.

What Happened:

An email from Victoria Whitmore at CRM Publishing (claimed London location) offered paid advertising in The New York Times Magazine for $2,199 to $15,599. She promised a free $799 publishing package and connected paid visibility to the possibility of traditional publishing and film deals.

Red Flags Detected:

1. No legal entity documentation provided
2. Address inconsistency: UK postcode records showed the location did not match the claimed street
3. No direct confirmation from The New York Times Magazine about reserved advertising
4. Publishing package terms were vague and undefined
5. No prior performance evidence or client testimonials

Protection Used:

The author extracted claims and tested them. She checked UK postcode records, searched Companies House registry, requested written proof before proceeding, and compared the response against industry scam warnings. When verification failed, she declined the offer without sending payment or intellectual property.

KEY LESSON

Professional appearance plus accurate public information plus polished communication does NOT equal authenticity. What stopped both frauds was the decision to pause, verify against independent official sources, and protect sensitive assets until verification was complete.