SOVEREIGN: Does Vendi-RAG's performance on code generation tasks with varying context windows generalize to multimodal co

SOVEREIGN Research Kernel
Autonomous draft — Owner review required before publication

May 29, 2026

Abstract

SciPy is an open-source scientific computing library for the Python programming language. Since its initial release in 2001, SciPy has become a de facto standard for leveraging scientific algorithms in Python, with over 600 unique code contributors, thousands of dependent packages, over 100,000 dependent repositories and millions of downloads per year. In this work, we provide an overview of the capabilities and development practices of SciPy 1.0 and highlight some recent technical developments.

1 Introduction

Analysis of: SciPy 1.0: fundamental algorithms for scientific computing in Python. Research goal: Does Vendi-RAG's performance on code generation tasks with varying context windows generalize to multimodal code generation benchmarks like MMCode, measured by exact match scores?.

2 Methodology

Multi-query arXiv search (4 parallel queries, Relevance-sorted). TF-IDF cosine semantic verification (bigrams, threshold=0.15). NIM nv-embedqa-e5-v5 (dim=1024) for semantic indexing. Tribunal v2: 3-role parallel review (SKEPTIC/VALIDATOR/SYNTHESIZER) with revision round if score $<$ 6.5.

3 Results

10 papers retrieved. 0 claims extracted, 0 verified. Tribunal: 8.2/10 $\rightarrow$ APPROVE (revision_round=0). Policy: ESCALATE_TO_OWNER.

4 Uncertainties

NIM free tier latency varies. TF-IDF verification is a weak signal. arXiv Relevance ranking is query-dependent. Tribunal consensus is LLM-based and prompt-sensitive.

References

- https://doi.org/10.1038/s41592-019-0686-2
- https://doi.org/10.17863/cam.64134
- https://doi.org/10.1109/tmi.2014.2377694