1 AI in attack and defence – Part 2

Name: _____ Class: _____ Date: _____

Vocabulary

Write each word three times.

English	Write it 3 times (English)		
deepfake			
phishing			
shared secret			
MFA			

Recall

- You have seen AI chatbots write fluent text and answer questions.
- You know a photo or video can be edited to look real.
- You understand that some emails try to trick you (phishing).

New ideas

Read these first. Each idea has a short worked example. Then do the Practice below.

■ 1 AI helps attackers

AI lets attackers make a **deepfake** (a fake voice or face) to impersonate someone, and write flawless **phishing** emails in any language. Bad grammar used to expose scams; AI removes that clue.

Worked example. A caller sounds exactly like your manager but is an AI deepfake asking you to transfer money.

2 Defending against AI attacks

Agree a **shared secret** word with close contacts to check identity, and enable **MFA** so a cloned voice alone cannot log in. Never trust AI output without checking a real source.

Worked example. You ask the "manager" a secret question only the real one would know - the deepfake fails.

3 AI helps defenders

A network makes millions of events a day - too many to read. AI sorts likely-malicious events from harmless ones and **alerts** staff in seconds. A human must still review AI advice, because AI only guesses probabilities.

Worked example. AI flags 12 suspicious events out of 20 million and alerts the team instantly.

Watch out: AI is dual-use: the same tool (a language model, code analysis) can help both an attacker and a defender. Read the question to see which side it means.

Practice

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

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

2.3 Why should you never type sensitive data into a chatbot? [2]

2.4 Explain why AI is useful for reading network logs. [2]

2.5 State one task that should be left to a human, not the AI, and why. [2]