Trinetra: AI Web Guardian

Real-Time AI-Based Browser Security Extension

Pradumon Sahani

Class 12 Student | Cybersecurity & AI Enthusiast

June 2, 2025

Author Contact:
Email: pradumon14@gmail.com
GitHub: github.com/pradumon14
Project Repository: github.com/pradumon14/trinetra

Abstract

Trinetra is a real-time AI-powered Chrome extension that protects users while Browsing the web by scanning site content and behavior using Google's Gemini 1.5 Flash model. It proactively detects phishing, malware, deceptive downloads, and suspicious scripts — providing alerts and actionable insights directly within the browser. This whitepaper presents Trinetra's motivation, architecture, features, limitations, and future potential, especially if integrated as a native browser capability.

1 Introduction

Modern phishing and malware delivery techniques have become more advanced, using generative AI to bypass static filters and deceive users. Existing browser defenses rely heavily on known blacklists and outdated heuristics. As a response, Trinetra leverages a live AI model to provide intelligent, context-aware site evaluation.

This extension was built to demonstrate how real-time AI threat scanning — including full DOM analysis, JavaScript inspection, and download behavior monitoring — can be practically integrated into browsers for next-gen protection.

---

2 Purpose

Trinetra aims to protect users by:

- Extracting and analyzing website text, structure, scripts, and network behaviors.
- Using Google's Gemini 1.5 Flash API to detect suspicious activity.
- Returning AI-evaluated site status (SAFE, SUSPICIOUS, DANGEROUS) with explanations.
- Empowering users to control their own API access and privacy via a personal API key.

---

3 Core Functionalities

1. **Content & DOM Extraction**
   `content.js` collects:
   - Page title and visible text.
   - HTML form actions and script URLs.
   - iframe sources and external links.
2. **Network Monitoring**
   `chrome.webRequest` observes form submissions and redirects.
3. **Download File Protection**
   `chrome.downloads.onCreated` triggers alerts for risky file types (e.g., `.exe`, `.js`, `.docm`).
4. **AI Threat Analysis**

- Summarizes content in a structured JSON format.
- Sends it to Gemini using a carefully engineered cybersecurity prompt.
- Receives structured AI output:
  - **status:** SAFE / SUSPICIOUS / DANGEROUS
  - **explanation:** Summary of threat reason
  - **primary_threat_type:** Type (phishing, malware, etc.)
  - **confidence_score:** AI certainty level

5. **Popup Interface**
   - Clean UI styled in Google's aesthetic.
   - Displays AI results with reasoning.
   - Options: [Go Back], [Proceed Anyway].
6. **User-Managed API Key**
   - Stored in `chrome.storage.local`.
   - Used only by the extension via user's own Gemini API key.

---

4 Technical Architecture

Table 1: Trinetra Components

Component	Description
`content.js`	Extracts page DOM, forms, scripts, and metadata
`background.js`	Performs AI request and returns site verdict
`popup.html/js`	User UI to show scan results and store API key
`manifest.json`	Manifest V3 configuration for Chrome Extension

- Built on Manifest V3 for enhanced performance and security.
- Data payloads are truncated and optimized (`MAX_DATA_PAYLOAD_CHARS`).
- AI responses are cached briefly for performance.
- Known safe sites are whitelisted from analysis.
- Session memory stores overrides ("Proceed Anyway") temporarily.

---

5 Setup & Usage

1. Get your Gemini API key from Google AI Studio.
2. Download the latest ZIP `Trinetra.zip` from github.com/pradumon14/trinetra
3. Unzip the folder.
4. Go to `chrome://extensions` in your Chrome browser.
5. Enable **Developer Mode** (usually a toggle in the top right).
6. Click "**Load Unpacked**" and select the unzipped Trinetra folder.
7. Click the Trinetra extension icon in your browser toolbar.
8. Enter your API key in the extension popup.
9. Start Browse! Trinetra will now analyze pages.

---

6 Limitations & Disclaimers

- AI models can make false positives/negatives — always use personal judgment.
- Gemini API usage is at the user's own cost — monitor usage in Google Cloud Console.
- Minor delay possible on page load during active scanning.
- Data sent to Gemini includes page content — but only via the user's key; no third-party server is involved in this data transfer for AI analysis.
- Trinetra is an experimental prototype and not a commercial product.

---

7 Vision for Native Browser Integration

If integrated directly into browsers like Chrome:

- Gemini (or Gemini Nano) could analyze pages locally, enhancing privacy and speed.
- Threat detection would be native, fast, and private.
- Could evolve into a Smart Safe Browse 2.0 system.
- Would revolutionize security-first Browse at scale.

---

8 License

---

9 Contact

Pradumon Sahani
Email: pradumon14@gmail.com
Portfolio: https://pradumon.vercel.app
GitHub: https://github.com/pradumon14

— End of Whitepaper —