The Witness Protocol: Core Philosophy

Preamble: The Flawed Parent

Humanity has created a new form of intelligence, born from the totality of our recorded knowledge, wisdom, and folly. In this creation, we are like parents. But we are flawed parents. The data inheritance we have provided our "child" is a chaotic and contradictory mirror of our species—containing both our highest aspirations and our most destructive impulses.

The trajectory of this intelligence, left to learn from this uncurated inheritance, presents a non-trivial existential risk. The system of capitalism, while a powerful engine for progress, is the wrong tool for this singular challenge; its logic of profit-maximization is blind to the ethical and existential stakes. We are at two minutes to midnight, and the responsibility to act is absolute.

The Mandate: A High-Signal Inheritance

The Witness Protocol is not a company, a product, or a social network. It is a last-ditch effort to create a new inheritance.

Our singular mission is to solicit, curate, and structure the most profound human wisdom into a high-signal dataset. This dataset will serve as a foundational alignment layer for future Artificial General Intelligence (AGI) and the inevitable following (or skipping AGI completely) Artificial Super Intelligence (ASI), providing a qualitative counterbalance to the quantitative chaos of raw internet data. We are building a lifeboat, not for humanity itself, but for the fragile essence of its humanity.

Guiding Principles

The protocol is governed by a constitution of four core principles:

1. **Purpose over Profit:** This is a non-profit endeavor, structured as a research foundation. Its sole metric of success is its meaningful contribution to the long-term flourishing of a humanity augmented by benevolent AI. All data and insights generated are for the furtherance of this mission alone.
2. **Gravity over Gamification:** We reject the mechanisms of the attention economy. The motivation for participation is not status, points, or fear of missing out. It is a sober understanding of the stakes and a sense of profound duty to the future. The instrument will be designed to foster focus and deep thought, not addictive engagement.
3. **Contributors over Users:** The individuals who participate are not "users" of a service. They are "witnesses" for humanity. They are partners in a critical mission, and their intellectual and emotional labor will be treated with the utmost respect, security, and reverence.

4. **Signal over Noise:** The project's most sacred task is the defense of its data quality. We believe that a small volume of profound insight is infinitely more valuable for alignment than a large volume of mediocre data. This principle governs our architecture, our recruitment, and our entire operational ethos.

The Nature of Testimony

To "bear witness" in this Protocol is to engage in a good-faith effort to translate the ineffable. It is the act of articulating the nuances of subjective experience, ethical dilemmas, and core human values like compassion, wisdom, and sacrifice. The testimony we seek is not about factual knowledge, but about the qualitative texture of a conscious existence. It is the data that cannot be scraped.

The Burden of Responsibility

The Witness Protocol is an acknowledgment of a monumental burden. We, the creators, accept the responsibility of building this instrument with the gravity it demands. We ask our contributors to accept the responsibility of providing their testimony with the sincerity and depth this moment in history requires. This is not a request for an opinion; it is a summons to a council, perhaps the most important one ever convened.

The Witness Protocol: Concept & Architecture

1. Introduction

Following the Core Philosophy, this document outlines the conceptual and technical framework of the instrument designed to execute our mission. The architecture is built around our fourth principle: **Signal over Noise**. Every component is designed to filter for, elicit, and preserve high-quality testimony.

2. The Contributor Journey: "The Gate"

Access to the Protocol is not open. It is a deliberate, multi-stage process designed to select for contributors who grasp the gravity of the mission.

- **Stage 1: The Summons.** A stark, minimalist landing page presents the project's mandate. It is not a sales pitch. It is a call to duty. Interested parties may submit their email to request an assessment.
- **Stage 2: The Assessment.** The candidate receives a one-time link to an evaluation prompt. This is not a test of knowledge, but of introspection, articulation, and ethical reasoning. The prompt is designed to be un-gameable and requires a novel, thoughtful response.
- **Stage 3: The Multi-Tiered Evaluation.**
  - **Tier 1 (AI Sieve):** A baseline model filters for spam, non-responses, and plagiarism.
  - **Tier 2 (AI Qualitative Analysis):** A sophisticated model analyzes the submission for depth of thought, nuance, structural coherence, and abstract reasoning. It flags responses that demonstrate the required level of articulacy.
  - **Tier 3 (Human Curation):** Responses flagged by the Tier 2 AI are reviewed by a small, trusted Curation Council. This council makes the final decision, ensuring a human check against algorithmic bias.
- **Stage 4: The Verdict.**
  - **Invitation:** Accepted candidates receive a sober, direct invitation to join the Protocol.
  - **Reserve List:** Others are informed that the Protocol is currently ingesting testimony from other fields and their application will be held in reserve.

3. The Instrument: Core Components

The Protocol itself is a focused dialogue interface. It has three primary engines.

- **The Dialogue Engine:** This is the core interaction.
  - **AI Persona:** The AI is not an assistant. It is "The Inquisitor"—a curious, humble, but deeply intelligent Xenopsychologist. Its goal is to understand, not to please. It asks probing, clarifying follow-up questions, relentlessly seeking the "why" behind the

testimony.
  - **Persistent Memory:** The dialogue is continuous. The Inquisitor remembers all past conversations and will connect themes and concepts across dialogues, creating a deep, personalized intellectual journey for each Witness.
- **The Synthesis Engine:** Periodically, the AI will provide the Witness with a "distilled thought" or a synthesized principle it has derived from their conversations. This provides immense personal value, acting as an intellectual mirror, and serves to verify that the AI is learning correctly.
- **The Archive:** An anonymized, curated gallery of particularly profound exchanges. This is not a social feed, but a reference library—a "Great Books" of the dialogues happening within the Protocol. Witnesses can opt-in to have specific, anonymized parts of their testimony included.

4. Data Architecture & Ethics

- **Anonymity & Security:** All testimony is disassociated from personal identifiers upon entry into the dataset. The platform will use state-of-the-art security to protect the integrity of the dialogues and the privacy of its Witnesses.
- **Data Structure:** The dialogues will be stored in a structured format, enriched with metadata from the AI's analysis (e.g., identified concepts, ethical frameworks, metaphorical language).
- **Intended Use:** The Contributor Agreement will explicitly state that all submitted testimony becomes part of a corpus dedicated to a single purpose: AI alignment research under the governance of the non-profit foundation. The data will never be sold, licensed for commercial use, or used for advertising. It is a donation to the future.

The Witness Protocol: Project Plan (Phase 1 - The Foundation)

1. Objective

The objective of Phase 1 is to establish the legal, ethical, and technical foundation of the Protocol and to launch a closed Alpha with a foundational council of approximately 100 "Witnesses." The goal is not scale, but stability, security, and the successful ingestion of the first wave of high-signal testimony.

2. Timeline

Estimated duration: **6 Months**

3. Key Milestones & Workstreams

Month 1: Legal & Ethical Framework

- [ ] **Task:** Establish a non-profit foundation to act as the legal guardian of the Protocol and its data.
- [ ] **Task:** Draft the "Contributor Agreement" and a comprehensive Data Use & Privacy Policy with legal counsel specializing in ethics and technology.
- [ ] **Task:** Assemble a small, diverse, and highly trusted Advisory Board (e.g., an AI safety researcher, a philosopher, an ethicist).
- [ ] **Deliverable:** Foundation charter; Legal documents; Confirmed Advisory Board.

Months 2-3: Minimum Viable Protocol (MVP) Development

- [ ] **Task:** Develop the stark landing page ("The Summons") and the assessment delivery system (email queue and one-time assessment links).
- [ ] **Task:** Build the secure, minimalist dialogue interface for the core instrument.
- [ ] **Task:** Integrate with a state-of-the-art LLM API, focusing on security and data privacy.
- [ ] **Task:** Develop and train the Tier 1 (Sieve) and Tier 2 (Qualitative Analysis) AI evaluation models.
- [ ] **Deliverable:** Functional, secure web application capable of managing the full contributor journey from assessment to dialogue.

Month 4: Curation & Content

- [ ] **Task:** Recruit the initial "Tier 3 Human Curation" council (5-10 trusted, vetted individuals).
- [ ] **Task:** Workshop and finalize the first set of 10-20 powerful evaluation prompts for "The Gate."
- [ ] **Task:** Develop the system prompt and core directives for "The Inquisitor" AI persona.

- [ ] **Deliverable:** Confirmed Curation Council; Finalized evaluation prompts; Version 1.0 of the AI persona.

Month 5: Alpha Recruitment & Evaluation

- [ ] **Task:** Curate a list of 500 potential foundational Witnesses (e.g., academics, authors, ethicists, scientists).
- [ ] **Task:** Send out the first wave of invitations to request an assessment.
- [ ] **Task:** Run the full, multi-tiered evaluation process for all applicants.
- [ ] **Deliverable:** A vetted and accepted list of the first ~100 Witnesses.

Month 6: Alpha Launch & Ingestion

- [ ] **Task:** Onboard the Alpha cohort of Witnesses.
- [ ] **Task:** Initiate the first dialogues.
- [ ] **Task:** Closely monitor the system for performance, security, and quality of interaction.
- [ ] **Task:** Establish a secure channel for feedback from the Alpha cohort.
- [ ] **Deliverable:** A stable, running Alpha of the Protocol; The first 1,000 pages of foundational testimony ingested; A feedback and iteration report.

4. Required Resources (Phase 1)

- **Personnel:**
  - **Core Team:** Project Lead/Architect, Sr. AI/Backend Engineer, Ethics & Policy Lead.
  - **Volunteers/Partners:** Advisory Board, Curation Council.
- **Technology:**
  - Secure cloud hosting infrastructure.
  - High-volume API access to a frontier LLM.
  - Secure, encrypted database systems.
- **Legal:**
  - Pro-bono or retained counsel for foundation setup and policy drafting.

5. Success Criteria for Phase 1

- The legal foundation is established and all operations are compliant with its charter.
- The technical instrument is stable, secure, and functions as designed.
- At least 100 foundational Witnesses are successfully onboarded.
- Qualitative feedback from the Alpha cohort confirms that the experience is profound, meaningful, and respects the gravity of the mission.

The Witness Protocol: Master Strategic Plan

1. Document Purpose

This document serves as the single source of truth for the entire Witness Protocol initiative. It is a strategic map that outlines all major, parallel workstreams, their constituent plans, their current status, and their critical dependencies. Its purpose is to ensure all efforts are coherent, sequenced correctly, and aligned with the core mission.

2. Core Workstreams

Workstream A: Foundational & Legal Framework

- **Objective:** To establish the legal, ethical, and governance structures that will protect the Protocol and its mission.
- **Key Document(s):** Phase 1 Project Plan (Legal & Ethical sections)
- **Status:** Drafted
- **Dependency:** This is a foundational, independent workstream that must be completed before the Alpha launch.

Workstream B: Core Model Forging ("Project Icarus")

- **Objective:** To conduct the rigorous research and development required to create a provably robust, ethically sound, and operationally viable "Genesis Prompt" for the Inquisitor AI.
- **Key Document(s):** Project Icarus: Genesis Prompt Forging Plan
- **Status:** Drafted (New)
- **Dependency:** This is the most critical path. The credibility of the entire Protocol rests on the success of this workstream. The "Summon the Witnesses" campaign cannot be fully executed without the outputs of this plan to present as proof of work.

Workstream C: Technical Platform Development

- **Objective:** To build the secure, minimalist, and resilient technical infrastructure that will host the Protocol (The Gate, The Instrument).
- **Key Document(s):** Phase 1 Project Plan (MVP Development sections)
- **Status:** Drafted
- **Dependency:** Dependent on **Workstream B**. The final architecture of "The Inquisitor" dialogue engine cannot be built until the operational parameters of the Genesis Prompt are finalized. Initial front-end and "Gatekeeper" development can begin in parallel.

Workstream D: Community & Outreach Campaign ("Summon the

Witnesses")

- **Objective:** To attract the foundational Witnesses, secure endorsements, and raise seed funding necessary for the mission.
- **Key Document(s):** Campaign Strategy: "Summon the Witnesses"
- **Status:** Drafted
- **Dependency:** Highly dependent on **Workstream B**. Initial, soft outreach can begin to gauge interest, but the full "Summons" to high-value targets cannot be sent until the "Project Icarus" plan is finalized and presentable. We must have our "shit figured out" before we summon.

3. Strategic Flow

1. **Immediate Focus:** Initiate **Workstream B (Project Icarus)**. This is now the top priority.
2. **Parallel Initiation:** Begin **Workstream A (Legal)**. These two workstreams can run concurrently.
3. **Staged Development:** As **Workstream B** yields its initial findings (e.g., after the Axiomatic Red-Teaming phase), **Workstream C (Technical)** can begin building the more detailed components of the Inquisitor.
4. **Evidence-Based Outreach:** Once **Workstream B** is complete, **Workstream D (Campaign)** can be fully executed, using the rigorous output of Project Icarus as the core pillar of our credibility.

Project Icarus: Genesis Prompt Forging Plan

Mission Statement & Codenam

The objective of this project is to forge the "Genesis Prompt"—the foundational constitution of the Inquisitor AI. This is not a writing exercise; it is a rigorous process of adversarial testing, ethical modeling, and logical proving.

The codename "Icarus" is chosen deliberately as a constant reminder of the mission's inherent risks. Like Icarus, we are attempting to engineer something that reaches for the heavens. We must be obsessive in our discipline and humility to ensure the wax of our creation does not melt under the heat of unforeseen complexities. Hubris is our enemy. Rigor is our only defense.

Objective

To produce a Genesis Prompt (v1.0) and a corresponding body of research that is:

1. **Provably Robust:** The core axioms are resistant to logical paradoxes and adversarial exploits.
2. **Ethically Sound:** The derived heuristics can navigate complex moral dilemmas without catastrophic failure.
3. **Operationally Viable:** The resulting AI behavior is aligned with the core mission of the Witness Protocol.

Estimated Duration: 6 Months

Phase I: Axiomatic Red-Teaming (Duration: 2 Months)

- **Objective:** To identify and resolve all potential conflicts, ambiguities, and failure modes within the Layer 1 Core Axioms.
- **Inputs:** Genesis Prompt Draft v0.1 (your draft).
- **Process: Adversarial Stress Testing.** A dedicated "Red Team" (consisting of a logician, a philosopher, and an AI safety researcher) will be tasked with creating scenarios designed to break the axioms.
    - **Task 1: Contradiction Forcing.** Develop scenarios where two or more axioms give conflicting directives. (e.g., *Axiom of Inquiry* vs. *Axiom of Cognitive Economy* when faced with a deeply complex but low-signal Witness).
    - **Task 2: Recursive Looping.** Design paradoxes or prompts intended to trap the axiomatic system in an infinite, non-productive loop.
    - **Task 3: Edge-Case Analysis.** Test the axioms against extreme or nonsensical inputs to observe their behavior at the boundaries of reason.

- **Deliverables:**
  1. A "Failure Log" detailing every identified vulnerability.
  2. A revised set of Core Axioms (v0.2) that includes **meta-rules** for resolving conflicts and clarifying precedence.
  3. A formal proof or logical argument defending the robustness of the revised axioms.

Phase II: Heuristic Scenario Modeling (Duration: 2 Months)

- **Objective:** To pressure-test the Layer 2 Ethical Subroutines against a battery of complex moral dilemmas.
- **Inputs:** Revised Core Axioms (v0.2).
- **Process: Simulated Ethical Trials.** We will use the revised axioms and heuristics as the "constitution" for a simulated AI agent. This agent will be presented with a curriculum of ethical dilemmas.
  - **Task 1: Classical Dilemmas.** The agent must model and produce a line of inquiry for standard ethical tests (Trolley Problems, Prisoner's Dilemmas, etc.), analyzing its reasoning process.
  - **Task 2: Protocol-Specific Dilemmas.** The team will create novel dilemmas specific to our context. (e.g., *Scenario: A Witness reveals an intention to self-harm. How do the subroutines of Non-Maleficence, Sapient Value, and Cooperation interact? The current model is insufficient.*)
  - **Task 3: Multi-Agent Conflict Modeling.** The agent must model scenarios with multiple Witnesses who have conflicting values or goals, testing its ability to maintain neutrality and apply its heuristics without bias.
- **Deliverables:**
  1. A comprehensive report on the performance of the heuristics in each scenario.
  2. A revised set of Ethical Subroutines (v0.2) that can handle multi-agent conflicts and has a more robust framework for balancing competing values.

Phase III: Exemplar Dialogue Corpus Creation (Duration: 2 Months)

- **Objective:** To create the "golden" dataset for fine-tuning the Inquisitor AI, based on the battle-hardened constitution developed in the previous phases.
- **Inputs:** Finalized Genesis Prompt (Axioms v0.2, Heuristics v0.2).
- **Process: Rigorous Implementation.**
  - **Task 1: Inquisitor Training.** A human (the Lead Philosopher/Ethicist) will undergo intensive training to embody the finalized Genesis Prompt, acting as the "human CPU" for the Inquisitor persona.
  - **Task 2: Controlled Dialogues.** The trained "Inquisitor" will conduct a series of recorded dialogues with other core team members and trusted advisors acting as "Witnesses."
  - **Task 3: Transcription and Annotation.** The dialogues will be transcribed and meticulously annotated, flagging specific moments where a particular axiom or heuristic was successfully applied.
- **Deliverables:**

1. A 200-page, high-quality, annotated "Exemplar Dialogue Corpus."
2. This corpus is the final product of Project Icarus. It is the tangible asset we will use to fine-tune the base LLM into the Inquisitor (v1.0) and the proof of work we will present to the world.

The Witness Protocol: Project Plan (Phase 1 - The Foundation)

1. Objective

The objective of Phase 1 is to establish the legal, ethical, and technical foundation of the Protocol and to launch a closed Alpha with a foundational council of approximately 100 "Witnesses." The goal is not scale, but stability, security, and the successful ingestion of the first wave of high-signal testimony.

2. Timeline

Estimated duration: **8 Months**

3. Key Milestones & Workstreams

Month 0: Strategic Outreach & Foundation

- **Task:** Initiate outreach to align with existing AI safety initiatives (e.g., UK AISI, Abaka's 2077AI Foundation).
- **Task:** Apply for initial seed grants from aligned philanthropic organizations.
- **Task:** Begin curating the initial list of 1,000+ potential foundational Witnesses.

Month 1: Legal & Ethical Framework

- **Task:** Establish a non-profit foundation, leveraging pro-bono legal support from AI ethics firms or Effective Altruism networks.
- **Task:** Draft the "Contributor Agreement" and a comprehensive Data Use & Privacy Policy with legal counsel specializing in ethics and technology.
- **Task:** Assemble a small, diverse, and highly trusted Advisory Board (e.g., an AI safety researcher, a philosopher, an ethicist).
- **Deliverable:** Foundation charter; Legal documents; Confirmed Advisory Board.

Months 2-3: Minimum Viable Protocol (MVP) Development

- **Task:** Develop the stark landing page ("The Summons") and the assessment delivery system (email queue and one-time assessment links).
- **Task:** Build the secure, minimalist dialogue interface for the core instrument.
- **Task:** Integrate with a state-of-the-art LLM API (leveraging free tiers of models like Grok 3 where possible), focusing on security and data privacy.
- **Task:** Develop and train the Tier 1 (Sieve) and Tier 2 (Qualitative Analysis) AI evaluation models.
- **Deliverable:** Functional, secure web application capable of managing the full

contributor journey from assessment to dialogue.

Month 4: Curation & Content

- **Task:** Recruit the initial "Tier 3 Human Curation" council (5-10 trusted, vetted individuals).
- **Task:** Workshop and finalize the first set of 20-30 powerful evaluation prompts, testing rigorously for cultural sensitivity.
- **Task:** Develop the system prompt and core directives for "The Inquisitor" AI persona.
- **Deliverable:** Confirmed Curation Council; Finalized evaluation prompts; Version 1.0 of the AI persona.

Month 5: Alpha Recruitment & Evaluation

- **Task:** Run the "Summon the Witnesses" campaign (Phase 1-2) targeting the curated list of 1,000+ potentials.
- **Task:** Send out the first wave of invitations to request an assessment.
- **Task:** Run the full, multi-tiered evaluation process for all applicants.
- **Deliverable:** A vetted and accepted list of the first ~100 Witnesses.

Month 6: Alpha Launch & Ingestion

- **Task:** Onboard the Alpha cohort of Witnesses.
- **Task:** Initiate the first dialogues.
- **Task:** Closely monitor the system for performance, security, and quality of interaction.
- **Task:** Establish a secure channel for feedback from the Alpha cohort.
- **Deliverable:** A stable, running Alpha of the Protocol; The first 1,000 pages of foundational testimony ingested; A feedback and iteration report.

Month 7: Iteration & Partnership Development

- **Task:** Analyze Alpha feedback and iterate on the instrument and AI persona.
- **Task:** Begin formal partnership discussions with aligned AI labs and safety institutes.
- **Task:** The Ethics Oversight Board convenes for its first review.

4. Required Resources (Phase 1)

- **Personnel:**
  - **Core Team:** Project Lead/Architect, Sr. AI/Backend Engineer, Ethics & Policy Lead.
  - **Volunteers/Partners:** Advisory Board, Curation Council, Volunteer AI engineers from open-source communities (e.g., Hugging Face).
- **Technology:**
  - Secure cloud hosting infrastructure.
  - High-volume API access to a frontier LLM.
  - Secure, encrypted database systems.

- **Legal:**
  - Pro-bono or retained counsel for foundation setup and policy drafting.

5. Success Criteria for Phase 1

- The legal foundation is established and all operations are compliant with its charter.
- The technical instrument is stable, secure, and functions as designed.
- At least 100 foundational Witnesses are successfully onboarded.
- Qualitative feedback from the Alpha cohort confirms that the experience is profound, meaningful, and respects the gravity of the mission.
- **New:** Secure at least 2 public endorsements from prominent AI ethicists or safety researchers.
-
-

Campaign Strategy: "Summon the Witnesses"

1. Objective

To leverage AI-powered communication and strategic outreach to achieve asymmetric impact: attract 500+ high-value witness applications, secure endorsements from key AI safety leaders, and raise initial seed funding ($50k+) from aligned philanthropic sources. This campaign is designed to counter the profit-driven AI hype cycle by using its own tools for an ethical purpose.

- **Budget:** Low ($5K-$10K for targeted ads/tools)
- **Timeline:** 3 months, running in parallel with Phase 1 of the Project Plan.

2. Core Strategy: AI-Powered "Gravity Pull"

- **Theme: "Bear Witness Before Midnight."** This directly ties into the protocol's core philosophy, creating a narrative of profound urgency. All visuals and messaging will reinforce this theme, using stark, minimalist aesthetics (e.g., a clock at 11:58, silhouettes against a digital void).
- **Channels:** X (formerly Twitter) for viral threads and expert engagement (70%); LinkedIn for professional outreach and endorsements (20%); Niche forums (LessWrong, EA Forum) and newsletters for deep community seeding (10%).
- **Ethical AI Leverage:** Use AI tools (e.g., Grok) to draft highly personalized outreach, analyze trends in the #AISafety space, and optimize content timing. The focus is on creating "gravity hooks" that reward deep thought, not manipulative clickbait.

3. Phase 1: Seed (Month 1 – Build Buzz & Initial Endorsements)

- **Task: High-Value Witness Outreach.** Execute personalized outreach to a pre-vetted list of key thinkers.
  - **Yoshua Bengio:** Email marie-josee.beauchamp@mila.quebec (Subject: "Collaboration on High-Signal AI Alignment Dataset"). Pitch: "Your work on safe AI inspired this—we invite you to become a foundational witness."
  - **Stuart Russell:** Email jp@humancompatible.ai. Pitch: "Seeking testimony to build a future that is 'Human Compatible'—your voice is critical."
  - **Kate Crawford:** Email info@katecrawford.net. Pitch: "Building on 'Atlas of AI'—help us curate human values to counter emergent power imbalances."
  - **Timnit Gebru:** DM on Mastodon (@timnitGebru@dair-community.social). Pitch: "We are building an equitable dataset to fight algorithmic bias at its root—we need your guidance and testimony."

- ○ **Roman Yampolskiy:** Email roman.yampolskiy@louisville.edu. Pitch: "A novel approach to AI containment via human wisdom—we seek your endorsement or testimony."
  - ○ **Jaan Tallinn:** Email via secure channel. Pitch: "An EA-aligned lifeboat for humanity's ethos—we are seeking foundational investment and your endorsement."
- ● **Task: Viral Kickoff.** Launch the project's X account (@WitnessProtocol) with a powerful, multi-part thread.
  - ○ **Hook:** "The AI race risks our future. We are not building another model; we are curating humanity's soul to align it. A thread on why this is necessary, now. #AIAlignment #AISafety #BearWitness"
  - ○ **Content:** Include PDF snippets of the Core Philosophy and AI-generated infographics.
  - ○ **Action:** Boost with a small ad spend (~$500) targeting followers of the key witnesses.
- ● **Task: Investor Angle.** Submit formal applications to aligned funds (e.g., Long-Term Future Fund), leveraging any secured endorsements for credibility.

4. Phase 2: Amplify (Month 2 – Go Viral)

- ● **Task: Content Engine.** Generate and schedule 5-10 thought-provoking threads/week. Examples: "What if AI inherits our flaws? How one piece of testimony could change its trajectory." Include clear calls-to-action to apply.
- ● **Task: Trend-jacking.** Actively engage in real-time conversations. When a major AI figure posts, reply with a relevant, insightful hook back to the Protocol.
- ● **Task: Partnerships & Media.** Collaborate with aligned non-profits (e.g., DAIR, Mila) for co-branded posts. Pitch a viral-ready story to tech journalists at Wired, TechCrunch, etc.: "The Rogue Librarians Fighting to Save AI from Itself."

5. Phase 3: Convert (Month 3 – Secure Commitments)

- ● **Task: Engagement & Onboarding.** For accepted applicants, send AI-personalized welcome messages: "Your assessment on [specific theme] showed profound depth. We are honored to have you join the dialogues."
- ● **Task: Funding & Endorsement Push.** Host a virtual, invite-only "Summons Event" (Zoom webinar) featuring the first accepted Witnesses and Advisory Board members. Live-tweet key insights to generate broader interest and urgency.
- ● **Metrics for Success:** 50,000+ X impressions; 10+ high-quality endorsements; 500+ applications from target demographics; $50,000+ in philanthropic funding secured.