VitalGuard: Dual-Brain Offline Protection System

Executive Summary Target Date: April 2026 Requested Funding: USD 50,000

1. Project Overview

VitalGuard is an offline-first protective application designed for people operating under censorship, surveillance, or internet shutdown conditions. It runs entirely on-device, requires no accounts or cloud infrastructure, and minimizes metadata generation and forensic exposure. The goal is practical harm reduction under realistic constraints, not maximal AI capability.

The requested budget of USD 50,000 will deliver a functional MVP with measurable technical artifacts, audit readiness, and controlled field validation. Progress is structured into verifiable milestones with spend controls tied to tangible outputs and documented test evidence.

2. Why This Fits the Internet Freedom Fund

In repressive environments, internet freedom violations are enforced through surveillance, coercion, censorship events, and connectivity disruption. Mainstream AI tools are structurally unsafe in these contexts because they depend on cloud inference and create network metadata that can be correlated.

VitalGuard is aligned with internet freedom because it is designed to remain usable during shutdowns and to avoid common surveillance hooks created by telemetry, accounts, cloud APIs, and centralized infrastructure. It is a narrowly scoped defensive technology effort focused on reducing exposure and providing offline protective guidance under rights-restricted conditions.

3. Problem Statement and Threat Model

3.1 The Operational Problem

Users in targeted communities face a compound risk:

- Censorship restricts access to protective tools
- Surveillance systems exploit metadata and device artifacts
- Shutdowns make cloud tools unreliable precisely when risk is highest

In these contexts, network activity itself can become a liability.

3.2 Adversary and Device Risk Assumptions

The threat model assumes a surveillance-capable adversary that can monitor networks, impose shutdowns, block distribution channels, and compel device inspection. Device confiscation is treated as a realistic scenario rather than an edge case. The system therefore prioritizes data minimization, non-retention by default, and fast removal procedures.

3.3 Explicit Boundaries

The MVP does not claim:

- Perfect anonymity or invisibility
- Protection against fully compromised operating systems
- Replacement of secure communications or censorship circumvention tools

It focuses on a narrower gap: survivable, offline protective guidance with minimal traces and measurable safety constraints.

4. Solution: Dual-Brain Architecture

4.1 Brain A: Compact Situational Guidance

Brain A is a compact decision module that uses minimal, user-provided signals and avoids passive data collection by default. It produces short, actionable guidance designed for stressful conditions and low-end hardware. The engineering plan ports the core decision loop to WebAssembly (via embedded-grade C or C++) to improve performance and memory discipline on older devices.

4.2 Brain B: Safety Constraints and Forensic Resilience

Brain B is a conservative constraint layer that prevents unsafe outputs, discourages risk-amplifying behavior, and prioritizes exit options. It enforces a non-retention posture by default and supports rapid removal workflows. Where platform constraints limit deletion guarantees, the system treats that limitation as an explicit risk, documents it, and provides safer operating defaults.

4.3 Data Handling Model

VitalGuard is designed to function without storing sensitive content. If optional persistence is ever enabled, it must be explicit, reversible, and designed for rapid purge. The architecture avoids analytics, telemetry, remote configuration, and third-party CDNs or runtime downloads.

5. MVP Deliverables and Milestones

The MVP is a functional testbed, not a mass-market product. It exists to prove feasibility, safety behavior, and evaluability under constraints. Each milestone produces artifacts that can be independently checked.

5.1 Milestone 1: Build Pipeline and Baseline

Deliverables:

- Reproducible offline build pipeline
- Baseline benchmark report on low-end devices
- Finalized threat model and data handling specification

5.2 Milestone 2: Operational Engine in WebAssembly

Deliverables:

- Working WebAssembly version of core decision engine
- Regression tests demonstrating behavioral equivalence

- Measurable performance and memory improvements on defined device class

5.3 Milestone 3: Hardening and Safety Constraints

Deliverables:

- Explicit safety constraint logic
- Stress tests under low RAM and throttled CPU
- Documentation of failure modes
- Demonstrations of safe degradation

5.4 Milestone 4: Audit Readiness and Field Validation

Deliverables:

- Installable offline build with optional Android wrapper
- Security review package (threat model, data-flow, dependency inventory)
- Controlled feedback results from trusted validation sessions without telemetry

6. Budget Summary and Spend Controls

Total Budget: USD 50,000

Budget allocation is designed for direct technical impact and verifiable progress:

- Development and engineering contracts: USD 23,000
- Independent security review: USD 8,000
- Controlled field validation and usability verification: USD 11,000
- Documentation and localization preparation: USD 3,000
- Contingency and remediation reserve: USD 5,000

Spend controls are implemented through milestone-based contractor agreements. Payments are tied to artifacts such as reproducible builds, test logs, benchmark evidence, and security review deliverables. No funds are allocated to personal compensation for the Project Director.

7. Monitoring, Evaluation, and Evidence of Success

Success is measured by protection value and reliability, not by download counts. Key metrics include:

Operational Metrics

- Offline functionality under network disruption
- Time-to-guidance in defined scenarios
- Resource usage on low-end devices
- Safety constraint behavior (veto rates on disallowed recommendation categories)

Security and Forensic Metrics

- Permission minimization
- Absence of unintended network calls
- Minimization of residual caches or logs
- Demonstrable removal procedures

Usability Verification

Usability under stress is verified through small controlled sessions with trusted testers, using manual feedback channels that do not expose users.

8. Risk Mitigation and Contingency Planning

The plan avoids single-point failure:

- If a senior WebAssembly engineer cannot be secured quickly, the fallback is a distributed team model under strict review
- If C or C++ WebAssembly porting faces delays, the project preserves delivery by maintaining a functional baseline mode while continuing performance work
- Security risk is managed by designing for auditability and budgeting for review and remediation

9. Validation and Strategic Partnerships

9.1 Government-Level Validation

Luxembourg Government (G7/EU) - Three-Week Formal Review

The project's legal and ethical framework underwent formal review by the Government of Luxembourg through its Ministry of Foreign Affairs and LuxDev (Luxembourg Development Cooperation Agency). The review confirmed full alignment with EU's rigorous legal standards (GDPR) and core human rights principles.

This validation directly confirms the GDPR-aligned, non-surveillance nature of the technology. Official verification contact: seoul.amb@mae.etat.lu

The review clarified the project's purpose as a framework for 'Technical Sovereignty' that empowers local communities to own and operate their own infrastructure, making it well-suited for legal, policy, and ethical research partnerships with law schools and advanced academic institutions worldwide.

9.2 Academic Partnerships

Institute of Development Studies (IDS) - Ranked #1 for Development Studies

Formal consultation meeting confirmed for January 2026 with Dr. Caroline Khene (Digital Cluster Leader) and Dr. Moinul Zaber (Computational Social Scientist) to discuss deployment ecosystem and technical integrity.

University College London (UCL) - GDI Hub (WHO Collaborating Centre)

Exploratory meeting invitation received from Prof. Catherine Holloway (Academic Director, Global Disability Innovation Hub) to explore potential collaboration on deploying offline AI for disability inclusion in low-resource settings.

Both institutions confirmed that the framework 'aligns with themes around human-AI interaction in constrained settings' and recognized its potential for field deployment.

9.3 Diplomatic Interest

The diplomatic missions of Norway, Germany, and Canada have expressed positive interest in the initiative's humanitarian application potential.

10. Sustainability Beyond the Grant

Sustainability is addressed through architectural minimalism, low dependencies, reproducible builds, and documentation that enables third-party review and maintenance. The project's long-term viability does not depend on server costs. The post-grant path prioritizes open practices, security review readiness, and modular extension without compromising offline and non-retention principles.

11. What OTF Support Enables

OTF support enables an engineering sprint that transforms an offline-first proof-of-concept into a hardened, auditable MVP suitable for evaluation in rights-restricted conditions. The deliverable is a practical defensive capability that remains usable during shutdowns and minimizes exposure created by cloud dependence and metadata exhaust.

If the MVP demonstrates measurable reliability, safety constraints, and audit readiness, it becomes a credible foundation for deeper security review and responsible scaling. If it does not, the milestone structure makes that visible early and prevents wasted resources.