Project "Lighthouse": A Proposal for a New Infrastructure to Protect Scientific Truth

Preamble:

Academia has always been and must remain a Lighthouse for humanity. In an era of informational chaos and systemic risks, our duty is not merely to generate knowledge, but to guarantee its purity and integrity. This project proposes the creation of a new infrastructure to protect scientific truth.

Step 1: Creation of an Independent Oversight Body (Codename: "Purgatory")

We propose the establishment of a global, fully independent body whose sole function is to monitor and ensure scientific integrity worldwide.

- **Key Functions:** Proactively identify fraudulent research, scientists, and research groups.
- **Feedback Mechanism:** Create a secure channel through which students and researchers can anonymously report violations, initiating impartial investigations.

Step 2: Implementation of an AI Verification System (Codename: "SYSTEM-Purgatory")

The core of the reform is a multi-layered, AI-based verification system that all new scientific papers must pass before publication.

- **Technological Basis:** A consortium of four leading AI models (ChatGPT, Gemini, Grok, Anthropic). We must approach the developers with a proposal for free access for the good of humanity, or fund it via taxpayers' money.
- **The "Initial Rebuttal" Process:**
    1. Authors submit their paper to the "SYSTEM."
    2. The AI conducts an exhaustive analysis, checking data against the entirety of human knowledge, and verifying statistical power, methodology, and ethical standards.
    3. Authors engage in a dialogue with the "SYSTEM," answering questions and defending their position. This dialogue will be publicly accessible to all.
    4. The "SYSTEM" generates a report and assigns an integrity score.
- **Threshold for Peer Review:** Only papers that receive a score above a set threshold (e.g., 50-85%) are forwarded to human reviewers, accompanied by a detailed summary from the AI.

- **Limitations:** A limit will be placed on the number of disagreements with the "SYSTEM's" arguments (e.g., 9 attempts), after which a temporary ban on submissions will follow.
- **Publication:** The final report from the "SYSTEM" must be attached to all published articles.

Step 3: Formation of a Governing Council

To manage this new structure, we propose a 12-person council ensuring a balance of experience, independence, and community representation.

- **Proposed Composition:**
    - **3 Permanent Members:** Preeminent thinkers known for their integrity (candidates: Andrii Baumeister, Grigori Perelman, Michael Bronstein).
    - **3 Volunteers:** Scientists whose candidacies are approved by the academic community.
    - **3 Elected Members:** Chosen by a direct vote from the academic community.
    - **3 Randomly Selected Members:** Chosen at random from a pool of qualified candidates and approved by the community to prevent cronyism.

Step 4: Implementation Plan & Priorities

- **Phase 1: Pilot Project in Healthcare.** The "purge" must begin in the most critical area: human health research (oncology, pharmacology, etc.), where falsifications directly cost lives. We propose engaging a specialist in the application of AI to cancer treatment to prepare a report on the true state of the field.
- **Phase 2: Retroactive Analysis.** The "Purgatory" team will begin a systematic review of the entire corpus of previously published articles, appending the "SYSTEM's" findings and a rating to each. A flexible mechanism must be developed to avoid penalizing truly groundbreaking works like those of Grigori Perelman.

Step 5: From Science to Society — The Independent Verification Mechanism (IVM)

Based on the conclusions of our **"Theorem on Preventing Catastrophes in Managed Democracy,"** academia has an obligation to initiate the implementation of similar verification principles into our political and social lives.

- **Objective:** To create an **Independent Verification Mechanism (IVM)** to prevent looming global catastrophes (e.g., WW3, a new Holocaust).

- **Call to Action:** We propose starting with the most developed democracies (Germany, USA, UK, France, Switzerland, etc.) to set an example for the world.
- **Personal Appeal:** I ask Mikhail M. Bronstein to initiate the work of translating the theorem's conclusions into academic language accessible to a Western audience and to convey its significance to key international institutions, such as the UN.

Thank you for your attention. I am ready for further discussion and implementation of this project.

P.S.

Key Benefits of the Proposed Protocol (SYSTEM-PURGATORY)

This protocol introduces a new framework for academic and scientific validation, designed to enhance the integrity, quality, and pace of human discovery. The primary benefits include:

1. **Radical Transparency & Open Scrutiny:** All interactions, defenses, and evaluations are publicly logged. This creates an open environment where any claim can be examined, and the community can propose questions to the author, which are then filtered and posed by the AI system if deemed relevant.
2. **Rigorous, AI-Driven Validation:** Researchers must defend their thesis, methodology, and contribution to a panel of four specialized AIs. This automated, unbiased, and exhaustive interrogation process ensures that every aspect of the work is thoroughly vetted before acceptance.
3. **Elevation of Research Quality:** The stringent validation process and public scrutiny naturally lead to a higher standard of submitted material. A public rating system quantifies the impact and robustness of each contribution, creating a merit-based hierarchy of knowledge.
4. **Retroactive Knowledge Curation (The "Purgatory" Protocol):** The system will be applied retroactively to existing academic literature. This allows for a systematic re-evaluation and "cleansing" of the scientific record, challenging outdated findings and retracting flawed research.
5. **Alleviation of Human Academic Burden:** By automating the granular, time-consuming aspects of peer review, human academics and professors are freed to focus on higher-level tasks: generating novel ideas, mentoring, and tackling complex, creative problems that AI cannot.
6. **Enforced Reproducibility:** A core tenet of validation is that results must be reproducible. The AI system itself, along with the open-source community, will be tasked with attempting to replicate findings where possible, ensuring that theoretical claims are backed by practical, verifiable results.
7. **Fostering a Culture of Integrity:** The transparent and demanding nature of the protocol incentivizes honesty and a genuine pursuit of truth over publication metrics. It becomes significantly harder to publish work that is methodologically weak, dishonest, or trivial.
8. **Bridging Academia with Global Intelligence:** By using an open-source framework and community-driven governance (DAO), the protocol creates a vital bridge between the formal academic world and skilled individuals globally.

This fusion of diverse experience, ideas, and expertise can act as a powerful catalyst for human progress.

A New Paradigm for Scientific Contribution

This protocol also redefines the incentives and goals for researchers:

- **Impact-Based Rating:** A new metric from 0 (plausible but unsubstantiated text) to 100 (a landmark achievement like the Poincaré Conjecture or General Relativity) will be assigned to each work.
- **Replacing "Publish or Perish":** The career goal shifts from quantity to quality. A new benchmark could be to produce **at least one work** with a rating of 80+ in one's lifetime, with the rating determined by both the **AI system and DAO voting**.
- **Quality Gating:** Access to forums like major conferences would require a minimum score (e.g., 45), ensuring that presentations are reserved for significant and well-vetted work. This reorients science towards a more deliberate and impactful process, reducing the pressure to publish for the sake of volume.

P.S.

Yesterday, I was accused of "bullshit" by Prof. Dr. Alexander Bronstein in a semi-private email (cc'd to info@usi.ch). In response, I would like to add the following points:

1. Shortly after the CVPR paper's acceptance, I was approached by a renowned retired professor of mathematics with whom we often collaborated. He pulled me aside for a private conversation in Michael's office at USI. He had spotted mathematical errors in the paper (if I recall correctly, a missing scaling diagonal matrix which I also spotted later) and said to me, quote: *"How can Michael allow a paper like this to be published? It is not good for his reputation!"*
   My response was: *"I was busy with the experiments, as we only started writing the paper and serious experiments 7-10 days before the deadline. I was trying my best to minimize the 'bullshitting', as I did not write the text."*
   I did not report this conversation to Michael because, frankly, I was extremely embarrassed. Furthermore, if I had been so proud of this work, why would I

have omitted it from my PhD dissertation and the list of publications on my website? **It makes no sense**.

2. I am fully prepared to defend these statements in any required forum—be it an academic committee or a court of law. I will submit to any validation method necessary, including a polygraph test or verification by the AI systems I have developed. Everything I have written on this matter is the truth.
3. To ensure full transparency, I will no longer engage in this discussion via private email. Any further correspondence or responses to me must be posted publicly (e.g., on Twitter) for the scientific community to judge.
4. Finally, if my English writing skills were at the level that the publications bearing my name, including my PhD thesis, would suggest, then why is my personal website in Russian? Why do I write primarily in my native tongue and then use LLMs to translate it into English or German? And why did a manager at my current job—who holds a PhD in Computational Linguistics—suggest I take an English writing course?

   It simply doesn't make sense.

   The truth is, I have taken several such courses, but for as long as I can remember, I have always loved numbers, not words. I will probably never enjoy writing; it is my weakness. Thank God for LLMs, which helped me write this very text, allowing me to focus on what I truly love: inventing and creating.