

Patrick Amadeus Irawan

✉ patrickamadeusirawan@gmail.com 🌐 Site in LinkedIn 🐙 Github

Summary

Research Engineer specializing in **multilingual & multimodal NLP** studies, and **backend engineering**. Proven track record of developing AI-driven solutions at scale and publishing at top-tier conferences. Previous accomplishments:

- Leading and contributing to multiple research projects accepted at **A-tier** NLP conferences (*i.e.*, EMNLP, NAACL, COLING).
- Completed **5+ internships** in Data Science and Software Engineering across diverse industries.
- Strengthened my academic foundation through **AI lab teaching assistantships** during my undergraduate studies.
- Earned multiple cross-disciplinary **awards (consulting, tech, cultural exchange, management)** at both national and international levels.

Education

MBZUAI (Incoming) Aug 2025
Ph.D. in Natural Language Processing

Institut Teknologi Bandung Aug 2020 - Oct 2024
B.Eng. in Computer Science

- GPA: 3.88/4.0 (Magna Cum Laude)
- **Notable Highlights:** AI & ML Teaching Assistant, Global Study @ UC Davis ([News](#) 📄), CSLeaders Scholarship, [Data Science Professional License](#) 📄

University of California, Davis July 2023 - Dec 2023
Global Study Program

- GPA: 3.94/4.0 (Academic Excellence)
- Accomplished 6 multi-discipline courses under UC Global Study Program
- **Notable Courses & Activities:** Management Tech, Computational Neuroscience, UC AI Symposium

Experience

Research Engineer Singapore, Remote
Singapore Management University **Feb 2025 - current**

- Part of the Multimodal SEA-Cultured LLM research team, spearheading the corpus creation provisioned by Prof. Chong-Wah Ngo
- Designing and executing large-scale multilingual user studies (projected 1M+ data points) to improve multimodal NLP datasets using Huggingface and CVAT.
- Leading a team of 3 users on implementing an end-to-end multimodal reasoning data consolidation pipeline (data collection, quality assurance, and deployment).

Software Engineer Germany, Remote
IT Bauschmiede **Nov 2024 - Jul 2025**

- Developing a project management system with real-time scheduling, invoicing, and geolocation, serving 2+ enterprise clients.
- Implementing a scalable full-stack solution using ReactJS & Django, reducing processing time by 20% and enhancing system reliability.

Research Engineer Intern Jakarta, ID
Ai@You **May - July 2024**

- Developed LLM/VLM-based software prototypes for 3+ automotive and SaaS industry clients.

- Co-led domain specific company's portfolio system design with chatbot platform & full-stack website using NextJS and LLM API services.
- Implemented end-to-end RAG system using Botpress, Streamlit, and Vector DB, including programmed its custom plugins to suit niche retrieval needs.

Data Scientist Intern

Jakarta, ID

Supertype

Jan – July 2023

- Redesigned topic extraction pipeline using Transformers & PyTorch, improving processing speed by 100% and reducing redundancy by 30%.
- Optimized Django REST + GCP backend, ensuring 95%+ uptime and handling more API calls per day.
- Improved a time-series forecasting model for lifetime prediction, increasing accuracy by 5%.

Software Engineer Intern

Jakarta, ID

Blibli

June – Dec 2022

- Expanded and debugged 4+ services using Java and Selenium with BDD practice, producing 150+ new QA cases.
- Optimized Kafka data ingestion pipeline by eliminating code smell, boosting test efficiency by 10%.
- Ensured resources' quality of multiple environments, reducing core automation service erroneous test rate by 35%.

Software Engineer Intern

Jakarta, ID

Ruangguru

Jan – April 2022

- Involved in intensive 3-week e2e BE training in Golang, design patterns, infrastructure, and monitoring, achieving an 85% task completion rate.
- Engineered a coding judge software for over 1k+ users using Golang, boosting the cohort's task completion rate by 15%.
- Optimized codebase through Golang semantics & templates research and application.

Publications

WorldCuisines: A Massive-Scale Benchmark for Multilingual and Multicultural VQA on Global Cuisines Jan 2025 - NAACL

Genta Indra Winata*, Frederikus Hudi*, *Patrick Amadeus Irawan**, David Anugraha*, Rifki Afina Putri*, *and others**

[WorldCuisines Webpage](#) [🔗](#)

ProxyLM: Predicting Language Model Performance on Multilingual Tasks via Proxy Models Jan 2025 - NAACL

David Anugraha, Genta Indra Winata, Chenyue Li, *Patrick Amadeus Irawan*, En-Shiun Annie Lee

[Arxiv](#) [🔗](#)

Towards Efficient and Robust VQA-NLE Data Generation with Large Vision-Language Models Jan 2025 - COLING

Patrick Amadeus Irawan, Genta Indra Winata, Samuel Cahyawijaya, Ayu Purwarianti

[COLING 2025 Index](#) [🔗](#)

Leveraging IoT and Machine Learning for Efficient Rice Stock Monitoring and Prediction Dec 2024 - APSIPA ASC

Nana Sutisna*, Aditya Prawira Nugroho*, Christopher Jeffrey*, *Patrick Amadeus Irawan**, Rizky Ramadhana*, Ronggur Mahendra*, *and others**

[IEEE](#) [🔗](#)

SEACrowd: A Multilingual Multimodal Data Hub and Benchmark Suite for Southeast Asian Languages June 2024 - EMNLP

Holy Lovenia*, Rahmad Mahendra*, Salsabil Maulana Akbar*, Lester James Validad Miranda*, Jennifer Santoso*, Elyanah Aco, **Patrick Amadeus Irawan**, and others
[EMNLP 2024 Index](#) [🔗](#)

Leadership, Incubation, & Volunteer Experience

Core Platform Team

Jan 2025 - current

[Grassroots Science](#) [🔗](#)

- Engineered an annotation platform designed to support 50+ concurrent global users for multilingual LLM development.
- Led the backend team in developing concurrent multi-turn RLHF platforms.

Data Analytics Lead

2021 – 2024

[Internnet Indonesia](#) [🔗](#)

- Led a team of 3 in data warehousing and analytics, securing 6 partnerships and funding, including collaborations with MNCs.
- Designed a mentee-mentor matchmaking algorithm, eliminating 95% of manual processing.

Director of Operations

Spring 2022

[GICC by Campus Cabinet \(KM ITB\)](#) [🔗](#)

- Managed an 8-person engineering team to develop a MERN-based website for 200+ participants.

Project Associate

2021 – 2022

[ShARE ITB](#) [🔗](#)

- Served as the data analytics champion, providing KYC service expansion insights for Verihubs across 4+ SEA countries.

Achievement & Awards

SEA Emeralds Fellow (1st Place in Exit Strategy Case)

Jakarta, Indonesia

Boston Consulting Group (BCG)

2024

Indonesian International Student Mobility Awards

UC Davis, USA

Ministry of Education, Indonesia

2023

Finalist of DOMO Case Competition (Higher Education)

Utah, USA

DOMO

2023

TASLA Merit-based Techno-preneurship Education Scholarship

Indonesia

Monk's Hill Ventures, Glints

2022

Top 10 of Data Mining Gemastik XV

Malang, Indonesia

Ministry of Education, Indonesia

2022

Technologies

Languages: Python, Golang, Java, SQL, JavaScript, TypeScript

Tools & Frameworks: PyTorch, Scikit-Learn, Huggingface, ReactJS, NextJS, FastAPI, Django, REST, Flask, Gin, Echo, Google Cloud Platform (GCP), Google BigQuery, Google Datastudio, Looker Studio, Tableau, PowerBI, Alicloud, Git, Jenkins