SIMERJIT KAUR NATH

Singapore · simernath05@gmail.com · simerjit001@e.ntu.edu.sg
LinkedIn · GitHub · Website · Google Scholar

MSc AI candidate at NTU (Jan 2027) who builds AI systems that hold up on real data. LLM applications and multi-agent pipelines, multimodal deep learning for city-scale digital twins at A*STAR, and efficient fine-tuning; first-author paper accepted at the SENTIRE Workshop, IEEE ICDM 2026. 2 years at NielsenIQ delivering research for global consumer brands and leading an 11-person team.

PUBLICATIONS

Nath, S. K., Cambria, E., Pen, H. *SenticLens: A Production Sentiment-Analysis Framework with Behavioral Auditing of LLM Affective Coherence.* SENTIRE Workshop at IEEE ICDM 2026 (ICDM is ranked CORE A*). Proceedings of the 2026 IEEE International Conference on Data Mining Workshops (ICDMW), IEEE; Scopus-indexed. **Accepted (in press).** Sentic Labcast (video)

EXPERIENCE

Graduate Researcher, SenticNet Lab, NTU *Aug 2025 – Present*

- Built **SenticLens**, an end-to-end sentiment and LLM evaluation framework ingesting live reviews from **3 platforms** (Reddit, Apple App Store, Google Play), scoring polarity, intensity and emotion, with guardrails flagging inconsistent model outputs before results reach downstream analytics.
- Across **24,000 audited records**, showed models disagree with themselves **twice as often** on real reviews as on synthetic ones (38.16% vs 18.49%) and contradict themselves outright **about 6× as often** (2.32% vs 0.37%), on every model and brand.
- Designed a **label-free consistency check** that needs no ground truth: on **10,000 public benchmark reviews with star ratings**, answers it flagged were wrong **9–25× as often** as unflagged ones.

Research Intern, A*STAR Centre for Frontier AI Research (CFAR) *Nov 2026 – Jan 2027*

Lead Software & AI Engineer (part-time), Kenza Living (Kenza AI) *Dec 2025 – Apr 2026, Aug – Oct 2026*

- Led the end-to-end build of a **bilingual (English/Arabic) community-management platform** for residential communities: **React Native** mobile app, **Node.js/Prisma** backend and **PostgreSQL on AWS**, now **live on the App Store, Google Play and the web**.
- Built an **LLM assistant** that turns resident messages into amenity bookings and service tickets, with server-side guardrails for security (authentication, role checks, rate limiting) and for the booking use case (booking details from server data, explicit confirmation, ticket fallback) to reduce incorrect actions.

Research Intern, A*STAR Institute of Advanced Intelligence and Computing (IAIC) *Apr – Aug 2026*

Computational Sustainability

- Built and validated the **4-source** urban-climate ingestion pipeline and set up the quality-control procedures.

Graduate Student Research Assistant, NTU School of Social Sciences (part-time) *Feb – Apr 2026, Aug – Oct 2026*

- Built a local, offline LLM workflow for confidential interview transcripts with automated alignment and scoring, using LLM-assisted extraction, HDBSCAN with UMAP and Ward-linkage clustering.

Research Associate → Research Analyst → Team Lead, NielsenIQ, Mumbai, India *Jul 2023 – Jul 2025*

- **Promoted twice in 2 years**, rated **Exceeded Goals** and named a **top performer**; led an **11-member** ARC Europe team as Team Lead.
- Delivered concept analysis and market research behind **100+ FMCG launches** for brands such as Reckitt and Nestlé, at **100% accuracy** and a **3.9/4.0** project rating, resolving data issues under tight client timelines.

SELECTED TECHNICAL PROJECTS

- **RAG grounding audit** (open source, github.com/simer05/rag-grounding-audit, 2026): tested **3 open-weight LLMs** on **500 HotpotQA questions** under 6 controlled evidence conditions. Reordering the same passages flipped **8–14%** of answers between right and wrong while accuracy stayed flat, and models answered without evidence **26–40%** of the time.
- **Agentbox**, open-source contributor (2026–present, github.com/savannahfeder/agentbox): desktop app that runs many Claude Code and Codex agents in parallel, each in its own git worktree. First contributor to bring **Windows support**: traced why agents could not start on Windows, then fixed how Claude Code and Codex launch, how tool servers and per-task folders work, and the in-app terminal; merged upstream.
- **AdaptLink, AI seller intelligence for Shopee** (Sea x OpenAI Codex Hackathon, Regional Finalist, team of 4, Jun 2026): came up with the product idea and built the seller-intelligence dashboard (TypeScript/Next.js), turning conversion, reviews, buyer questions, reply times and competitor prices into a product health score and up to **5 margin-safe actions**; designed a **6-agent** seller workflow with model routing.

EDUCATION

Master of Science in Artificial Intelligence *Aug 2025 – Jan 2027 (expected)*

Nanyang Technological University, College of Computing and Data Science, Singapore

Bachelor of Science in Computer Science *Jul 2020 – Jun 2023*

SIES College, University of Mumbai, India · CGPA 9.63/10 (WES 3.86/4.00) · ***Silver Medalist***

AWARDS, LEADERSHIP AND SERVICE

Regional Finalist, Sea x OpenAI Codex Hackathon (2026) · Top 6 Finalist, American Express GenAI Hackathon (2025) · 2nd Place, TE Connectivity x AWS Hackathon (2025) · Bravo! Superstar, NielsenIQ/ARC Excellence Awards · Vice Chair, Operations and Logistics, TEDxNTU 2026 · Communication Coach, NTU CommCube (2026–Present): **95.6%** session satisfaction

TECHNICAL AND RESEARCH SKILLS

AI/ML: LLM evaluation and observability, guardrails, behavioural auditing, prompt engineering, supervised fine-tuning (SFT) with LoRA/QLoRA/PEFT, quantisation and 4-bit inference, deep learning, multimodal learning, NLP, sentiment and emotion analysis, diffusion models, physics-informed ML, time-series analysis, inference optimisation

ML systems: local and self-hosted deployment, multi-agent and agentic systems, synthetic-vs-real evaluation, surrogate and reduced-order modelling for digital twins, experimental design, responsible AI, reproducibility

AI coding tools: Claude Code, OpenAI Codex, Cursor, GitHub Copilot; hands-on with Claude, GPT, Gemini, DeepSeek and open-weight models

Tools: Python, TypeScript, React Native, Node.js, Next.js, PostgreSQL, PyTorch, TensorFlow, Hugging Face, scikit-learn, pandas, SQL, Git, Docker, AWS, Linux/HPC, Tableau, Power BI, Advanced Excel