Tao Ren (Tempest)

AI Engineer & Researcher | Agents, structured generation, and evaluation

tar118@pitt.edu | tempest.fun | GitHub | LinkedIn

Profile

AI engineer and independent researcher building agents for document-heavy workflows and software engineering. Work spans insurance AI at Newfront, an internal coding agent at TikTok, and research on precise editing, structured code generation, and tool-using agents. Bring a full-stack engineering background to connecting model capabilities with tools, evaluation, production services, and human review.

Agent Engineering Experience

Newfront | Senior Full-stack Software Engineer

October 2025 - Present

- Work on Nova Agent, an internal AI coworker for insurance operators, supporting document-heavy tasks, contract generation, company-style email drafts, and human-in-the-loop review.
- Develop agent workflows involving more than 50 PDFs per task, policy and submission documents, generated artifacts, and internal services; account for permissions, domain constraints, and operator review.
- Work on agent system design involving tools, skills, memory, document processing, and structured outputs. Experience with quoting and placement services informs integration with insurance workflows.

TikTok | Software Engineer, Social and User Relations

Prior to October 2025

- Built an internal LLM coding agent from scratch for Objective-C-to-Swift migration, connecting code generation with iterative quality checks and custom evaluation harnesses.
- Iterated on prompts and generated code to assess practical usefulness for engineers performing migrations.
- Built consumer social interaction features and mobile product systems, grounding developer tooling work in production engineering requirements.

Tira AI | Co-Founder

June - August 2025

- Co-founded an AI-native execution startup; drove product architecture, implementation, and early go-to-market work.
- Developed product execution workflows from initial concept through implementation, combining full-stack engineering with rapid product iteration.

Core Capabilities

- **Agent systems:** Tool use, document workflows, memory, task decomposition, structured outputs, and human-in-the-loop review.
- **Evaluation and research:** Code-generation evaluation harnesses, prompt iteration, precise editing, AST-guided generation, and real-time AI systems.
- **Implementation:** Python, TypeScript, JavaScript, SQL; React, Next.js, React Native, Node.js, and Python backend frameworks.
- **Production foundations:** API integration, PostgreSQL, MySQL, Redis, Docker, AWS, Linux, and Git.

Selected Research

Co-authored work on structured generation, editing, tool use, and runtime constraints.

FineEdit | Findings of EMNLP 2025

Bridging the Editing Gap in LLMs: FineEdit for Precise and Targeted Text Modifications

Research on precise instruction-driven editing across code, LaTeX, and database languages, connecting specialized editing models with evaluation of structured modifications.

TreeDiff | SURGeLLM at ACL 2026

TreeDiff: AST-Guided Code Generation with Diffusion LLMs

Research incorporating abstract syntax tree structure into diffusion-based code generation to address syntactic boundaries and hierarchical dependencies.

Infant Agent | arXiv preprint, 2024

Infant Agent: A Tool-Integrated, Logic-Driven Agent with Cost-Effective API Usage

Research on tool-integrated agents with task-aware functions, memory retrieval, and cost-aware API usage for multi-step tasks.

RED | RTSS 2023

RED: A Systematic Real-Time Scheduling Approach for Robotic Environmental Dynamics

Research on robotic inference scheduling under changing environmental conditions, connecting AI execution with latency and runtime constraints.

Selected Product

AgentShelf | AI Commerce Readiness Console

github.com/tempest2023/AgentShelf

Built a hackathon product exploring inspectable AI commerce workflows: product readiness, shopping-query simulation, structured content, and launch preparation. The demo includes a mock publishing flow.

Earlier Engineering & Research

- **University of Pittsburgh | Research Assistant, Full-stack Engineer | From August 2021:** Built rehabilitation research platforms for data collection, visualization, wearable integrations, and configurable study workflows in Prof. Dan Ding's lab.
- **Ansys | Full-stack Software Engineer Intern, Digital Twin | May - August 2022:** Built Twin Deployer UI, APIs, and data models; helped move Electron-based file operations toward browser-compatible deployment workflows.
- **Ant Group / Alibaba | Software Development Engineer | July 2020 - June 2021:** Built low-code financial storefront infrastructure, live-streaming components, and backend services for Alipay; contributed to an OceanBase migration.

Education

- **University of Pittsburgh** - MS, Information Science | 2021 - 2023
- **Southern University of Science and Technology** - BEng, Computer Science and Technology | 2016 - 2020