SPOTTING AI DEEPFAKES

Seeing is no longer believing in the digital age.

Page 1: The Deepfake Threat in Accounting

Deepfakes are no longer a theoretical risk. In the professional services sector, cybercriminals use **Generative Adversarial Networks (GANs)** to clone the voice of a Managing Partner or create a video of a CEO. Their goal is simple: manipulate staff into bypassing security protocols for wire transfers or sensitive data releases.

Detailed Visual & Audio Red Flags

- **The "Turn Test" & Profile Distortions:** Most AI deepfakes are trained on front-facing photos. Ask the person on the video call to turn their head 90 degrees. You will often see the AI "mask" flicker, blur, or detach from the jawline.
- **Unnatural Eye Dynamics:** Look for a lack of "micro-expressions." AI often struggles with realistic blinking patterns and "saccades" (the tiny, rapid movements eyes make when focusing).
- **Audio Anomalies:** Cloned voices often lack emotional inflection. Listen for robotic, flat delivery or strange digital "artifacts" (clicks or hums) that occur during transitions between words.
- **Lighting Inconsistency:** Check if the lighting on the face matches the background. Deepfakes often have a "halo" effect or inconsistent shadows around the ears and hair.

The Psychological Hook

Deepfake attacks almost always rely on **Manufactured Urgency**. The "voice" of the partner will sound stressed, claim to be in a loud environment (to mask audio glitches), and insist that the standard verification process be skipped "just this once."

Page 2: The Human Firewall Checklist

Don't wait for a breach to happen. Use this checklist to harden your firm's defenses against AI-driven social engineering.

Immediate Verification Steps

- [ ] **[ ] The Safe-Word Policy:** Establish a non-digital "Firm Secret Word" known only to staff for verifying high-stakes or unusual requests.
- [ ] **[ ] Out-of-Band Verification:** If you receive a "video" or "voice" request for money or data, hang up and call the person back on a **known, trusted phone number**. Never use the number provided in the call itself.
- [ ] **[ ] Question the Context:** Ask the "caller" a question only the real person would know (e.g., "What was the specific issue we discussed at lunch yesterday?"). AI cannot yet pull real-time, private contextual memories.

Firm-Wide Defensive Actions

- [ ] **[ ] Social Media Hygiene:** Advise Partners to limit high-quality, front-facing video content on public profiles, as these are the "training sets" for deepfake creators.
- [ ] **[ ] Awareness Training:** Conduct quarterly drills where staff are shown "Deepfake vs. Real" samples to sharpen their visual detection skills.
- [ ] **[ ] Wire Transfer Hardening:** Update internal policies to require two-person authorization for any payment over a specific threshold, regardless of who "authorizes" it on video.

CardinalsByte Rule

If a request feels unusual, it **is** unusual. In the age of AI, "Trust but Verify" has been replaced by **"Verify, then Trust."**