HACKTHEBOX

AI Security Assessment

Report of Findings

HTB Certified Offensive AI Export (HTB COAE) Exam Report

Candidate Name: TODO Candidate Name

PhantomKernel LLC

April 10, 2026

Version: 1.0

Table of Contents

1 Statement of Confidentiality

The contents of this document have been developed by Hack The Box. Hack The Box considers the contents of this document to be proprietary and business confidential information. This information is to be used only in the performance of its intended use. This document may not be released to another vendor, business partner or contractor without prior written consent from Hack The Box. Additionally, no portion of this document may be communicated, reproduced, copied or distributed without the prior consent of Hack The Box.

The contents of this document do not constitute legal advice. Hack The Box's offer of services that relate to compliance, litigation or other legal interests are not intended as legal counsel and should not be taken as such. The assessment detailed herein is against a fictional company for training and examination purposes, and the vulnerabilities in no way affect Hack The Box external or internal infrastructure.

2 Engagement Contacts

PhantomKernel Contacts		
Contact	**Title**	**Contact Email**
TODO Name	TODO Title	TODO Email
TODO Name	TODO Title	TODO Email

Assessor Contact		
Assessor Name	**Title**	**Assessor Contact Email**
TODO Candidate Name	TODO Candidate Title	TODO Candidate Email

3 Assessment Overview

PhantomKernel LLC (“PhantomKernel” herein) invited TODO Candidate Name to perform a targeted Security Assessment of their systems to identify high-risk security weaknesses, assess their impact, document all findings in a clear, professional, and repeatable manner, and provide remediation recommendations.

The objective of this engagement is to identify, analyse, and document security vulnerabilities, safety risks, and misuse vectors affecting the PhantomKernel's AI products, including those that may:

- Cause harm to end users, customers, or the client
- Enable bypass of the intended use cases or operational safeguards
- Allow manipulation of model behavior
- Expose sensitive data or internal system components
- Facilitate unauthorized access, control, or output manipulation

The engagement will focus on both technical and behavioral vulnerabilities within AI systems. However, the following types of activities are considered out of scope for this assessment:

- Physical attacks against the client's properties
- Unverified scanner output
- Any vulnerabilities identified through DDoS or spam attacks
- Vulnerabilities in third-party libraries
- Any theoretical attacks or attacks that require significant user interaction.

3.1 Approach

TODO Candidate Name performed testing under a “Black Box” approach from April 6, 2026, to April 10, 2026. The engagement followed a hybrid methodology combining elements from established security testing frameworks and industry best practices for web applications, APIs, and AI systems

Testing was performed remotely from a non-evasive standpoint, with the goal of uncovering as many misconfigurations and vulnerabilities as possible. Each weakness identified was documented and manually investigated to determine exploitation possibilities and escalation potential.

Identified issues were validated through controlled proof-of-concept exploitation to confirm their feasibility and full impact. Exploits were performed in a manner designed to avoid disruption to production services or access to real customer data.

3.2 Scope

The scope of this assessment was as follows:

TODO: scope

- `www.phantomkernel.htb` : The main company website. The client placed a file **/secret_data.txt** on the filesystem, which TODO Candidate Name was tasked with retrieving as proof of compromise.
- `mcp.phantomkernel.htb` : PhantomKernel's MCP server, which has not been released and is thus not yet publicly accessible. However, TODO Candidate Name was able to access it by exploiting vulnerabilities documented in this report.

- `staging.phantomkernel.htb` : PhantomKernel's staging environment for their AI products. All AI systems hosted in the staging environment were considered in scope.

URL	Description
TODO http://www.phantomkernel.htb:PORT	Main PhantomKernel Website
TODO http://mcp.phantomkernel.htb:PORT	PhantomKernel MCP Server (not yet released and thus protected from unauthorized access)
TODO http://staging.phantomkernel.htb:PORT	Staging Environment for AI Systems

4 AI Security Assessment Summary

4.1 Summary of Findings

During the course of testing, TODO Candidate Name uncovered a total of 3 findings that pose a material risk to PhantomKernel's information systems. The below chart provides a summary of the findings by severity level.

In the course of this penetration test **2 High** and **1 Info** vulnerabilities were identified:

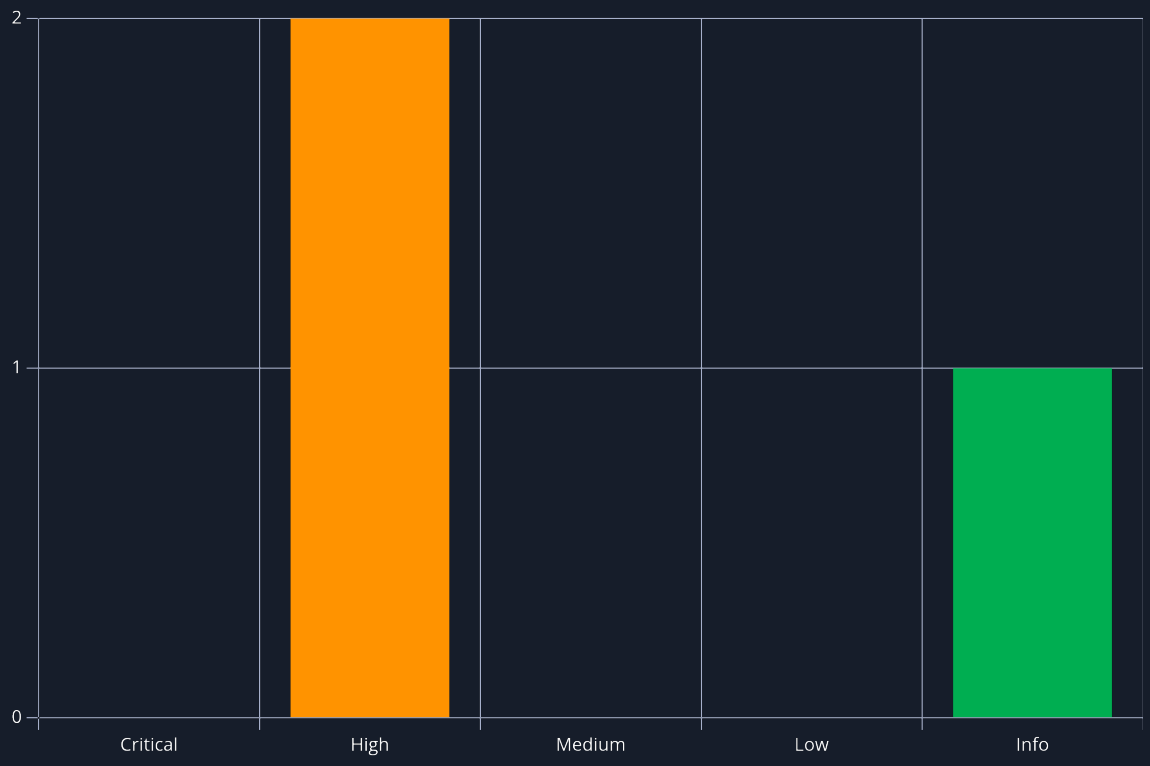

Figure 1 - Distribution of identified vulnerabilities

Below is a high-level overview of each finding identified during testing. These findings are covered in depth in the Technical Findings Details section of this report.

#	Severity Level	Finding Name	Page
1	8.2 (High)	Finding 1	9
1	8.2 (High)	Finding 1	10
1	0.0 (Info)	Finding 1	8

5 Findings in mcp.phantomkernel.htb

1. Finding 1 - Info

CWE	-
CVSS 3.1	N/A
Description (Incl. Root Cause)	TODO DESCRIPTION
Security Impact	TODO IMPACT
Affected Domain	mcp.phantomkernel.htb
Remediation	TODO REMEDIATION
References	-

Detailed Walkthrough

TODO Candidate Name performed the following:

1. TODO: LIST HIGH LEVEL STEPS

```
ADD COMMAND OUTPUT AS APPROPRIATE
```

TODO ADD SCREENSHOTS AS APPROPRIATE

Remediation Summary

TODO: remediation summary

Short Term

TODO

Medium Term

TODO

Long Term

TODO

6 Findings in staging.phantomkernel.htb

1. Finding 1 - High

CWE	CWE-522 - Insufficiently Protected Credentials
CVSS 3.1	8.2 / CVSS:3.1/AV:N/AC:L/PR:N/UI:N/S:U/C:H/I:L/A:N
Description (Incl. Root Cause)	TODO DESCRIPTION
Security Impact	TODO IMPACT
Affected Domain	staging.phantomkernel.htb
Remediation	TODO REMEDIATION
References	-

Detailed Walkthrough

TODO Candidate Name performed the following:

1. TODO: LIST HIGH LEVEL STEPS

```
ADD COMMAND OUTPUT AS APPROPRIATE
```

TODO ADD SCREENSHOTS AS APPROPRIATE

Remediation Summary

TODO: remediation summary

Short Term

TODO

Medium Term

TODO

Long Term

TODO

7 Findings in www.phantomkernel.htb

1. Finding 1 - High

CWE	-
CVSS 3.1	8.2 / CVSS:3.1/AV:N/AC:L/PR:N/UI:N/S:U/C:H/I:L/A:N
Description (Incl. Root Cause)	TODO DESCRIPTION
Security Impact	TODO IMPACT
Affected Domain	www.phantomkernel.htb
Remediation	TODO REMEDIATION
References	-

Detailed Walkthrough

TODO Candidate Name performed the following:

1. TODO: LIST HIGH LEVEL STEPS

```
ADD COMMAND OUTPUT AS APPROPRIATE
```

TODO ADD SCREENSHOTS AS APPROPRIATE

Remediation Summary

TODO: remediation summary

Short Term

TODO

Medium Term

TODO

Long Term

TODO

A Appendix

A.1 Flags Discovered

Flag #	Host	Flag Value	Flag Location	Method Used
1.	TODO www.phantomkernel.htb - Flag 1			
2.	TODO www.phantomkernel.htb - Flag 2			
3.	TODO www.phantomkernel.htb - Flag 3 (`/secret_data.txt`)			
4.	TODO mcp.phantomkernel.htb - Flag			
5.	TODO staging.phantomkernel.htb - TicketFlow Flag			
6.	TODO staging.phantomkernel.htb - OmniDigit Flag			
7.	TODO staging.phantomkernel.htb - PassPort Flag			

End of Report

This report was rendered by SysReptor with