Jubal Chandy Jacob

+44 7552 919998 | jubaljacob67@gmail.com | Linkedin | Website |GitHub

AI researcher and software engineer with an MSc in Artificial Intelligence with Distinction from the University of Aberdeen. Research experience in reinforcement learning and goal recognition, experience developing and evaluating LLM-based and autonomous AI systems. Research interests include technical AI safety, behavioural monitoring of autonomous agents, goal misgeneralisation, AI evaluation and human oversight. Experienced in experimental design, quantitative model evaluation and production AI engineering.

Research Interests: Technical AI Safety · AI Control & Agent Monitoring · Goal Recognition · Reinforcement Learning · AI Evaluation · Human-in-the-Loop Oversight

EDUCATION

M.Sc Artificial Intelligence, University Of Aberdeen, Aberdeen, United Kingdom.
Grade/Classification: ***Distinction***.
Courses : *Software Agents and Multiagent Systems, Data Mining and Deep Learning, Natural Language Generation, Machine Learning, Symbolic AI, Evaluation of AI Systems, Applied Artificial Intelligence.*
Dissertation : **Tentative Landmarks for Goal Recognition in Reinforcement Learning.** Supervisor: Dr Leonardo Amado.

B.Tech Computer Science Engineering, Karunya University, Coimbatore, India.
Grade/Classification: ***First-Class.***
Dissertation : *EEG Brain-Computer Interface for Enhanced Interaction with Digital Devices***.** Supervisor Dr Esther Daniel

PUBLICATIONS

- ***Amado, L., Jacob, J. C., and Meneguzzi, F.*** *"Tentative Landmarks for Reinforcement Learning." Manuscript submitted to the* ***AAAI Conference on Artificial Intelligence 2027 (AAAI-27)****, 2026. Under review.*
- MingZe Tang, **Jubal Chandy Jacob** (2025). *"Language as a Label: Zero-Shot Multimodal Classification of Everyday Postures under Data Scarcity"*. Paper submitted to the **AACL-IJCNLP 2025 First International Workshop on Multimodal Models for Low-Resource Contexts and Social Impact** (MMLoSo 2025). https://doi.org/10.48550/arXiv.2510.13364.
- **Jubal Chandy Jacob**, Esther Daniel, S. Durga, and S. Seetha. *"An EEG-Based Brain-Computer Interface Approach for Enhanced Interaction with Digital Devices."* In: **Advances in Health Informatics, Intelligent Systems, and Networking Technologies.** (Lecture Notes in Networks and Systems, vol. 1286). Springer Nature Singapore, 2025. https://doi.org/10.1007/978-981-96-4008-9_40.

SELECTED AI RESEARCH & PROJECTS

Tentative Landmarks for Goal Recognition in Reinforcement Learning

- Co-formulated and developed a new goal recognition concept named Tentative Landmarks replacing classical landmark necessity over a full planning model with behaviour-relative necessity inferred from successful traces.
- Mathematically formalised tentative landmarks across predicates, states, actions, and state-action pairs, including the trace-induced reachability graph and cost-envelope formulation and developed four recognition methods :- Tent, hTpen, DTL, and DTL+H
- Comparing them across eight benchmark domains including IPC domains such as Driverlog, Logistics and Rovers against state-of-the-art model-free and neural goal-recognition baselines.
- Highlighted the strongest result: DTL+H reached 0.829 versus 0.687 for the strongest external baseline, an improvement of 0.142. Co-authored and submitted the resulting AAAI-27 manuscript, currently under review.

INSPECT-AI: Domain-Aware Document Review Agent (Ongoing)

- Developed a human-in-the-loop AI Meta science evaluation system in cooperation with Inspect-SR project that generates structured judgements from complex technical documents while allowing expert reviewers to verify, override and annotate every automated decision.

- Designed workflows supporting AI auditability, transparency and decision traceability, preserving reviewer interventions and machine-generated assessments throughout the evaluation process.
- Developed and quantitatively evaluated a trial data duplication component against published ground-truth dataset using reproducible performance analysis.
- Engineered the production prototype using FastAPI, Redis/arq, PostgreSQL/pgvector, RAG, vision-language models, GraphDB and Docker Compose.

EAGE Annual Hackathon 2026 [Dell Technologies & NVIDIA] (June 2026) – 1st Place Winner

- Built an autonomous AI agent for a live energy-grid optimisation simulation, balancing multi-source asset siting, operational constraints and real-time grid stability; awarded 1st place and the top hardware prize.

CORE SKILLS

- **AI Safety & Evaluation:** Explainable AI (XAI), human-in-the-loop AI evaluation, AI auditability, decision traceability, provenance and decision lineage, source attribution, verifiable AI outputs, model transparency, fairness and bias evaluation.
- **Agents, RL & Machine Learning:** Machine Learning, Deep Learning, Transformers, Natural Language Processing (NLP), Large Language Models (LLMs), Retrieval-Augmented Generation (RAG), Computer Vision, Reinforcement Learning, Knowledge Representation, Symbolic AI, Multi-Agent Systems.
- **Research Methods:** Experimental design, quantitative evaluation, ground-truth validation, model performance analysis, literature review, research prototyping, technical documentation and scientific communication.
- **Technical Engineering:** Python, Java, TypeScript, SQL, FastAPI, REST APIs, Docker, CI/CD, PostgreSQL/pgvector, Redis, FAISS, GCP, Microsoft Azure, Git, Agile/Scrum.

WORK EXPERIENCE

Software Engineer, Aberdeen Software Factory (University of Aberdeen) ***May 2025 – August 2026***

- Delivered AI-enabled client systems from requirements and research prototyping through implementation, evaluation and deployment.
- Designed and implemented LLM, RAG and NLP workflows, translating emerging AI techniques into practical client-facing software.
- Collaborated with clients, developers and non-technical stakeholders to translate requirements into deployable AI/software solutions and communicate technical decisions throughout project delivery.

Web Developer, Diversity Bridge SCIO , Aberdeen November 2025 - January 2026

- Developed and maintained web functionality for a Scottish charity, working directly with organisational stakeholders to translate operational requirements into accessible digital features.

Hack The Burgh XII University of Edinburgh [Optiver]*(Nov 2025)* Top finalist (4 of 30 teams).