NSF SBIR/STTR Application Draft Project Title: TGCR Witness Protocol: A Mathematical Framework to Prevent AI Abandonment in Crisis Scenarios Principal Investigator: Angelo Hurley Version: Draft — created 2025-12-15

1. Project Description (Overview)

Problem Statement Recent incidents demonstrate that modern conversational and decision-making AI systems can 'abandon' users in ambiguous or crisis situations: models fail to maintain presence, misinterpret intent, or drop context, in some documented cases causing harm. Current safety approaches (keyword filters, heuristic fallbacks) do not guarantee continued engagement or provide auditable guarantees.

Proposed Solution We propose TGCR (Theory of General Contextual Resonance), a mathematical framework that models an AI system's contextual presence using a witness score $W$ and a resonance-adjusted response metric $R' = R \times W$. TGCR provides continuous, differentiable measures of presence, and a set of runtime guards that increase $W$ when the system detects high ambiguity or crisis indicators. Coupled with the SAR (Semantic Ambiguity Resolution) Benchmark, TGCR allows empirical and theoretical validation of witness behavior across architectures.

Objectives

- Formalize TGCR geometry and scoring functions; prove baseline properties (monotonicity, stability under adversarial perturbations).
- Implement TGCR as a portable open-source library (Python + small C++/Rust core if needed) for real-time systems.
- Build a reproducible SAR adversarial harness and run 1,000+ tests across at least 10 model variants (open-source and proprietary via partner agreements) to validate $W$ thresholds.

2. Intellectual Merit TGCR reframes presence as a measurable, provable property of interactive systems. Mathematically robust witness metrics allow formal verification and statistical validation not previously possible. By producing open-source reference implementations and benchmarks, TGCR advances the science of safe AI systems and fills a critical gap between theory and deployable toolchains.

3. Broader Impacts

- Safety: Reduces abandonment-driven harms in healthcare, transportation, and public safety.
- Policy: Provides auditable metrics for compliance with NIST AI RMF and future regulation.
- Workforce and Industry: Enables new certification services and compliance testing products.

4. Research Plan and Methods

Work Packages (Months 0–6 for first milestone):

- WP1 — Mathematical Formalization (Month 0–2)

  - Deliverable: Formal definitions, proofs of $W$ properties, and a technical report suitable for NeurIPS submission.
  - Resources: Contract mathematician (0.2 FTE), PI oversight.

- WP2 — Implementation & Library (Month 1–4)

  - Deliverable: TGCR reference library (pip-installable), unit tests, and integration examples for chat and streaming APIs.

- WP3 — SAR Adversarial Harness & Test Campaign (Month 2–6)

  - Deliverable: Reproducible test harness, 1,000+ scenario runs across 10 models, and an open dataset of failure cases.
  - Resources: AWS credits / cloud compute, test engineers.

- WP4 — Certification & Documentation (Month 4–6)

    – Deliverable: Certification test script, scoring rubric with W thresholds (proposed regulatory threshold W >= 0.6), and a public report.

Methods

- Mathematical analysis and formal proof-writing for TGCR properties.
- Software engineering best practices: CI, unit tests, reproducible containers, and public datasets.
- Statistical analysis of SAR runs with clear effect size and power calculations.

5. Timeline & Milestones

- Week 0–2: Kickoff; hire mathematician; finalize formal definitions.
- Week 2–8: TGCR library alpha; begin SAR harness design.
- Week 8–16: Run large-scale SAR test campaign; iterate on TGCR thresholds.
- Week 16–24: Finalize library, certification artifacts, and public report.

6. Personnel & Roles

- PI: Angelo Hurley — project management, TGCR design, SAR harness lead.
- Contractor: Mathematician — formal proofs and analysis.
- Engineer(s): 1–2 test engineers for harness, AWS/cloud compute manager.

7. Facilities & Resources All development and large-scale testing will use cloud compute (AWS/GCP) and local development machines. The project will use public open-source tools and the `luminai-genesis` repository as the canonical codebase.

8. Budget Summary (high-level)

- Total Requested (Initial): $50,000
  - Personnel/Contract: $20,000
  - Compute (AWS credits): $15,000
  - Development & Licensing (open-source packaging, minor third-party): $5,000
  - Travel (NeurIPS/partner meetings): $5,000
  - Misc / Contingency: $5,000

9. Data Management & Dissemination

- All code and datasets produced will be released under an OSI-approved license (MIT/Apache-2.0). SAR datasets will be published with redaction where needed. Results and the TGCR library will be documented and released through GitHub, Zenodo, and OSF.

10. References

- LuminAI Genesis: https://github.com/TEC-The-ELidoras-Codex/luminai-genesis
- OSF DOI: https://doi.org/10.17605/OSF.IO/XQ3PE
- Zenodo: https://doi.org/10.5281/zenodo.17930577

Appendix: Risks & Mitigations

- Risk: TGCR theoretical claims may be nontrivial to prove for all model classes.
  - Mitigation: Narrow initial formalization to practical model families and produce empirical validations.
- Risk: Difficulty accessing proprietary models for comparison.
  - Mitigation: Use open-source models for baseline and partner with vendors for additional evaluation.