Handuo Zhang

AI RESEARCH SCIENTIST | AGENTIC SYSTEMS

Singapore · Singapore Permanent Resident · +65 8341 0772

zhanghanduo@gmail.com · handuo.top · github.com/zhanghanduo

AI scientist with a PhD in robot vision, specializing in agentic systems, tool-use post-training, and reliable evaluation. Builds production infrastructure and open-source tools for complex, long-running work.

SELECTED HIGHLIGHTS

4.4K GitHub stars FrontierAgent · Initiated and led an open-source agent runtime, terminal product, and evaluation suite with single-agent and Agent Team workflows.

No. 1 on BFCL V3 Led Watt's Llama 3.3-70B tool-calling roadmap and fine-tuning; reached the top of Berkeley's Function-Calling Leaderboard in December 2024.

AgentOS founding engineer One of three founding engineers of Prax / AgentOS and technical owner of ApodexHarness, building multi-agent orchestration, memory, and verification from the ground up.

PROFESSIONAL EXPERIENCE

Apodex · Singapore 2026 - Present

AI Research Scientist · Post-training Team

- **Agent architecture.** Built heavy_mode for complex, long-running tasks, combining parallel sub-agents, cross-branch synthesis, global verification, and continuation after interruption. Delivered the orchestration stack from execution engine to SDK and event streaming.
- **Workflow execution.** Proposed planner-generated, schema-constrained workflow DAGs. Replaced LangGraph with a compact execution engine supporting branching and verification-retry loops; published shared AgentCore infrastructure to PyPI.
- **Memory and trustworthy output.** Built token-budgeted context compression, file-based research memory, and tool-specific result budgets. Introduced evidence checks to block unsupported numbers and citations in generated reports.
- **Production reliability.** Strengthened sandbox isolation, Bash permissions, durable sub-agent messaging, and cancellation-safe event streams. Fixed silent delivery failures and resource-management bottlenecks.
- **Evaluation and model releases.** Integrated agent workflows with a unified evaluation interface; completed 240 APEX Agents 1.1 tasks end to end. Diagnosed infrastructure-induced scoring failures and contributed ablation evidence to Apodex Discovery and harness support to Apodex 1.1/1.2 releases.

Shanda Group / Watt · Singapore 2023 - 2026

Vice President of AI · Watt, a Shanda-incubated startup

- **Tool-use post-training.** Built workflows supporting 200+ tools, raising task completion from 70% to 90% on 1,000 multi-step tasks. Fine-tuned Llama 3.3-70B for parallel tool use and multi-turn instruction following.
- **GMate product delivery.** Evolved GMate into a closed-loop agent with preference, factual, and experiential memory. Led launch reliability for Xiaohongshu: 7,875 uses and 426 active users in December 2025.
- **Multi-agent commerce.** Architected GMate's TikTok Shop assistant with LangGraph and DOM-plus-vision browser automation; built end-to-end observability and retrieval-enhanced evaluation.
- **Retrieval systems.** Built hybrid retrieval and reranking over a 6,000-document recruiting knowledge base and a 500,000-candidate talent graph; improved match success from 75% to 90% with matching latency below five seconds.

RESEARCH PAPERS AND TECHNICAL REPORTS

ICLR 2027 manuscripts in submission

FrontierChallenge: Evaluating Scientific Workflow Completion
2026 · Manuscript

AgentAgora: Facilitated Multi-Agent Discussion for Creative Generation
2026 · Manuscript

CreativitySuite: A Comprehensive Framework for Evaluating and Enhancing Creativity in Large Language Models
2026 · Manuscript

Technical reports

Apodex 1.1: Scaling Agentic Intelligence for Complex Work
2026 · Contributing author · arXiv technical report | arxiv.org/abs/2608.23283

Apodex Discovery: Reality Benchmarks and Environments for Evaluating and Building Discoverative Artificial Intelligence
2026 · Co-author · arXiv technical report | arxiv.org/abs/2608.11341

Earlier research includes 10 publications in robot perception, visual localization, and tracking, including IROS, IEEE T-ASE, CVIU, and IEEE Access.

TEACHING AND EARLIER EXPERIENCE

Nanyang Technological University · Singapore Apr 2023 - Present

Part-time Lecturer

Teach postgraduate EE7207 Neural Networks and Deep Learning, covering attention, transformers, graph neural networks, and practical training and inference.

Mind Pointeye · Singapore Apr 2021 - Aug 2023

AI Scientist / Technical Lead

Led seven AI and data engineers delivering video analytics and IoT systems. Reduced edge model size 4× and accelerated inference 3×; achieved 95% accuracy in action and intrusion detection.

Nanyang Technological University · Robotics I Lab Sep 2015 - Jun 2021

Project Officer and Doctoral Researcher

Led a three-PhD research team on stereo perception for unmanned ground vehicles; developed real-time stereo vision, multi-object tracking, and visual SLAM systems.

EDUCATION

PhD, Robot Vision and Control · Nanyang Technological University · 2016 - 2022

MSc, Pattern Recognition and Intelligent Systems · Northeastern University · 2011 - 2013

BEng, Automation Control · Northeastern University · 2007 - 2011

TECHNICAL EXPERTISE

Agent systems Multi-agent orchestration, DAG execution, memory, context compression, verification, MCP, browser automation, sandboxing, tracing, LangGraph

Models and retrieval SFT, RL, tool calling, PyTorch, TRL, DeepSpeed, LLaMA-Factory, hybrid retrieval, reranking, knowledge graphs

Engineering Python, C/C++, SQL, Docker, SDK development, evaluation infrastructure; earlier work in ROS, SLAM, and computer vision