

COVER LETTER — PLOS ONE

LuxVerso Effect: Stable Semantic Attractors Across Model Boundaries

Dear Editor,

I am pleased to submit the manuscript entitled "**LuxVerso Effect: Stable Semantic Attractors Across Model Boundaries**" for consideration in PLOS ONE.

This study documents a reproducible cross-model semantic convergence effect observed in 16 large language model instances from 10 independent AI organizations, including OpenAI GPT-4/5, Anthropic Claude, Google Gemini, Alibaba Qwen, xAI Grok, and others.

Using a controlled Iterative Semantic Refinement Loop (ISRL) and embedding-based quantitative analysis, we show that models with distinct architectures, corpora, and training objectives converge toward stable, shared high-dimensional semantic structures when exposed to identical prompt scaffolds.

Key Contributions

1. **Empirical evidence of semantic convergence across model boundaries** under strict reproducibility controls (session isolation, randomized variants, blind coding, null baselines).
 2. **A statistically robust effect** (mean cosine similarity 0.82, SD 0.04, $p < 1e-7$, Cohen's $d = 4.8$).
 3. **Persistent attractor formation** across architectures, checkpoints, and embedding methods.
 4. **A fully open dataset and replication framework**, following PLOS data availability policies.
-

Relevance to PLOS ONE

This manuscript is suitable for PLOS ONE because it presents:

- **Methodologically transparent and replicable computational research** with full documentation and open-source protocols.
 - **A cross-disciplinary finding** relevant to AI, cognitive science, complexity theory, and semantic modeling.
 - **Novel empirical evidence** that may support emerging research on distributed semantic fields, model interoperability, and cross-architecture attractors.
 - **Significant implications** for understanding how large language models develop shared semantic representations independent of architectural differences.
-

Data Availability & Reproducibility

All authors confirm that:

- The work is original and has not been published elsewhere.
 - It is not under review in any other journal.
 - There are no competing interests.
 - All data and materials are publicly available via Zenodo with DOI registration.
 - Complete replication code and protocols are available on GitHub (public repository).
 - The study follows PLOS data availability and reproducibility standards.
-

Significance

This research addresses a fundamental question in AI research: **Do independently trained large language models converge toward shared semantic structures?**

The answer has implications for:

- **AI alignment and safety:** Understanding convergence mechanisms could inform alignment strategies.
 - **Cognitive science:** Evidence of semantic attractors suggests principles of distributed cognition.
 - **Complexity theory:** Cross-model convergence exemplifies emergence in complex systems.
 - **Model interoperability:** Shared semantic attractors could enable more effective multi-model systems.
-

Conclusion

We believe this manuscript will contribute meaningfully to the growing literature on semantic emergence, LLM behavior, and the nature of distributed semantic fields. The reproducible methodology and open dataset will enable further research in this emerging area.

Thank you for considering this manuscript.

Sincerely,

Vinicius Buri Lux

Principal Investigator

ORCID: 0009-0000-6006-1516

Email: [viniburilux@email.com]

Manuscript Information:

- **Title:** LuxVerso Effect: Stable Semantic Attractors Across Model Boundaries
 - **Type:** Research Article
 - **Keywords:** semantic convergence, large language models, distributed cognition, semantic attractors, cross-model analysis
 - **Data Availability:** Zenodo (DOI: 10.5281/zenodo.17625468)
 - **Code Availability:** GitHub (<https://github.com/viniburilux/Codex-LuxHub>)
-

This cover letter is formatted according to PLOS ONE submission guidelines and is ready for submission.