

Gourob Debnath

AI Software Engineer and Graduate Researcher, EARL Research Network - Agentic AI, LLM Security

gourobdn15@gmail.com | [linkedin.com/in/gourob-debnath-b41730165](https://www.linkedin.com/in/gourob-debnath-b41730165) | github.com/Gourob47 | [Portfolio](#) | [ORCID: 0009-0002-2794-9374](#) | [ResearchGate](#)

Research Interests

Agentic AI and multi-agent system architecture, including multi-agent orchestration, long-term memory systems, tool-use safety, and governance frameworks for production LLM deployments. LLM security, including jailbreaking, prompt injection, and contextual deception, with implications for AI safety in regulated industries such as healthcare and finance.

Education

B.Sc. in Computer Science & Engineering

Metropolitan University, Sylhet | CGPA: 3.54 / 4.00

Publications

- “Agentic Artificial Intelligence: A Survey of Architectures, Evolutionary Dynamics, and Governance,” submitted to Information Processing & Management (Elsevier), 2026. (Under review)
- “Code Poisoning Through Misleading Comments: Jailbreaking Large Language Models via Contextual Deception,” IEEE ICCIT 2025. Designed a novel jailbreak attack via contextual comment poisoning, evaluated across GPT-4, LLaMA, and Gemini using a Python-based framework (PyTorch, OpenAI API).

Dataset: LLM Comment Vulnerability Dataset (200 adversarial prompts across 7 harm categories), published on Zenodo, DOI: [10.5281/zenodo.15786008](https://doi.org/10.5281/zenodo.15786008), supporting the ICCIT 2025 paper above.

Professional Experience

Graduate Researcher, EARL Research Network (earl-research.org)

- Member of a multi-lab collaborative research network. Co-authored work with network researchers now at University of Tsukuba, University of Houston, and other institutions.

Software Engineer, SJ Innovation LLC | 2024 – Present

- Designed the multi-agent architecture behind Sara, a production AI healthcare assistant serving 500+ users, orchestrating specialized agents for intent analysis, medical context summarization, retrieval, and response generation using the OpenAI Assistants/Chat Completions APIs with structured outputs and guardrails.
- Contributed to Model Context Protocol (MCP) integration connecting agents securely to internal healthcare APIs, and integrated Mem0 as a long-term memory layer enabling persistent, personalized context across sessions.
- Built asynchronous FastAPI backend services with Auth0/JWT authentication, PostgreSQL and Amazon DocumentDB storage, and Braintrust-based tracing for AI observability and prompt evaluation.
- Contributed to an AI-powered multimodal content platform (ImagineCreate AI) and an enterprise financial platform (ICR Capital), building AI-assisted analysis and a bidding management system.

Selected Projects

AssesGen : AI-Powered IELTS Question Generation & Evaluation Platform *(In development, building for production)*

- Building a multi-agent AI system for automated IELTS question generation and evaluation, powered by locally hosted models (Gemma for generation/evaluation, a local TTS model for speech) running on-device rather than external cloud APIs, with agent orchestration for content generation, scoring, and feedback pipelines.

Founder and Developer, IELTS Ready

- Designed and built a multi-tenant SaaS platform for IELTS preparation with automated scoring, mock/partial examinations, instructor dashboards, and role-based access control.

Technical Skills

- AI/LLM: OpenAI Assistants & Chat Completions APIs, multi-agent systems, prompt engineering, tool calling, structured outputs, guardrails, MCP, Mem0, local/on-device model deployment (Gemma, TTS), multimodal (image/video) generation
- Languages: Python, JavaScript/TypeScript, PHP, C++
- Backend & Data: FastAPI, Node.js/Express, Django, REST APIs, async programming, PostgreSQL, MongoDB, Amazon DocumentDB, MySQL
- Infrastructure: AWS (incl. Secrets Manager), Docker, Auth0/JWT, Braintrust, Git/GitHub

Achievements

- Solved 750+ problems across competitive programming judges. Participant, ICPC Dhaka Preliminary (multiple years).
- Judge & Problem Setter, MU Intra Programming Contest.