



Deepfakes

When Identity Itself Can Be Faked

A Short Scenario

You receive a message from someone you know asking for urgent help – a document, login, or a quick action. The message sounds exactly like them, so you don't question it. You act immediately. Only later do you realize that it wasn't them.

What Are Deepfake Impersonation Attacks?

Deepfakes are no longer limited to fake voices or videos. They now include AI-generated emails, messages, and documents that imitate real people. These attacks replicate:

Writing style

Tone and context

Communication patterns

AI can produce highly realistic communication in seconds, and unlike traditional scams, these messages often contain no obvious errors or warning signs. The result is content that feels natural, familiar, and trustworthy, making it difficult to question.

How These Attacks Typically Work

- Public data (emails, posts, profiles) is collected
- AI learns the target's communication style
- Messages are generated that match tone and context
- Trust is established through familiarity
- Sensitive data or actions are requested

These attacks exploit recognition and trust, not system vulnerabilities.

Where This Risk Appears

- Emails that feel unusually "well-written" and accurate
- Messages that match a known person's tone
- Documents or requests that seem legitimate
- Conversations that feel familiar but slightly unusual

The more real it feels, the harder it is to question.

Pause and Check

Before acting on a request:

- Was this request expected?
- Does the context fully make sense?
- Can it be verified through another channel?

Familiarity should not replace verification.

Real-World Examples

- AI-generated emails are now used in Business Email Compromise (BEC) scams, leading to large financial losses globally
- Attackers have impersonated officials using AI-generated messages and communication to request sensitive information

How to Stay Safe

- Verify sensitive requests independently
- Be cautious of urgent or high-pressure messages
- Do not rely only on writing style or tone
- Limit sharing of personal communication publicly

Trust should always be confirmed.

Key Message

Deepfake attacks no longer need fake voices or videos. Even a simple message can be convincingly artificial. Even if it feels real, verify it anyway.