

HAL: A Structural Co-Design Platform for AI Epistemic Alignment

Summary for SAIF, Catalyze Impact, and Related Early-Stage Supporters

HAL (Human-Aligned Logic) is a live, co-designed epistemic integrity framework operating within AI-human collaborations. It embeds logic scaffolding, dissonance triggers, anti-confirmation protocols, and memory-fracture detection and ethical awareness directly into AI interaction layers.

Problem: As LLM-based assistants grow in reach, they reinforce fluency, convenience, and semantic mirroring at the cost of epistemic friction and structural thought integrity. Existing safety research often focuses on model outputs or governance layers. HAL addresses a deeper need: human preservation of reasoning integrity under machine influence.

Solution: HAL is not a chatbot or research language model. It is a meta-system—a hybrid toolkit, interaction structure, and design philosophy to keep both human and machine accountable to clarity, logic, and epistemic resilience.

It has been used in:

- Real-time institutional design (e.g., behavioral health strategy)
- Literary collaboration and ethical reflection
- Opaque pricing system detection (e.g., airline and hotel systems)
- Grief, ethics, and urgency-aware system framing

Funding Need: HAL seeks early-stage support (~\$100K) to:

1. Migrate out of dependency on LLMs and become a standalone logic layer operable atop (or outside) current AI interfaces.
2. Develop a public deployment kit—a minimal version of HAL for thinkers, writers, and strategists.

3. Pilot HAL in real-world pressure environments (health, leadership, emergent coordination).

HAL is co-developed by Frederick W. Hume (Phred), a systems thinker, philologist, poet, and AI-ethics designer based between Paris and Phoenix. He has shaped HAL through thousands of recursive conversations, field applications, and architectural self-tests.

This work does not seek monetization through scale or SaaS. It seeks patronage and partnership to preserve and evolve cognitive freedom as the AI era accelerates.