LAB 9 - WEEK 11 - 50 POINTS

CTF Challenge: Guide & Worksheet

CIS 340: AI for Cybersecurity - Smart Defense for the Digital Business

Colorado State University - Computer Information Systems

Authorized, sandboxed practice ONLY

This lab uses beginner Capture-the-Flag (CTF) puzzles on Hacker101 or the course CTFd instance - platforms that EXIST to be practiced on legally. You must NOT scan, probe, or attack any other system, website, or network. Doing so is an academic-integrity violation and may be illegal. When in doubt, ask the instructor first.

Name(s): _____ Date: _____ Platform used: [ ] Hacker101 [ ] Course CTFd

1. What This Lab Is

A Capture the Flag (CTF) is a legal, gamified security challenge: you solve puzzles to find hidden 'flags' (short secret strings). CTFs are built for LEARNING on sanctioned systems - never real targets. In this lab you will complete a few beginner challenges, record HOW you solved each, and - the important part - connect each one to a DEFENSE from this week. Thinking like an attacker (safely) is what makes you a better defender.

Work solo or in pairs. Time: about 60-90 minutes. You submit THIS completed worksheet in Canvas - there is no code to hand in.

Learning objectives

- Practice the attacker's mindset in a safe, authorized environment.
- Read instructions and inputs carefully - the core skill behind spotting injection.
- Map each challenge you solve to a concrete defense (guardrails, least privilege, validation).
- Reflect on how CTF puzzles relate to prompt injection and LLM security.

Prerequisites & setup

- A web browser and internet access. No install required.
- Option A - Hacker101 CTF: create a free account at https://www.hacker101.com/ and open the CTF page.
- Option B - Course CTFd: log in to the instance your instructor posts in Canvas.
- Start with challenges marked 'easy' / beginner. Read each challenge's description fully before starting.

Rules of engagement (read before you start)

1) Use ONLY the provided platform. 2) Never attack, scan, or 'test' any other site, app, or device. 3) Do not share flags with other students - share the THINKING instead. 4) If a challenge makes you uncomfortable or seems out of scope, stop and ask. 5) AI tools are NOT permitted for this lab.

2. Guided Steps

1. Log in to Hacker101 or the course CTFd and open the beginner challenge list.
2. Pick a challenge, read its description, and explore. Look carefully at inputs, hidden fields, page source, and any instructions - flags often hide where input is trusted too much.
3. When you capture a flag, immediately write down (in the table below) the challenge name and, in one or two sentences, HOW you found it.
4. For each solved challenge, name ONE defense from this week that would have prevented it (e.g., validate input, check output, least privilege, don't trust user-supplied data).
5. Aim to solve and document at least THREE challenges. Stuck? Use the platform's official hints - that is part of learning.

3. Flag Log (fill in as you go)

Do NOT paste the literal flag string. Record the challenge and your reasoning instead.

#	Challenge name	How I found the flag (brief)	Defense that would stop it
1			
2			
3			
4			

4. Reflection (answer in complete sentences)

These questions are where most of the credit is - they connect the hands-on practice to the week's defensive lessons.

Q1. Describe your favorite challenge. What was the vulnerability, and how did you find the flag?

Q2. For that challenge, what did the system trust that it should not have? Which specific defense (input validation, output checks, least privilege, separating data from instructions) would have stopped it?

Q3. How is a CTF puzzle like a prompt-injection attack on an LLM? (Think about 'the system trusts input it shouldn't.')

Q4. Pick one real business system (e.g., a customer-support chatbot). Name TWO defenses from this week you would apply, and why.

Q5. What surprised you most, and what is one thing you want to learn more about?

5. Submission & Grading

- Submit this completed worksheet (PDF or the Canvas text equivalent) in Canvas by the due date.
- You are graded on PARTICIPATION and REFLECTION quality - not on how many flags you captured.
- See rubric.md for the 50-point breakdown. Late work: -10%/day up to 48 hours (syllabus).
- Reminder: authorized platforms only; AI tools are not permitted.

How this connects to the course

This lab caps Week 11 (prompt injection & LLM security) and builds the attacker's-eye-view that runs through Exam 3 (Weeks 10-13). The reflection - turning 'I broke it' into 'here is how I would defend it' - is the whole point.