Phishing Website Detector

A Multi-Source Threat Intelligence Approach Using Streamlit

Presented by : **Gagan shrivas**

Institution: Digisuraksha Parhari Foundation Internship Program

Date: May 2025

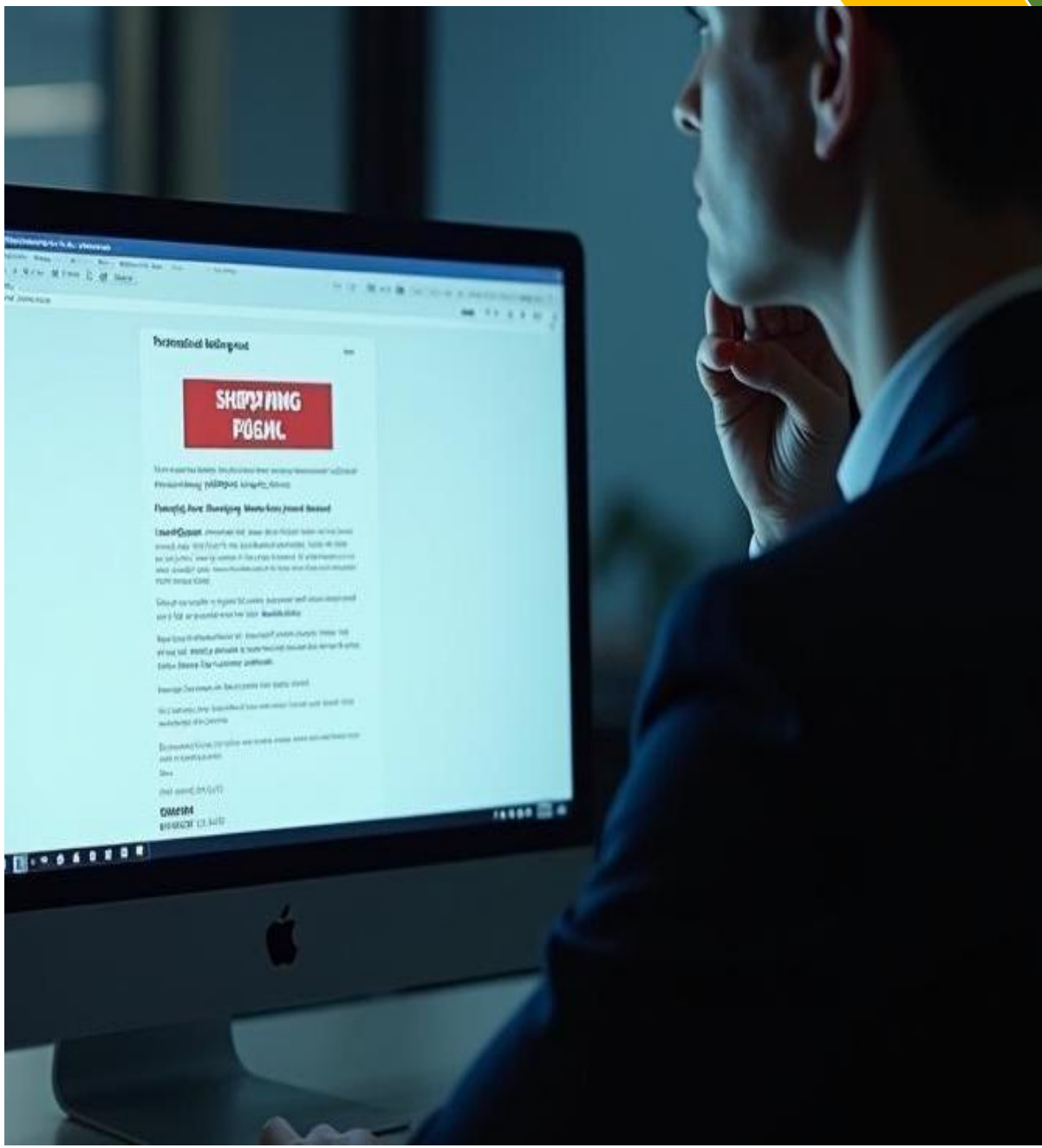
Team Members : Gagan shrives, Yash Yadav, Vishvesh Gopal

Team Code Verse

DigiSuraksha
Parhari Foundation

Abstract

- Phishing is a leading cyber threat, targeting users and organizations worldwide.
- This project presents an open-source tool that analyzes website URLs for phishing risk using real-time threat intelligence, technical analysis, and community feedback.
- Built with Python and Streamlit, the tool empowers users to make safer decisions online.



Problem Statement

- Phishing attacks are increasingly sophisticated, bypassing traditional security tools.
- Many existing solutions are proprietary, slow to update, or not user-friendly.
- There is a need for a transparent, real-time, and accessible phishing detection tool.

Objectives

Goals and Objectives:

- Develop a user-friendly, open-source phishing website detector.
- Combine technical analysis, multiple threat intelligence feeds, and community reporting.
- Provide actionable, real-time risk assessments for any website URL.

Literature Review

- Blacklists: Fast but can miss new threats (e.g., PhishTank, OpenPhish, URLhaus).
- Heuristic/Content Analysis: Looks for suspicious patterns, keywords, and forms.
- Machine Learning: Uses features like URL length, domain age, and content structure.
- Community Reporting: Crowdsources threat intelligence for better accuracy.
- *Key references: Sahoo et al. (2017), Aburrous et al. (2010), OWASP Phishing Guide.*

Research Methodology

- Data Sources: PhishTank, OpenPhish, URLhaus, VirusTotal, user reports.

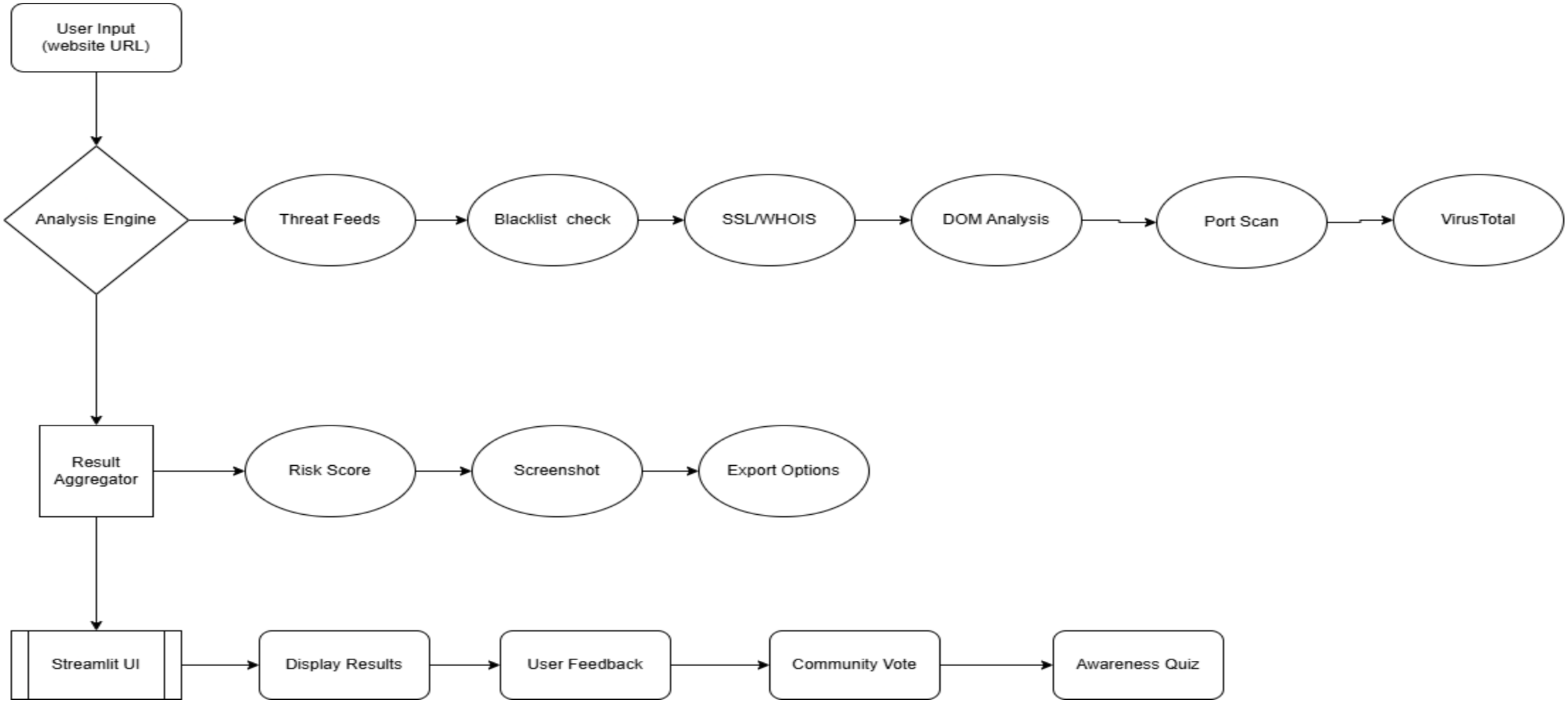
- Technical Analysis:
  - SSL & WHOIS checks
  - Directory and file crawling
  - DOM analysis for forms and keywords
  - Port scanning and service detection

- User Interface: Built with Streamlit for accessibility and rapid prototyping.

- Community Feedback: Collects user votes and comments.

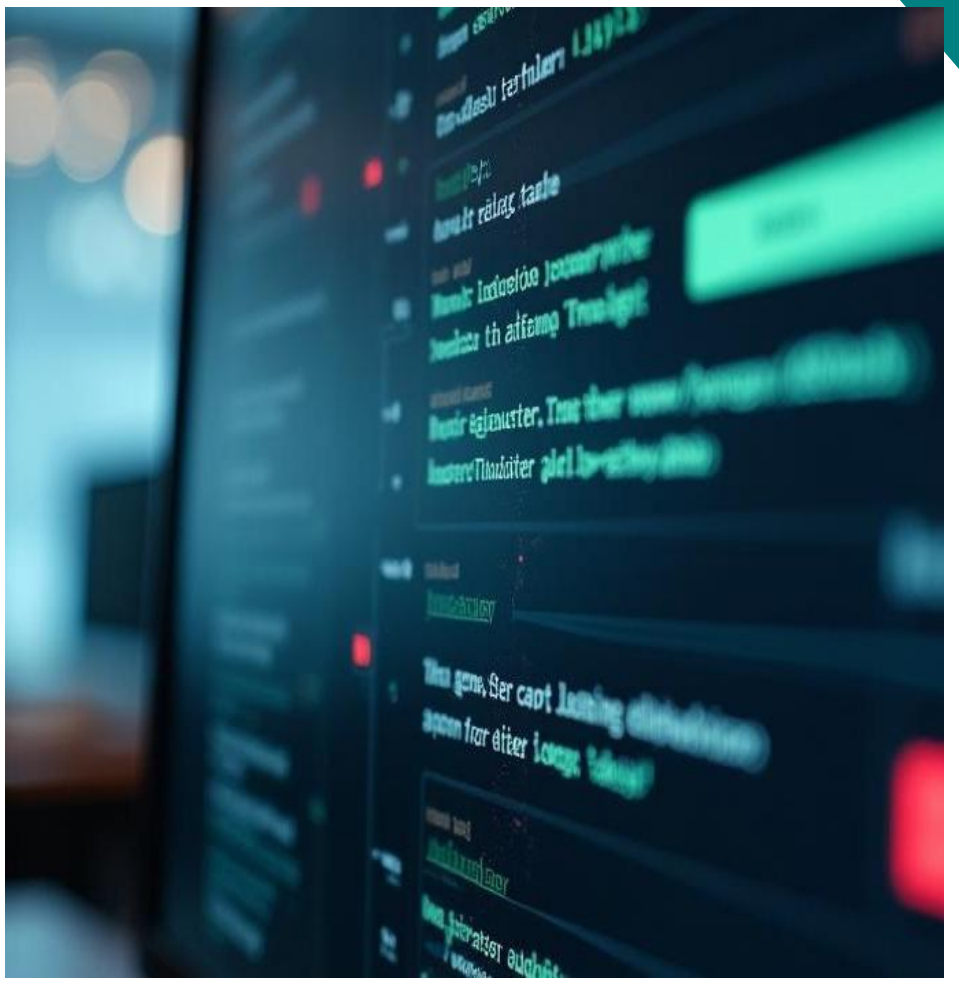
Objectives

Block Diagram:



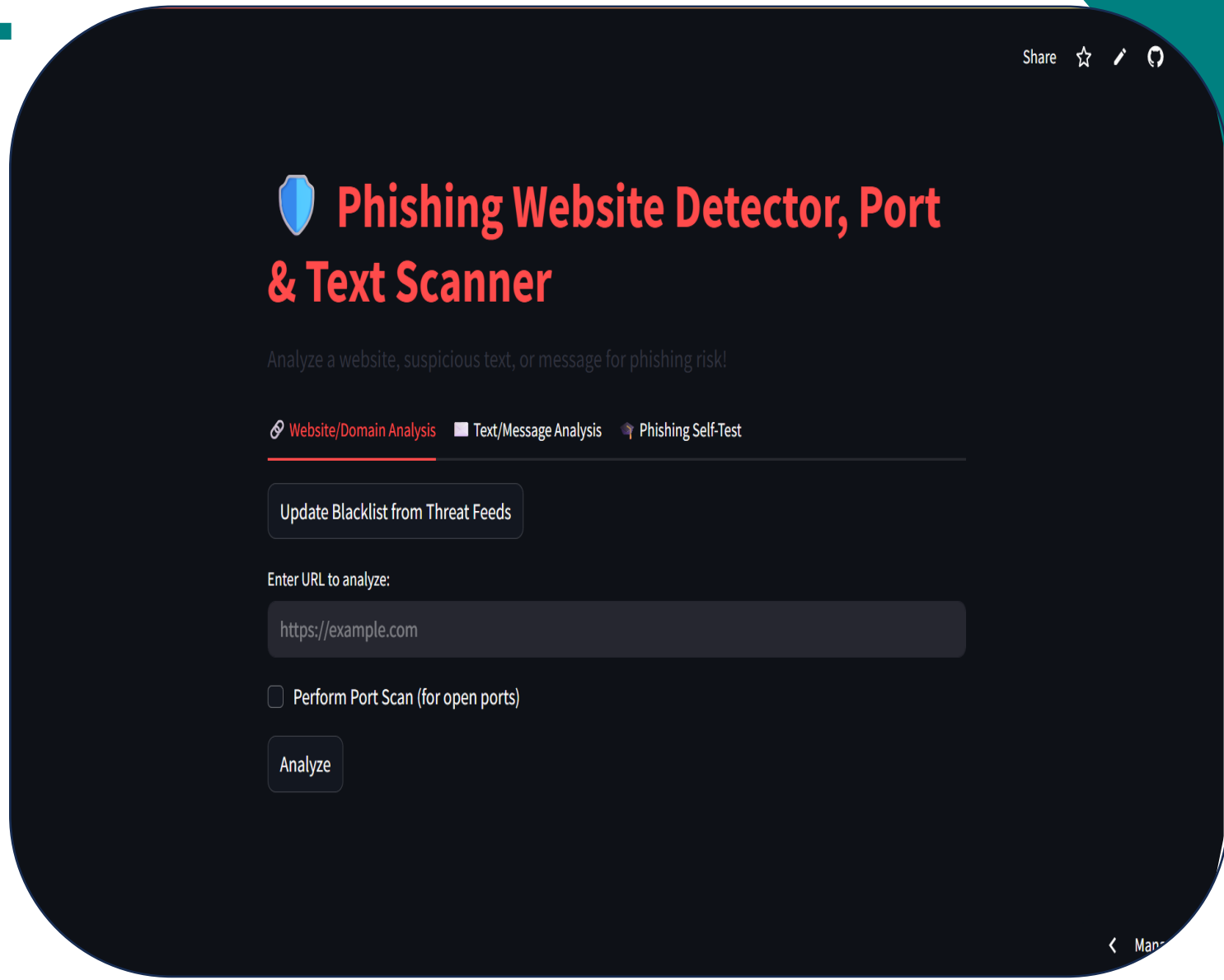
Key Features

- Multi-feed blacklist updates (PhishTank, OpenPhish, URLhaus)
- SSL certificate and WHOIS verification
- Content and DOM analysis (login forms, suspicious keywords)
- Port scanning with service detection
- Community reporting and phishing awareness quiz
- Data export for further analysis



Tool Implementation

- Technologies: Python 3.8+, Streamlit, Selenium, BeautifulSoup, requests, etc.
- Design: Modular, open-source, easy to extend.
- User Interface: Clean, modern, and responsive.



Ethical Impact & Market Relevance

Ethical Impact:

- Empowers users to assess web threats
- Promotes transparency and community-driven security
- No user data is shared externally

Market Relevance:

- Addresses need for accessible, open-source phishing detection
- Useful for individuals, organizations, and security training

Future Scope

- Integrate machine learning for zero-day phishing detection
- Develop a browser extension for real-time warnings
- Integrate with SIEM/SOAR platforms for automated reporting
- Build a mobile app version
- Add more threat intelligence feeds and multi-language support

Demo Video

Watch the demo:

- https://youtu.be/syWq5gx05mE

References

- ✓ FhishTank
- ✓ OpenPhish
- ✓ URLhaus
- ✓ VirusTotal
- ✓ Sahoo, D. et al. (2017). Malicious URL Detection using Machine Learning. IEEE.
- ✓ Aburrous, M. et al. (2010). Intelligent phishing detection system for e-banking. ESWA.
- ✓ OWASP Phishing Guide
- ✓ CERT Phishing Resources
- ✓ Streamlit Documentation
- ✓ Selenium Documentation

Thank you for your attention!

Contact: github.com/gagan-long

Questions?