AP CYBERSECURITY • EXERCISE SHEET

1.4 AI-Based Cybersecurity Attacks

Name: _____ Class: _____ Date: _____ **Total: 11 marks**

Objective

Build the skills to answer exam questions on how adversaries use AI, including **deepfakes** 深度伪造 and AI-written **phishing** 钓鱼.

You must be able to:

- explain how AI augments attacks (deepfakes, flawless phishing, data poisoning)
- describe defenses: a **shared secret** 共享秘密, MFA, verifying AI output

1 Worked examples

Study these first. Each one shows the method for a task used later.

How AI helps attackers

AI can clone a voice or face (a deepfake) to impersonate someone on a call, and can write phishing emails in perfect, native-sounding language - removing the bad grammar that once exposed scams.

Defending against AI attacks

Agree a shared secret word with close contacts to verify identity, enable MFA so a cloned voice cannot log in, and never trust AI output without checking a reliable non-AI source.

2 Practice

2.1 Explain how AI makes phishing emails harder to spot. [2]

2.2 State one defense against a deepfake voice call and explain how it helps. [2]

2.3 Why should you not type sensitive data into a chatbot? [2]

3 Exam-style questions

3.1 An AI clone of a person's voice, used to impersonate them, is called a [1]

- **A** honeypot
- **B** deepfake
- **C** keylogger
- **D** baseline

3.2 Which defense best stops an adversary who has only cloned an employee's voice? [1]

- **A** a longer password
- **B** multifactor authentication
- **C** a brighter screen
- **D** a faster network

3.3 An employee receives an urgent video call from the "CEO" asking to transfer money.

(a) State what kind of attack this could be. [1]

(b) Describe one way the employee could verify the caller's identity. [2]

4 Go further

- work through the **1.4 AI-Based Cybersecurity Attacks** lesson on the **Learn** page;
- read the **Introduction to Security** section of the AP Cybersecurity handout on the **Know** page.

Solutions

2.1 AI writes in perfect, native-sounding language, removing the poor grammar that used to reveal a scam.

2.2 a shared secret question only the two parties know; the deepfake cannot answer it, exposing the fake.

2.3 some AI tools feed user input back into training, and an adversary could later extract that data.

3.1 B. a deepfake clones a voice or face.

3.2 B. a second factor blocks a voice-only attack.

3.3 (a) an AI deepfake impersonation. (b) ask a prearranged shared-secret question that only the real CEO would know; or verify through a separate known channel.