

Austin Senna Wijaya

(+1) 646-295-5341 | asw2215@columbia.edu | [linkedin.com/austin-senna/](https://www.linkedin.com/austin-senna/) | [Github](#) | [Portfolio](#)

EDUCATION

Columbia University

Bachelor of Science: Computer Science

New York, NY

Aug 2024 - Expected May 2027

- GPA: 4.25/4.0; Full-Tuition Scholarship; Dean's List
- Relevant Coursework: Natural Language Processing, Data Structures & Algorithms, Systems Programming, Linear Algebra & Probability, Competitive Programming.

EXPERIENCE

Data Agents Research Assistant

Columbia Data, Agents, and Processes Lab

Dec 2025 – Present

New York, NY

- Contributing to DataSeek, a system providing infrastructure for deep-research agents to operate over TBs of multimodal data. Submitted a benchmark paper evaluating agents on this process to **ICML 2026**.
- Re-architected a legacy discovery tool into a scalable Agentic GraphRAG engine; reduced retrieval complexity to $O(N \log K)$ via LSH and Min-Heap Top-K pruning, increasing accuracy by 10% across a 10TB S3 data lake.
- Engineered a parallelized benchmark evaluation framework using **AWS Bedrock** to test autonomous data tasks and multi-step reasoning (downloading, filtering, SQL querying); added agentic tool capabilities to **DeepSeek R1**.
- **Tech Stacks:** Python, SQL, AWS (EC2, S3, Bedrock), Docker

Machine Learning Research Assistant

Columbia Zuckerman Institute

Dec 2025 – Present

New York, NY

- Trained linear probes via Ridge Regression on external audio benchmarks to decompose Qwen-Audio-2-7B's latent space into orthogonal feature vectors (paralinguistic and linguistic), mapping these embeddings to ECoG neural signals.
- Engineered an end-to-end, GPU-optimized pipeline deployed on **HPC clusters (NCSA Delta)** to parallelize Text-to-Speech (TTS) generation and extract penultimate layer embeddings across hundreds of thousands of inferences.
- **Tech Stacks:** Python, PyTorch, HuggingFace, Scikit-learn, NumPy

Data Engineering Intern

Ruangguru (Biggest Indonesian EdTech Platform)

Jun 2025 – Aug 2025

Jakarta, Indonesia

- Consolidated millions of payment logged events into user sessions in **BigQuery**, producing a product-level funnel used to diagnose major drop-off points.
- Automated large-scale Google Review scraping and sentiment analysis with **Puppeteer** and **HuggingFace**, enabling continuous customer satisfaction tracking.
- Processed hundreds of thousands of tutor-scheduling events with **Apps Script** and **BigQuery** to generate a real-time interactive **Looker Studio** tutor availability heatmap used by operations teams.
- **Tech Stacks:** Google Cloud (BigQuery, Looker Studio), Apps Script, Puppeteer, HuggingFace

PROJECTS

ResearcherX: AI-Powered IDE & GraphRAG Engine

[Github](#)

- Developed an AI-powered IDE for academic writing that performs real-time logical contradiction linting; orchestrated a dual-routing pipeline (**FastAPI**, **litellm**) to optimize API costs and latency between local 4B and frontier models.
- Engineered an asynchronous hybrid-search engine (**Neo4j**, **LanceDB**) featuring node-level provenance to mitigate LLM hallucinations and **Cypher**-based garbage collection to maintain strict graph integrity.
- **Tech Stacks:** Python (FastAPI), TypeScript (Next.js), Neo4j (Cypher), LanceDB, litellm, ProseMirror

Untukmu Karyamu (Tencent Hackathon)

[Github](#)

- 2nd Place (Best Code, Best Product Idea, Popularity Award) at **Tencent Kepler Plan S3 Competition 2025**.
- Architected a generative web-builder utilizing **Next.js** and **GeminiAPI** to automate site deployment to **Tencent Edgeone**, minimizing launch latency and infrastructure overhead for MSMEs.
- **Tech Stacks:** Next.js, Supabase, Puppeteer, Edgeone Pages

TECHNICAL SKILLS

Languages: Python, SQL, JavaScript, Java, C, Rust

Libraries & Frameworks: Pandas, NumPy, PyTorch, Scikit-learn, Puppeteer, Selenium, BeautifulSoup, React, Next.js, Node.js

Cloud & Databases: Google Cloud (BigQuery, Looker Studio), AWS (EC2, S3, Bedrock), Supabase, PostgreSQL, MongoDB

Tools: GitHub, Docker, Claude Code, n8n, Copilot, Arduino, Notion

HONORS AND SCHOLARSHIPS

Hack@Brown 2026: Finalist and awarded 'Strongest Product Thinking', working with **Google Ventures** and **Partiful**.

Clash of Champions Season 2: Top 9 in a reality academic survival show for top Indonesian students.

Indonesia Maju Scholarship: Full-ride scholarship recipient (Ministry of Education of Indonesia), 1/350 recipients.

55th International Chemistry Olympiad 2023: Final Stage (10th Rank) for National Team Selection.