

Project Proposal Document

**PAPILLON**

Kerem ELMA  
 Mehmet ESKİ  
 Sabri Mert PİŞKİN  
 Mete Cem TURAN

**1. INTRODUCTION**

This document presents two project proposals developed by the PAPILLON group for the BİL495 Graduation Project. The first proposal aims to develop an AI-driven cybersecurity framework, while the second focuses on an intelligent system that automatically generates notes from lecture videos. For each proposal, this document details the problem definition, project objectives, scope, target user profile, foreseen challenges, and the data sources to be utilized.

The project management and development lifecycle will be structured in accordance with the principles outlined in the ISO/IEC/IEEE 12207 standard for software lifecycle processes to ensure a systematic approach to planning, execution, and quality assurance.

Github link:<https://github.com/SabriMertPiskin/Papillon>

Trello link:<https://trello.com/b/WtLSndob/papillon-project-board>

| Term | Description |
| --- | --- |
| Honeypot | A decoy computer system set up to attract and trap attackers, gathering information about their methods and activities. |
| Zero-Day | A vulnerability in a software or hardware that has been discovered but not yet patched by the vendor. |
| Speech-to-Text | A technology that converts spoken language into written text. |

| Abbreviation | Description |
| --- | --- |
| AI | Artificial Intelligence |
| ML | Machine Learning |
| DL | Deep Learning |
| NLP | Natural Language Processing |
| UI | User Interface |

| Acronym | Full Expansion |
| --- | --- |
| CTI | Cyber Threat Intelligence |
| SOC | Security Operations Center |
| CVE | Common Vulnerabilities and Exposures |
| SIEM | Security Information and Event Management |
| IDS | Intrusion Detection System |

## 

### **1. Project Title**

AI-Driven Cybersecurity Framework for Intelligent Threat Detection and Analysis

### **2. Problem Definition & Objectives**

● **Problem Definition:** A Cyber Threat Intelligence (CTI) platform enables organizations to monitor, analyze, and respond to cyber attacks in real time. It tracks which assets, servers, ports, or endpoints are exposed and identifies existing vulnerabilities (CVEs) that require immediate attention. By linking attacks to their place of origin, the platform can identify patterns and trends in global threats. CTI combines raw data and actionable intelligence by integrating threat feeds, vulnerability databases, honeypots, and security logs. For instance, it can notify security teams of suspicious activity, forecast possible attack vectors, and suggest particular mitigations.  
With cyber threats in the form of malware, phishing attacks, and social engineering attacks based on deepfakes increasing exponentially, security operations performed manually are no longer sufficient. SOCs are bogged down by alert fatigue, response lag time, and the volume of threat intelligence data. Zero-day vulnerability detection and simulating real attacks to verify defense systems also remain very complex and time-consuming.  
 This project will develop an integrated AI-powered Cyber Threat Intelligence (CTI) platform that automates different elements of cyber defense; including malware classification, phishing detection, intrusion detection, vulnerability simulation, and password strength analysis. With the incorporation of honeypot data, CVE analysis, and real-time threat feeds, the platform enhances the SOC's situational awareness and decision-making.

● **Objectives:**

o Develop an AI-driven CTI platform integrating multiple cybersecurity components.

o Implementing a centralized threat intelligence and honeypot system to collect and analyze realtime attack data.

o Utilizing NLP to assess and prioritize the latest CVE data for vulnerability management.

o Building AI models for:

❖ Malware detection and classification

❖ Phishing detection

❖ Intrusion detection system

❖ Password strength estimation

o Simulating various vulnerable attacks in a controlled lab environment

o Creating a SOC assistant to summarize and visualize security alerts.

o Integrating all components into a unified platform that visualizes security insights and improves threat response efficiency.

o Evaluating the system’s effectiveness through accuracy, detection rate, and false/positive metrics against benchmark datasets.

### **3. Scope**

### ● Preprocessing and gathering cybersecurity data such as network traffic logs, phishing emails, malware samples, password hashes, deepfake images and CVE records.

### ● Applying machine learning and deep learning models to detect anomalies and classify threats.

### ● Designing a centralized dashboard to visualize, report, and monitor detected threats in real-time.

### ● Having a honeypot and threat intelligence module to carry out simulated attacks and gather behavior data.

### ● Developing a smart SOC assistant which processes alerts and facilitates decision making.

### ● Creating a cyber-attack simulation laboratory environment.

**3.1 Out of Scope:**

* Developing automated remediation actions (e.g., blocking, deleting, or quarantining) for detected threats. The system's focus will be on detection, analysis, and reporting.
* The deployment and performance testing of the developed tools on live, large-scale enterprise networks. The system will be tested in a controlled laboratory environment with simulated data.
* The development of cybersecurity solutions that require specialized hardware (e.g., FPGAs, custom servers).
* Transforming the developed platform into a commercial product or preparing a corresponding UI/licensing infrastructure.

### **4. User Profile**

● **Target Users:**

· **Security Operations Centers (SOCs):** For real-time monitoring, automated alert analysis, and decision support through the AI SOC assistant.

· **Cybersecurity Analysts:** To analyze network traffic, detect phishing campaigns, and identify malware using AI-enhanced tools.

· **System and Network Administrators:** To proactively detect anomalies and mitigate potential intrusions in enterprise networks.

· **Threat Intelligence Teams:** To collect and interpret threat data from honeypots and open-source intelligence feeds.

· **Researchers and Academic Institutions**: For studying AI applications in cybersecurity, developing improved detection algorithms and simulating scenarios.

### **5. Anticipated Challenges & Constraints**

● **Technical hurdles**

o Building an end-to-end AI-based cybersecurity system is a problem that involves multiple complex technical challenges. Developing stable machine learning and deep learning models relies on good-quality and well-balanced data, which may be tough for obtaining in terms of particular types of attacks. Integrating multiple modules — such as network intrusion detection, phishing detection, and threat intelligence — into one platform also involves interoperability, data synchronization, and scalability issues of the system. Further, supplying real-time performance under heavy data loads is computationally expensive, requiring careful optimization of the algorithms and model architectures.

● **Time/resources limits**

o The project will be built over an academic semester, which restricts the time spent on large-scale data collection, fine-tuning of the model, and comprehensive testing. Computational capacity, including GPU time for training deep learning models, can also limit experimentation scale. Additionally, running the system in a realistic enterprise setting will not be possible because of infrastructure and security constraints. Consequently, all testing will have to be done within simulated or controlled settings using publicly accessible datasets.

● **Ethical or regulatory considerations**

o As the project involves gathering and analyzing cybersecurity data, ethical handling of potentially sensitive information is critical. All datasets used will be either publicly available or anonymized to ensure data confidentiality and ethical standards of research. The system will not engage in or imitate any unauthorized penetration or data exploitation activities. In addition, respect for cybersecurity research ethics and relevant regulatory codes such as GDPR and institutional research regulations will be maintained throughout the project's lifespan

### **6. Data Sources**

### **●** **CICIDS2017, which includes modern attack types [2], will be used for anomaly detection.**

### **o** **Benchmark intrusion detection datasets containing realistic network traffic, including DoS, brute force, botnet, and web attacks[13].**

### **●** **NSL-KDD, an improved version of the KDD Cup 99 dataset [1], will be used for anomaly detection.**

### **o** **A widely used dataset for evaluating intrusion detection systems, useful for testing baseline models[12].**

### **●** **Custom Traffic Simulations:**

### **o** **The team will use tools like Wireshark and Metasploit to generate controlled attack traffic in a virtual lab environment.**

### **●** **Deepfake and Real Images:**

● This dataset contains a large collection of both authentic and AI-generated (deepfake) facial images. It will be used to **train and validate the deepfake detection model**, enabling the system to distinguish between real and manipulated images with high accuracy[3].

### **●** **Enron Fraud Dataset:**

● This dataset consists of approximately 450,000 emails from the Enron Corporation, labeled for fraudulent activity. It will be utilized to train and evaluate the phishing detection model, enabling the system to identify and classify deceptive email communications indicative of corporate fraud schemes[4].

### **●** **Classification of Malwares:**

● This dataset contains features extracted from the header fields of Portable Executable (PE) files. It is designed for **malware detection and classification**, allowing models to differentiate between various types of malicious PE files based on their structural characteristics[5].

### **●** **Password Strength:**

● This dataset contains a large collection of real-world passwords. It will be used to train and evaluate the AI-based password strength auditor, helping the model distinguish between weak and strong passwords based on structural patterns and common usage[6][7][8].

### **●** **Real-Time Stream Data**

### **o** If feasible, anonymized network telemetry from test environments or honeypots may be incorporated to evaluate model performance in dynamic, real-world-like conditions.