

Wilfried Ariel Mulyawan | 陳順福

Singapore | +65 8954 5377 | arielwilfried0@gmail.com | linkedin.com/in/wilfriedariel

RESEARCH INTERESTS

Quant Finance and FinTech, AI Real-World Models, LLM and VLM Benchmarking, Machine Unlearning, AI Ethics and Accountability.

EDUCATION

Ph.D. Candidate, College of Computing and Data Science, Nanyang Technological University (NTU), Singapore 2026 – Present

- Advisor: Prof. Lin William Cong (Founding Director, GIFTS). Co-advisor: Prof. Sean Du (Forbes 30 Under 30 Asia, Science).
- Lab affiliations: GIFTS, ADEFT Lab, Radio Lab. Lab Coordinator for I3 Lab.

B.Sc. in Engineering Management, Bandung Institute of Technology (ITB), Indonesia

2020 – 2025

Semester Exchange, International Advanced Technology Program, NTUST, Taiwan

Aug 2023 – Dec 2023

RESEARCH EXPERIENCE

Ph.D. Researcher, GIFTS Lab, Nanyang Technological University, Singapore

May 2026 – Present

- Research on AI real-world models for finance and economics, machine unlearning, and LLM benchmarking / question answering.
- Main project: AI in Finance and Economics. Delegated Financial Cognition, studying correlated errors across AI financial advisors and what they imply for portfolio-level risk when many investors delegate to the same models; also microstructure/macroeconomics.
- Coordinator for the I3 Lab: run cross-lab coordination and reporting across GIFTS, ADEFT and Radio Lab research tracks.

AI Research Engineer, Singapore Management University, Singapore | Hybrid

Feb 2025 – Sep 2025

- Built a 1,000+ question multimodal dataset benchmarking Vision-Language Models on ASEAN cultural reasoning, leading data collection, annotation, and validation under Professor Ngo Chong Wah and the SMU Multimedia Research Team.
- Co-authored "Seeing Culture: A Benchmark for Visual Reasoning and Grounding," accepted at EMNLP 2025 Main Conference, the first benchmark stress-testing VLMs on visually grounded reasoning across nine Southeast Asian cultures.
- Developed a teacher-facing grading platform built on multimodal LLMs that auto-evaluates student submissions and surfaces learning gaps, reducing grading time and giving students faster feedback.
- Ran evaluation cycles on a multilingual model to improve cultural relevance, analysing failure cases and reporting weekly findings on model weaknesses that fed directly into the next training iteration.

Lab Assistant, Lab of Innovation and Corporate System Development, ITB

Apr 2023 – May 2024

- Built an adaptive crypto trading system on a self-hosted server, training XGBoost and other scikit-learn models on real-time market data with automated re-learning cycles that improved entry and exit timing on leveraged positions.
- Investigated how consumers respond to AI-generated marketing versus human-made content using PLS-SEM and SEM-MGA across two real working industry case studies in Indonesia.
- Taught independent classes in valuation, financial modelling, and risk assessment, and supported faculty in finance and economics courses through quiz design, grading, and student mentoring.

Research Intern, CITI Lab, NTUST (Taiwan Tech), Taipei, Taiwan

Sep 2023 – Dec 2023

- Took over a graduating master's student's Outcome Economy project under Professor Shuo-Yan Chou, building predictive maintenance and financial models for subscription-based smart home ownership.
- Modelled opportunity cost and inventory tradeoffs for the lab's RMFS warehouse robotics project using the Newsboy framework.
- Co-developed demand forecasting models using the Northern Goshawk Optimization approach, with experiments and visualizations rolled into the lab's downstream publications and industry partner reports.

PUBLICATIONS

Peer-Reviewed

- Burak Satar, Zhixin Ma, Patrick Amadeus Irawan, Wilfried Ariel Mulyawan, et al. "Seeing Culture: A Benchmark for Visual Reasoning and Grounding." EMNLP 2025 Main Conference. First benchmark stress-testing Vision-Language Models on visually grounded reasoning across nine Southeast Asian cultures.

Working Paper Under Review, AAI - 2027

- **"Where Liability Lands When Automated Trading Systems Cause Harm: An Auditable, Multi-Coder Map of the Enforcement Record."** Codes 21 publicly documented liability outcomes involving automated trading in the US, UK and China from 2010 to 2025, with a four-model coder panel reporting inter-coder reliability. Finds that no court has adjudicated a system's own decision as the legal wrong, and that liability migrates from individual fault to firm-level ex ante control duties as automation deepens.
- **"Auditing Entity Redaction in Retrospective LLM Scoring: A Randomized Token Ablation and a Provenance Anomaly."** Shows that the standard contamination control of redacting company names moves LLM scores by up to 10 points, while a preregistered randomized minimal pair isolating the name token alone moves them by under one point. Also documents a provenance anomaly in which a serving alias routed paired experimental arms to different model versions.
- **"Raw Error-Correlation Levels Do Not Identify Monoculture in Binary Forecasts: Null-Specification Sensitivity in LLM Forecast Ensembles."** Shows that on binary outcomes every forecaster's signed error contains the realized outcome, so high pairwise error correlation cannot identify shared mechanism. Two defensible accuracy-matched nulls return opposite significant verdicts on the same 30 forecast conditions across six models and three vendors.
- **"When Robustness Checks Aren't Enough: Capacity, Signal, and Implementation Confounds in Adversarial Evaluation."** States two diagnostic lemmas showing that ratio-based robustness metrics can be won by a capacity-zero predictor and that an adversary becomes vacuous as task signal vanishes. An asset-pricing case study documents a Bonferroni-significant 38% harm that passes every within-pipeline check yet fails to transport to a matched pipeline.

Previous Technical Reports

- New Approach to Smart Home Subscription Systems in the Outcome Economy Era. NTUST CITI Lab.
- RMFS Opportunity Cost on Inventory Management. NTUST CITI Lab.

INDUSTRY EXPERIENCE

Arbital (Y Combinator 2026), Web3 / On-chain Quant Trading, Remote

Dec 2025 – Aug 2026

Head of Quantitative (May – Aug 2026), Quantitative Analyst (Feb – May 2026), Portfolio Manager (Dec 2025 – Feb 2026)

- Own the full math, model and strategy framework for perpetual DEX market making: fair-value pricing, funding-rate arbitrage, Avellaneda-Stoikov market making adapted for on-chain execution, and order-book imbalance signals.
- Lead strategy and operations for a systematic portfolio across on-chain prediction markets, owning portfolio construction, position sizing, and risk frameworks across multiple concurrent strategies.
- Develop quant strategies grounded in Avellaneda-Stoikov market making, Kelly criterion sizing, Black-Scholes pricing, Markov chains, Monte Carlo simulation, and Sharpe and Sortino-based risk evaluation.
- Designed the core trading infrastructure end to end: real-time on-chain data pipelines, multi-agent execution layers, dashboards, and configuration and risk controls, with a focus on clean execution and slippage minimisation.
- Built strategies with low pairwise correlation across distinct sectors of the on-chain prediction market, calibrating win rate and risk-reward ratio so each strategy stayed proportional to the portfolio's risk budget.
- Lead a cross-institutional research team spanning NYU, NTU, and HKU across the full alpha-to-deployment loop, from hypothesis generation to live capital deployment.

Assistant Analyst, Indonesia Ministry of Public Works and Housing | Part-time

Mar – Jun 2024

- Audited Indonesian regulations affecting the ministry, mapping how legislative changes impacted policy, decision-making, and operating procedures across the strategic planning division.
- Analysed the ARO matrix organisational structure, surfacing workflow bottlenecks and recommending realignments that clarified scope and ownership across teams.
- Drafted revised business models from existing legal and regulatory texts, integrating those insights into new service delivery and stakeholder engagement plans.

Digital Operations Manager, Radian Jaya, Indonesia | Hybrid

Sep 2020 – Jun 2025

- Led daily operations across inventory, fulfilment, customer experience, and creative content, and authored a lean returns and operations playbook that reduced friction on customer issues.
- Built a data-driven pricing and barbell strategy spanning budget and premium tiers, paired with targeted ads, SEO, and paid social campaigns that grew annual sales by roughly 20%.

LEADERSHIP & ENTREPRENEURSHIP

Co-Founder and COO, Ollifant Bebe (Babycare Brand), Indonesia

Oct 2024 – Present

- Co-founded a baby and kids care brand, running consumer research with parents on hair care needs, preferences, and pain points, and driving organic reach and brand visibility from launch.
- Run operations across product development, supplier coordination, branding, and go-to-market, taking new products from concept to shelf and selling 150+ units in the first month.

Founder and CEO, Miarsa (Student Housing Review Platform), Indonesia

2021 – 2022

- Founded a community-driven housing platform covering what listing sites left out: real student experiences of the places they live in across multiple Indonesian cities.
- Led on-the-ground research interviewing residents on budget, living conditions, guest policies, safety, and landlord behaviour, turning raw interviews into structured listings.
- Grew the platform to 150+ active student users, validating product-market fit through repeat use and word-of-mouth referrals.

TECHNICAL SKILLS

- Machine Learning: supervised and unsupervised learning, time-series forecasting, statistical modelling, regime detection, feature importance analysis, NLP, model evaluation, benchmark construction.
- AI and Deep Learning: Large Language Models, Vision-Language Models, multimodal reasoning, multi-agent system design and orchestration, RAG pipelines, dataset and annotation pipelines.
- Data Science and Quant: algorithmic trading system development, CLOB execution layers, Web3 broker API integration, on-chain analytics, real-time market data pipelines, financial modelling.
- Systems: interactive dashboard design, self-hosted server deployment, end-to-end ML deployment, automated data pipelines.
- Programming and Tools: Python, PyTorch, scikit-learn, Pandas, NumPy, SQL, React, TypeScript, Vercel.

HONORS & AWARDS

- IISMA Scholarship: 1 of 14 awardees selected from over 12,000 national applicants for the Indonesian Government's flagship student exchange program at NTUST, Taiwan (2023).
- PMW (Program Mahasiswa Wirausaha) Entrepreneurship Award and national startup funding, for founding Ticketoad, a secondhand event ticket marketplace tackling scalping and unsafe peer-to-peer resale (2022).
- 2nd Place Regional Physics (2018) and 3rd Place Regional Astronomy (2017), National Science Olympiad, DKI Jakarta.

LANGUAGES AND STANDARDIZED TESTS

Indonesian: Native | English: IELTS Band 7 | Mandarin: HSK 3 | GRE: Quantitative 161