Geo-Boundary Evidence Pack Generator (PRD)

---

1. About

The Geo-Boundary Evidence Pack Generator is an AI-powered compliance automation platform designed for high-stakes enterprise networking. It provides verifiable, auditor-ready evidence that sensitive data traffic remains within defined geographic boundaries (sovereign borders). In a world of increasing regulatory complexity, including GDPR, CCPA, and Saudi Data Law, this product transforms raw network telemetry into strategic compliance intelligence, ensuring “Zero-Trust Geo-Compliance.”

---

2. Market Insights

The market for sovereign networking is expanding rapidly as nations demand localized data residency. Current solutions provide “best effort” routing but lack the **verifiable evidence** required by GRC (Governance, Risk, and Compliance) teams to satisfy legal audits.

Competitor Analysis

- **Traditional Monitoring (Cisco ThousandEyes, SolarWinds):** These tools provide excellent path visualization but are designed for engineers. They lack automated, natural-language reporting for non-technical legal auditors.
- **Cloud-Native Solutions (AWS/Azure Sovereignty Zones):** These services lock users into a specific cloud ecosystem. Wayne’s World’s primary advantage is its neutrality and ability to verify paths across multi-cloud and hybrid environments.

Market Analysis

- **High Demand:** Concentrated in the Finance, Government, and Healthcare sectors.
- **Growth Trend:** The “Sovereign Cloud” market is projected to grow at a 25% CAGR as mid-market firms follow the lead of enterprise giants in adopting localized infrastructure.

Technology Analysis (Leveraging AI)

- **Path-to-Text Synthesis:** Utilizing LLMs (Llama) to translate complex JSON traceroute data and BGP path attributes into plain-English executive narratives.
- **Anomaly Forecasting:** Employs time-series models to predict when a path *might* flap outside a boundary due to upstream provider changes, allowing for proactive traffic management.

Customer Segments

- **Tier 1:** Multinational Banks with strict data residency requirements.
- **Tier 2:** Government and Public Sector agencies leading Sovereign Cloud initiatives.
- **Tier 3:** SaaS Providers serving highly regulated industries.

User Personas

- **Compliance Officer "Claire":** Needs to provide a monthly report to regulators proving data never left Germany. She is legally responsible for residency but does not understand BGP or technical logs.
- **Network Architect "Nigel":** Needs to verify that the "Sovereign Mesh" is behaving as configured and requires automated alerts on near-misses.

---

3. The Problem

Problem Statement

Compliance officers in regulated industries currently rely on manual, fragmented, and technical network logs to prove data residency to potential customers. This lack of verifiable proof leads to a delay in the enterprise sales cycle and a **decrease in gross monthly sales** due to stalled international contracts.

Pain Points

- **Manual Extraction:** Network engineers spend many hours per month manually parsing logs for auditors.
- **The "Technical Wall":** Auditors struggle to trust raw CLI output; they require synthesized, explainable data.
- **Reactive Compliance:** Violations are often discovered months later during audits rather than in real-time.

Hypotheses and Mission Statement

- **Hypothesis:** By automating the correlation of network telemetry with geo-spatial intelligence using AI, we will reduce audit preparation time and increase executive confidence in sovereign compliance.
- **Mission:** To empower global enterprises with **verifiable trust** by transforming complex network telemetry into automated, plain-English compliance evidence.

---

4. The Solution

Leveraging AI: Strategic Integration

1. **Synthesis:** AI condenses millions of technical path updates into a 3-page executive summary.

2. **Ambiguity Management:** Interpreting “fuzzy” geo-data where IP geolocation may be inconsistent across providers.
3. **Predictive Guardrails:** Identifying patterns that suggest a routing change is imminent that would violate a boundary.

Feature Prioritization (RICE Matrix)

Formula: (Reach x Impact x Confidence) / Effort = RICE Score (1-10 Scale)

Feature / Initiative	Reach	Impact	Confidence	Effort	**Score**
1. Cryptographic Evidence Signing	10	9	9	3	**270.0**
2. Phase 1: Cloud-Native Beta (VPC Logs)	10	7	9	4	**157.5**
3. “Plain-English” AI Path Synthesis	10	9	7	5	**126.0**
4. Phase 2: “Sovereign-Local” Appliance	8	10	8	7	**91.4**
5. Real-Time Geo-Violation Alerts	6	8	8	6	**64.0**
6. Phase 3: Fabric-Native Verification	10	10	5	10	**50.0**

AI MVP & Roadmap Strategy

The product follows a 3-phase evolution to balance speed-to-market with absolute data sovereignty:

- **Phase 1 (Cloud Beta):** A cloud-native SaaS that analyzes VPC Flow Logs (metadata). This proves the AI “Narrative Engine” in a low-friction environment.
- **Phase 2 (Sovereign-Local MVP):** A customer-managed virtual appliance designed for absolute data isolation. Deployment options include:
    - **Virtual Edge:** Running as a Virtual Network Function (VNF).
    - **Customer Bare Metal:** Direct installation on physical hardware running in a datacenter or in a colocation cage
    - **Private Cloud:** A containerized deployment (Docker/K8s) within the customer’s perimeter.
- **Phase 3 (Network-Native):** Full integration into **Wayne’s World Fabric** nodes for real-time, hardware-level verification.

Technical Architecture

- **Control Plane:** A localized GRC Dashboard for defining boundaries and generating reports.
- **Data Ingestion:**
  - *Phase 1:* Pulls VPC Flow Logs from Google/Azure via read-only APIs.
  - *Phase 2:* Ingests NetFlow/IPFIX telemetry directly from local or virtual routers.
- **Processing:** Localized Small Language Models (SLMs) like Llama-3-8B performing RAG (Retrieval-Augmented Generation) against local telemetry.
- **Security:** Air-gapped capable; no telemetry data ever leaves the customer's managed environment.

---

5. Requirements

Functional Requirements

- **FR-1 (Ingestion):** System MUST pull VPC Flow Logs (Cloud) and NetFlow/IPFIX (Local) while ignoring packet payloads to maintain privacy.
- **FR-2 (Synthesis):** The LLM engine MUST correlate telemetry with Geo-IP data to translate raw network hops into human-readable narratives.
- **FR-3 (Reporting):** Users MUST be able to generate a cryptographically signed PDF report for any time period with one click.
- **FR-4 (Verification):** Every report MUST include a unique QR code or link for auditor authentication.
- **FR-5 (Sovereignty):** In Phase 2, all AI inference MUST occur locally within the customer's perimeter with no internet requirement.

Non-Functional Requirements

- **NFR-1 (Security):** System MUST ensure 100% data isolation in Phase 2; no telemetry data can "phone home."
- **NFR-2 (Performance):** Monthly reports (up to 10M flows) MUST be generated in under 60 seconds.
- **NFR-3 (Explainability):** The AI MUST provide "Source Grounding," allowing auditors to verify the raw logs behind any narrative conclusion.

---

6. Measuring Success

- **North Star Metric:** Total "Sovereign Volume" Certified—the total TB of traffic verified as compliant.
- **Business Metric:** Reduction in the Enterprise Sales Cycle (Target: -20%).
- **Quality Metric:** Faithfulness Score—ensuring LLM summaries perfectly match the underlying network logs.

---

7. Roll-out Strategy

1. **Internal Alpha:** Testing on Wayne's World's own internal network.
2. **Marketplace Beta (Phase 1):** Launching for Google/Azure Cloud customers to validate the AI Narrative Engine.
3. **Sovereign Launch (Phase 2):** On-prem and Virtual Edge deployment for Tier 1 Financial and Government clients.
4. **Global GA (Phase 3):** Fully integrated Fabric-Native compliance for all Wayne's World customers.