I. Project Information

Title

Mine-Safe India: Autonomous AI-Driven Data Sanitization & Forensic Verification Utility

II. Proposal Details

Problem Statement

The Indian e-waste sector operates on a flawed **"Custody-First, Wipe-Later"** model. Consumers hand over functional devices containing live PII (banking tokens, IDs) to collection agents, but data destruction only occurs 7–10 days later at centralized warehouses. This creates a **"Logistics-Gap" Vulnerability** a 168+ hour window where data is exposed to theft or cloning during transit. This security flaw causes consumer "Data Anxiety" and creates massive liability for recyclers, as no verifiable proof of deletion exists at the point of handover.

Objective

To deploy a portable "Edge-Computing" utility that enforces a **"Wipe-First, Custody-Later"** protocol:

- **Sanitize:** Execute DoD 5220.22-M cryptographic wipes immediately at the collection point.
- **Verify:** Use a Random Forest ML model to validate 0% data recoverability (entropy analysis).
- **Certify:** Generate an instant, tamper-proof "Digital Tombstone" (QR Code) for the customer.

Proposed Solution

Mine-Safe India is a forensic utility that runs on the "Hunter-Killer" architecture:

- **The Killer (Sanitization Engine):** Bypasses the OS cache/file system to inject high-entropy random data directly into physical storage sectors (NVMe/SSD). This renders magnetic recovery impossible.

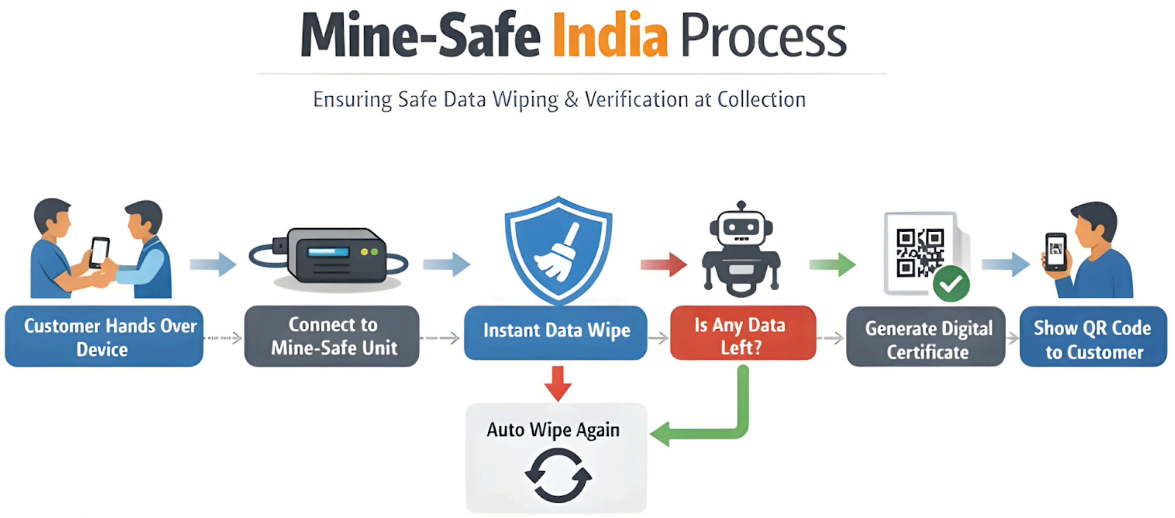
- **The Hunter (AI Auditor):** Unlike standard tools that assume a wipe worked, our AI acts as an adversarial auditor. It scans random drive sectors post-wipe, calculating **Shannon Entropy**. If the ML model detects structured data patterns (indicating failed erasure), it automatically re-triggers the sanitization loop.
- **The Chain of Custody:** Upon passing the audit, the system cryptographically signs the event log and issues a public verifiable QR code, closing the liability gap before the device moves.

Methodology

The development lifecycle follows the **NIST 800-88** guidelines for media sanitization:

Phase	Description
Phase 1	**HAL Development** - Building the Hardware Abstraction Layer (Python) to gain root-level access to storage controllers for direct block writing.
Phase 2	**AI Training** - Training the RandomForestClassifier on a synthetic dataset of 10k file headers (PDF/JPG) vs. high-entropy noise to ensure accurate auditing.
Phase 3	**Integration** - Merging the backend agent with a Streamlit-based "Compliance Dashboard" for non-technical operators.
Phase 4	**Validation** - Performing adversarial recovery attacks using forensic tools (e.g., Autopsy) to validate the "Certificate of Destruction."

Architecture


Mine-Safe India Process
Ensuring Safe Data Wiping & Verification at Collection
Customer Hands Over Device
Connect to Mine-Safe Unit
Instant Data Wipe
Is Any Data Left?
Generate Digital Certificate
Show QR Code to Customer
Auto Wipe Again

The system uses a **Client-Kernel Architecture**. The **Frontend (Streamlit)** is a localized web server providing a simplified UI for the field agent. The **Backend (Python Kernel)** runs as a daemon process with elevated privileges. It manages the `os` and `shutil` libraries for I/O operations. The **AI Inference Engine** runs locally (Edge AI), ensuring no sensitive data leaves the device during the verification process. *See below for the Conceptual Data Flow Diagram.*

Tech Stack

- **Programming Languages:** Python 3.10+
- **AI/ML Frameworks:** Scikit-learn (Random Forest), NumPy (Entropy Math), Pandas.
- **Databases:** SQLite (Local Audit Logs), JSON (Session Telemetry).
- **Cloud Platforms/Services:** N/A (Designed for Air-Gapped / Edge Execution for Security).
- **Other Key Tools:** Core Logic: `os`, `shutil` (Low-level I/O); Interface: Streamlit (Local Web Server); Security: PyCryptodome (Hashing), QRCode (Certificates); **Standards:** NIST 800-88, CERT-In Guidelines, Digital Personal Data Protection (DPDP) Act 2023..

III. Submission Details

Field	Detail
Submitter Name	Sarvesh Patil
Team Members	sai Katkade , Kaivalya Patil , Avani