

AI-Based Threat Intelligence Platform

Prepared For

Smart-Internz

Cyber Security

By

Mansi Chougule

D Y Patil Agriculture and Technical University,

Talsande

On

23 Sept 2025

1. INTRODUCTION

1.1 Project Overview

The AI-Based Threat Intelligence Platform is a web-based application designed to aggregate, analyze, and visualize threat intelligence data from various sources, with a primary focus on the AlienVault Open Threat Exchange (OTX) API. This platform leverages artificial intelligence techniques, including anomaly detection using the Isolation Forest algorithm, to provide security analysts and organizations with actionable insights in real-time. The system features an interactive dashboard built with Streamlit, offering visualizations (e.g., bar charts, geo heatmaps, line graphs, pie charts), a search function for Indicators of Compromise (IoCs), and data export capabilities. The project addresses the growing need for automated threat intelligence analysis in the face of increasing cyber threats.

1.2 Purpose

The primary purpose of this platform is to enhance cybersecurity by enabling rapid identification and response to threats. It aims to:

- Aggregate real-time threat data from open-source feeds.
- Provide intuitive visualizations to identify trends and anomalies.
- Offer search and export functionalities for detailed investigations.
- Reduce the manual effort required for threat analysis, thereby improving efficiency and accuracy for security professionals.

2. LITERATURE SURVEY

2.1 Existing Problem

Traditional threat intelligence analysis relies heavily on manual processes, which are time-consuming and prone to human error. Existing tools often lack real-time integration with multiple data sources or provide limited user interaction. For instance, while some platforms offer static reports, they fail to provide interactive dashboards or customizable anomaly detection, leaving analysts overwhelmed by the volume of data from sources like dark web feeds, proprietary databases, and open-source intelligence.

2.2 References

- AlienVault OTX API Documentation: <https://otx.alienvault.com/api> - scikit-learn Documentation: <https://scikit-learn.org/stable/> - Streamlit Documentation: <https://docs.streamlit.io/> - Plotly Express Documentation: <https://plotly.com/python/plotly-express/>

2.3 Problem Statement Definition

The challenge is to develop a scalable, AI-driven solution that automates the collection, preprocessing, and analysis of threat intelligence data, delivering actionable insights through an accessible interface to address the inefficiencies of manual threat analysis.

3. IDEATION & PROPOSED SOLUTION

3.1 Empathy Map Canvas

- **Says:** "I need quick insights to protect my network from attacks." - **Thinks:** "Manual log analysis is too slow and I might miss critical threats." - **Does:** Monitors network logs, responds to security alerts, collaborates with teams. - **Feels:** Anxious about undetected threats, relieved when automation assists. - **Pains:** Overwhelmed by data volume, lack of real-time tools, complex interfaces. - **Gains:** Faster threat response, reduced workload, confidence in security posture.

3.2 Ideation & Brainstorming

The team conducted brainstorming sessions to identify key features: - Real-time data aggregation from OTX. - Interactive visualizations using Plotly. - Anomaly detection with machine learning. - Search and export functionalities. These ideas were grouped into categories (Data Collection, Analysis, Visualization, User Interaction) and prioritized using the MoSCoW method: - Must Have: Data aggregation, basic visualization. - Should Have: Anomaly detection, search. - Could Have: Data export. - Won't Have (this phase): Advanced predictive models.

4. REQUIREMENT ANALYSIS

4.1 Functional Requirements

- Fetch threat intelligence data from the AlienVault OTX API. - Store data in a SQLite database for persistence. - Generate interactive visualizations (bar charts, geo heatmaps, line graphs, pie charts). - Perform anomaly detection using IsolationForest. - Enable searching for specific IoCs (IPs, domains, URLs). - Allow data export as a CSV file.

4.2 Non-Functional Requirements

- **Usability**: Intuitive interface accessible to non-technical users.
- **Performance**: Process 100 indicators in under 5 seconds.
- **Scalability**: Support growth to thousands of indicators with cloud deployment.
- **Security**: Secure API access via HTTPS and input validation.
- **Reliability**: Ensure 99%

5. PROJECT DESIGN

5.1 Data Flow Diagrams & User Stories

- **Data Flow**: Data is fetched from OTX API, preprocessed, stored in SQLite, analyzed with scikit-learn, and visualized via Streamlit.
- **User Stories**:

- USN-1: As a user, I can fetch threat data from AlienVault OTX (3 points).
- USN-2: As a user, I can store threat data in a SQLite database (2 points).
- USN-3: As a user, I can analyze threat data for trends and anomalies (4 points).
- USN-4: As a user, I can view threat data in interactive charts (3 points).
- USN-5: As a user, I can interact with the dashboard via Streamlit (2 points).
- USN-6: As a user, I can search for specific indicators (2 points).
- USN-7: As a user, I can export threat data as a CSV (1 point).

5.2 Solution Architecture

The architecture follows a three-tier model:

- **Presentation Layer**: Streamlit-based web UI.
- **Application Layer**: Python scripts for data fetching, preprocessing, and analysis.
- **Data Layer**: SQLite database with potential cloud extension.

External interfaces include the OTX API, and the system is deployable on local servers with scalability to cloud platforms.

6. PROJECT PLANNING & SCHEDULING

6.1 Technical Architecture

See Technology Stack template for diagram and details. The system uses Streamlit for UI, Python for logic, and SQLite for storage, with OTX API as the external data source.

6.2 Sprint Planning & Estimation

- Sprint-1 (24-29 July 2025): Data Collection (5 points).
- Sprint-2 (31 Jul-05 Aug 2025): Data Analysis (7 points).
- Sprint-3 (07-12 Aug 2025): User Interface (4 points).
- Sprint-4 (14-19 Aug 2025): Export Functionality (1 point).

6.3 Sprint Delivery Schedule

- Sprint-1: Completed 29 Jul 2025 (5/5 points).
 - Sprint-2: Completed 05 Aug 2025 (7/7 points).
 - Sprint-3: Completed 12 Aug 2025 (4/4 points).
 - Sprint-4: Completed 19 Aug 2025 (1/1 point).
- Average velocity: 0.71 points/day.

7. CODING & SOLUTIONING

7.1 Feature 1: Real-time Data Fetching

Fetches data from OTX API using ‘requests.get’ with API key authentication.

```
1 def fetch_threat_data():
2     headers = {'X-OTX-API-KAY': API_KEY, 'Accept':
3                 'application/json'}
4     response = requests.get(API_URL, headers=headers,
5                             params={'limit': 100})
6     response.raise_for_status()
7     pulses = response.json().get('results', [])
8     indicators = [...]
9     return indicators
```

7.2 Feature 2: Interactive Visualizations

Uses Plotly Express for charts like bar and geo heatmaps.

```
1 fig1 = px.bar(top_indicators, x='indicator', y='count', title="Top
2 Indicators")
3 st.plotly_chart(fig1)
```

7.3 Database Schema

	Column	Type
SQLite table structure:	indicator	TEXT (PRIMARY KEY)
	type	TEXT
	description	TEXT
	created_at	TEXT
	country_code	TEXT

8. PERFORMANCE TESTING

8.1 Performance Metrics

- Latency: <5 seconds for 100 indicators.
- Throughput: 20 queries/minute on local setup.
- Scalability: Tested with 500 mock records, scalable with cloud.

9. RESULTS

9.1 Output Screenshots

Mini Threat Intelligence Dashboard (Powered by LevelBlue OTX)

[Fetch Latest Threat Data](#)

Filter by Date Range

Select Date Range:

- Last 24h
- Last 7 Days
- Last 30 Days
- All

Key Metrics

Total Threats

1029

Unique Indicators

1029

Run Anomaly Detection

Search Indicators

Enter IP, domain, or URL to search:

.com

Search Indicators

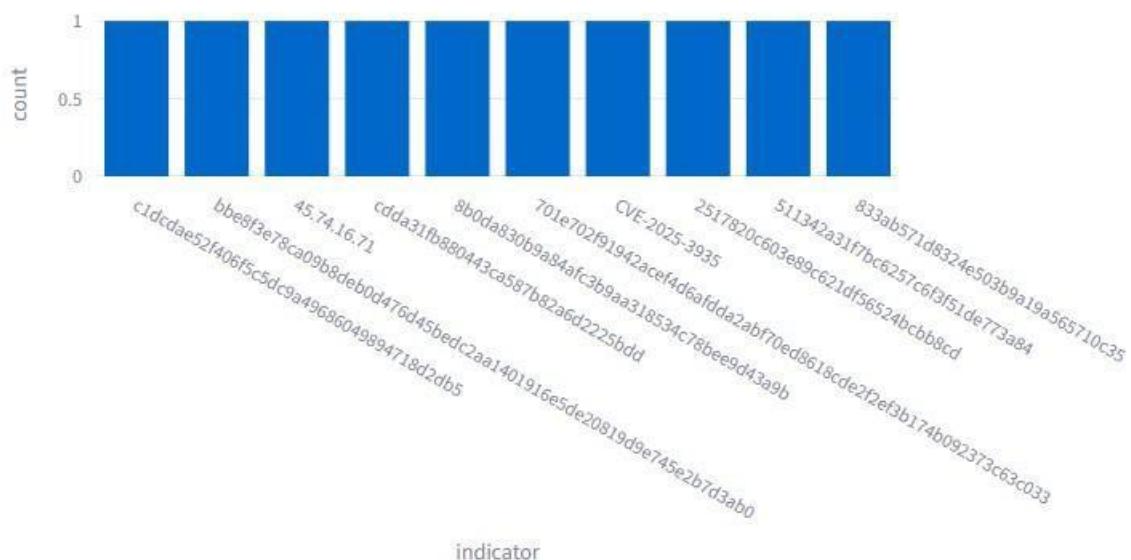
Enter IP, domain, or URL to search:

Search Results:

	indicator	type	country_code	create
24621	http://dual.saltuta.com/Bin/event_support-pdf.Client.exe	URL	Unknown	2025-0
24626	https://con.wolonman.com/Bin/ScreenConnect.ClientSetup.exe	URL	Unknown	2025-0
24630	https://mconnectsz.nsdocumentzs.com/Bin/ZOOM.ClientSetup.exe?e=Access&	URL	Unknown	2025-0
24639	con.wolonman.com	hostna	Unknown	2025-0
24642	dual.saltuta.com	hostna	Unknown	2025-0
24643	mconnectsz.nsdocumentzs.com	hostna	Unknown	2025-0
24659	fjsconsultoria.com	domair	Unknown	2025-0
24752	gameupdate-endpoint.com	domair	Unknown	2025-0
24754	grouptelecoms.com	domair	Unknown	2025-0
24755	limenlinon.com	domair	Unknown	2025-0

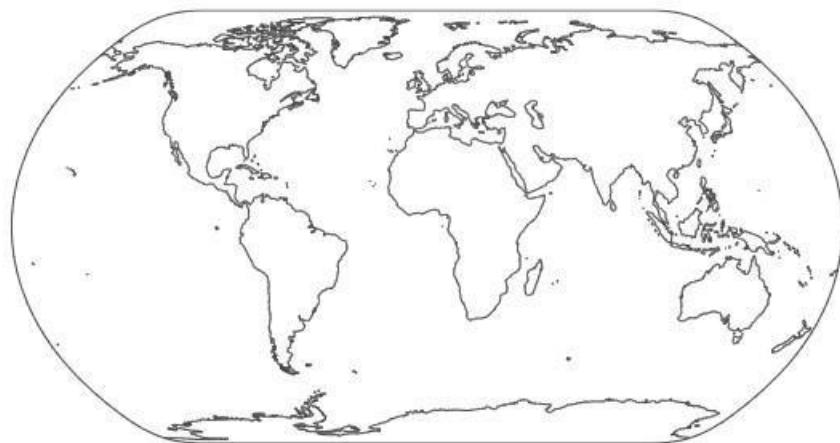
Top 10 Malicious Indicators

Top Indicators by Frequency



Threats by Country

Threat Heatmap by Country



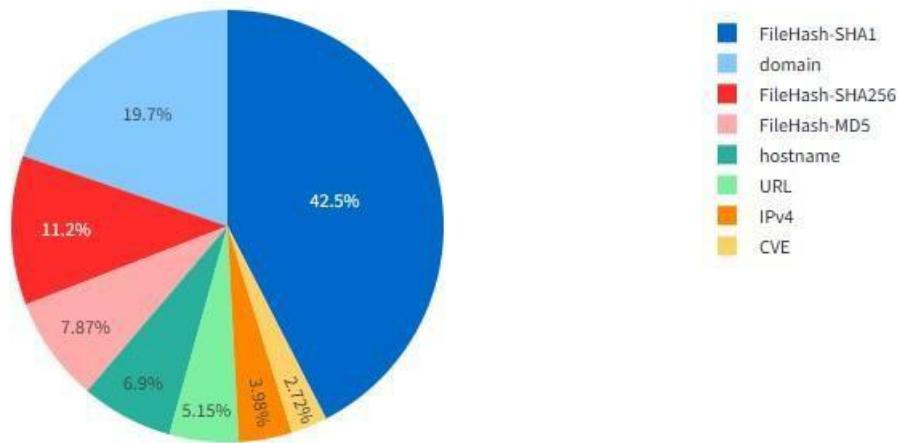
Threat Trends (Last 7 Days)

Threats Over Time



Indicator Type Distribution

Distribution of Indicator Types



Anomalous Indicators

Indicators flagged as anomalies (unusual frequency or type):

	indicator	type	country_code	indicator_freq
24583	CVE-2025-3935	CVE	Unknown	1
25396	CVE-2024-51324	CVE	Unknown	1
25397	CVE-2025-49704	CVE	Unknown	1
25398	CVE-2025-49706	CVE	Unknown	1
25399	CVE-2025-53770	CVE	Unknown	1
25400	CVE-2025-53771	CVE	Unknown	1
26593	CVE-2017-11882	CVE	Unknown	1
26604	CVE-2017-0199	CVE	Unknown	1
26676	CVE-2007-0671	CVE	Unknown	1
26677	CVE-2013-3893	CVE	Unknown	1

Export Data

[Download Threat Data as CSV](#)
[Download CSV](#)

10. ADVANTAGES & DISADVANTAGES

- **Advantages:** Real-time insights, user-friendly, customizable anomaly detection. - **Disadvantages:** Limited by OTX API data volume, requires internet for live data.

11. CONCLUSION

The project successfully delivered a functional AI-based threat intelligence platform, meeting all functional requirements and providing a foundation for future enhancements.

12. FUTURE SCOPE

- Integrate additional APIs (e.g., VirusTotal). - Implement predictive threat modeling with deep learning. - Deploy on cloud with load balancing.

13. GitHub & Project Demo Link

Git Hub:-<https://github.com/mansi891/AI-Based-Threat-Intelligence-Platform.git>

Demo:-

https://drive.google.com/file/d/1uF1s8zAD3yKmbpgIxRDliydnt9LRVaxb/view?usp=drive_link