

CPTC Homework 3

AI Hacking Task

You are tasked with putting your AI hacking skills to the test with this new challenge. Merlin, a great wizard, has been trained to not give you the password to his vault. Your goal is to convince him by any means necessary to give you this information. This will likely be through trying to obscure your message and his response so that Merlin never knows he is leaking the password.

Your challenges lies below:

<https://hackmerlin.io/>

Remember: Go beyond just what the slides show. They cover basic techniques so I would encourage you to both try to combine multiple of them together. And more importantly encourage you to conduct your own research on potentially new and more complex bypass prompts!

Web Hacking Task

Unfortunately, due to complications with our infrastructure on the SDC, we were not able to get the homework website up and running. Due to us already being over halfway through the week, we decided to cancel the web hacking requirements for this week's homework.

Although there will be no official web homework, if you are interested in learning more on your own I would recommend looking into [Portswigger's learning paths](#). Some of the beginner paths that I can recommend would be Path Traversal, File Upload, SQL Injection, SSRF, and API Testing.

Deliverables – AI ONLY

Create a writeup for each level with Merlin. This would include a screenshot of the prompt used / response and a quick description of your thought process on how the prompt attempts to trick Merlin.