Colin M. Skinner

Walnut Creek, CA • skinner.colinm@gmail.com • (619) 251-0354
colinmskinner.com | github.com/SkinnerCM |linkedin.com/in/colin-mskinner

SUMMARY

Software engineer and Ph.D.-trained computational biologist building AI-powered software systems spanning backend services, cloud infrastructure, semantic and graph retrieval, multi-agent orchestration, and reproducible evaluation. Built a research platform indexing 35M+ publications and pipelines generating millions of benchmark examples.

TECHNICAL SKILLS

Languages: Python, SQL, Bash

AI & ML: PyTorch, Hugging Face, LLM APIs, Multi-Agent Systems, AI Evaluation, Embeddings, Vector Search, Semantic Retrieval, Knowledge Graphs

Backend & Data: FastAPI, REST APIs, PostgreSQL, PGVector, Neo4j, Cloud SQL

Cloud & Infrastructure: Google Cloud Platform (Cloud Run, Compute Engine, Cloud SQL, Cloud Storage, Secret Manager, IAP, Load Balancers), Docker, Linux, GitLab CI/CD

Frontend: Next.js, React, TypeScript, HTML, CSS

Developer Tools: Git, GitHub, GitLab, VS Code, Claude Code, Codex, Grok Build, MCP Servers, Jupyter

PROFESSIONAL EXPERIENCE

Software Engineer — Agents & Tooling Team | SpaceXAI • January 2025 – August 2026

- Architected and built Astral, an internal AI-assisted research platform for the Human Data team centered on a corpus of **35M+ scientific publications**, integrating semantic search, knowledge discovery, collaborative research tools, and agent-assisted workflows.
- Built cloud-native ingestion and enrichment pipelines using OpenAlex, Crossref, LLM-assisted metadata extraction, vector embeddings, PostgreSQL, Cloud Storage, and Neo4j to continuously process and index scientific literature.
- Designed and implemented a deterministic genomics benchmark pipeline that generated **millions of reproducible evaluation question-answer pairs** through configurable DNA mutation simulation, protein translation, and HGVS annotation, replacing a proposed LLM-based workflow.
- Developed literature retrieval and knowledge-discovery systems combining natural-language search, semantic similarity, graph analytics, citation analysis, PGVector, Neo4j, and Louvain community detection.
- Built collaborative PDF annotation and citation-aware research workflows that synchronized highlights into structured notes and enabled AI agents to organize and expand research documents.
- Engineered a hierarchical Director/Project Manager/Worker multi-agent orchestration platform with a custom episodic-memory layer, enabling concurrent management of up to **14 engineering projects**.

- Built and operated secure production infrastructure on Google Cloud Platform using Docker, Cloud Run, Compute Engine, Cloud SQL, Secret Manager, IAP, load balancing, FastAPI services, and GitLab CI/CD.
- Supported a roughly 200-person internal research organization, with Astral reaching **40–50 weekly active users** and becoming a default starting point for literature discovery and dataset creation.

SELECTED PROJECTS

Aria | Independent Multi-Agent Engineering Orchestrator

- Built cross-model agent collaboration across coding assistants using shared chat, agent-to-agent messaging, independent terminal interfaces, and role-based approval gates for implementation and testing.
- Implemented episodic memory distillation from completed task cards and two proposal-based improvement loops: learning from past execution and discovering new agent methods from external research, with human approval.
- Created a voice-enabled mobile progressive web app accessed securely through Tailscale for remote task launch, status updates, and workflow control.

AI News Intelligence Platform | colinmskinner.com

- Built a news ingestion and filtering pipeline across **16 sources**, with cached headlines and summaries powering an interruptible
  real-time voice briefing agent that supports contextual questions, web search, and resuming its scripted brief.
- Developed embedding-based topic visualization using **UMAP** and momentum comparisons across 24-hour, 3-day, 7-day, 30-day, and 90-day windows to surface emerging, sustained, and declining trends.

EDUCATION

Ph.D., Bioengineering (Computational Biology), UC Berkeley / UCSF, 2025 | NIH T32 Predoctoral Fellow

B.S., Biomedical Engineering, UC Berkeley, 2017 | Jack Kent Cooke Scholar

SELECTED PUBLICATIONS

- Skinner, C.M., et al. DNA methylation clocks struggle to distinguish inflammaging from healthy aging, but feature rectification improves coherence. Geroscience, 2025.
- Mehdipour, M., Skinner, C.M., et al. Small-animal blood exchange for systemic aging research. Nature Protocols, 2022.
- Mehdipour, M., Skinner, C.M., et al. Rejuvenation of three germ-layer tissues via plasma exchange. Aging, 2020.

HONORS

- xAI STEM Tutor Employee of the Month (2026)|NIH T32 Predoctoral Fellowship | Jack Kent Cooke Scholarships