ALEJANDRO DE LA CRUZ

Chiclayo, Lambayeque, Peru - +51 930 116 108 - alejantony.22@hotmail.com

LinkedIn - GitHub

PROFESSIONAL SUMMARY

Research-oriented Systems Engineer with 2+ years building scalable ML infrastructure and simulation environments for agentic AI systems. Specializing in the designing training environments inspired by world models and JEPA (Joint Embedding Predictive Architectures), with strong interest in neurosymbolic AI approaches for robust reasoning. Currently expanding expertise in reinforcement learning environment design and LLM fine-tuning through hands-on implementation. Proven track record implementing ML pipelines and building rigorous evaluations measuring genuine model capabilities beyond standard benchmarks. Demonstrated high agency by leading NASA-nominated AI project from concept to global recognition. Deep technical experience across distributed systems, containerization (Docker, Kubernetes), and production ML operations, balancing research exploration with engineering rigor.

TECHNICAL SKILLS

AI & Machine Learning Research: World Models and representation learning, JEPA (Joint Embedding Predictive Architectures), Neurosymbolic AI systems, Reinforcement Learning environment design, Long-horizon agentic task planning, Model capability evaluation frameworks, LLM fine-tuning & alignment

ML Frameworks & Tools: PyTorch, TensorFlow, Hugging Face ecosystem, Generative models (diffusion, GANs), Custom training pipeline implementation

Infrastructure & DevOps: Docker, Kubernetes (AKS, OpenShift 4), CI/CD pipelines, Nginx, Ubuntu Server administration, Distributed systems, VM orchestration

Observability & Monitoring: Dynatrace, Grafana, Kibana, Prometheus, Production incident management, Performance optimization

Programming Languages: Python (3.10+, production-grade), Rust (systems programming), JavaScript/TypeScript (Three.js, React, Angular), Java, PHP

Databases & Storage: PostgreSQL, MySQL, Oracle PL/SQL, SQL Server, Database optimization for ML workloads

RELEVANT EXPERIENCE

Lead Research Engineer & Infrastructure Lead | MAWJAC AI (NASA Space Apps Challenge - Global Nominee)

October 2025

- Designed and implemented novel training environment for ML models to classify exoplanets using NASA astronomical datasets (3,000+ confirmed exoplanets), translating complex astrophysical parameters into interactive 3D simulation spaces
- Built rigorous evaluation framework measuring genuine model capability in ambiguous scenarios, with a system that visualizes prediction uncertainty and handles edge cases where training data conflicts, going beyond standard accuracy metrics
- Architected procedural world generation engine where physical variables (orbital period, stellar radius, equilibrium temperature) deterministically define environment geometry, an approach inspired by world model architectures where latent representations encode environmental dynamics for agent reasoning
- Owned end-to-end infrastructure: containerized ML pipeline with Docker, implemented real-time data processing for astronomical datasets, achieved sub-200ms inference latency despite complex 3D rendering requirements

- Demonstrated high agency by independently identifying research direction, forming cross-functional team, and shipping production system that achieved global recognition, selected as NASA Space Apps Global Nominee from 60,000+ participants

Systems Engineer – RPA Operations & Infrastructure | NTT DATA

July 2024 – Present | Remote

- Manage production infrastructure for international banking clients (Banco Pichincha Ecuador, Banco de Crédito del Perú), reducing incident response time by 40% through automated monitoring and comprehensive service documentation
- Operate distributed systems across OpenShift 4 and Azure Kubernetes Service, debugging complex microservice interactions in real-time using Dynatrace, Grafana, Kibana, developing deep expertise in production system reliability
- Built automated incident detection pipelines querying Oracle PL/SQL and SQL Server databases, and optimized log analysis queries reducing MTTR from hours to minutes for critical production issues
- Developed and implemented RPA automations for project management and traceability across Global Delivery Network Spain (GDN Spain) using Power Platform, streamlining cross-team workflows for international operations
- Architected RPA system for rapid, prioritized notification of critical production incidents for Alicorp (Peru's largest consumer goods supply chain), automating escalation workflows reducing response latency for P1 incidents
- Lead war room coordination and technical training sessions for global teams, demonstrating ability to operate under high-pressure scenarios requiring rapid iteration and cross-team collaboration

Independent ML Engineer & Technical Consultant | Freelance

October 2023 – Present | Remote

- Delivered 6 end-to-end AI/ML and software solutions for businesses across e-commerce, healthcare, and finance, owning technical architecture decisions, evaluated trade-offs between research exploration and production constraints
- Implemented LLM-powered automation systems and generative AI applications, and designed evaluation frameworks measuring real-world task performance beyond benchmark metrics, ensuring production reliability
- Built scalable ML infrastructure using Docker Compose, React/Vue.js, Django, PostgreSQL, Nginx on Ubuntu Server, and established CI/CD pipelines enabling rapid iteration from research prototypes to production deployments

Backend Software Engineer | Grupo Consigueventas Inversiones

May 2023 – August 2023 | Remote

- Developed attendance management system using Test-Driven Development (TDD) methodology, implementing comprehensive unit tests and refactored codebase achieving 85%+ code coverage, demonstrating commitment to code quality and maintainability
- Architected RESTful API services and database schema for Laravel-based system, balancing technical rigor with rapid iteration in agile environment, shipping features weekly while maintaining clean architecture
- Mentored and onboarded junior backend developers, and established code review practices and pair programming sessions, fostering collaborative problem-solving culture aligned with technical excellence

Full-Stack Software Engineer | Universidad Nacional Pedro Ruiz Gallo

January 2024 – April 2024 | Hybrid

- Developed legal document management system for university's Legal Advisory Office using Laravel and Vue.js, improving administrative efficiency by 60% through workflow automation and real-time tracking
- Architected RESTful API services and database schema, balancing engineering rigor (test-driven development, comprehensive documentation) with rapid feature delivery in agile environment

Volunteer Software Engineer | Universidad Nacional Pedro Ruiz Gallo - Nursing Faculty

May 2023 – October 2024 | Volunteer

- Owned end-to-end development of thesis project management system for graduate specialty programs, demonstrating high agency by independently driving project from requirements gathering through production deployment over 17-month period
- Designed complete system architecture including database schema, application logic, and user interfaces, and created comprehensive documentation and user manuals, showcasing ability to ship production-ready systems with minimal supervision

CURRENT RESEARCH & ACTIVE LEARNING

Actively developing expertise in areas critical to advancing agentic AI systems through hands-on implementation and self-directed study:

- **Reinforcement Learning Environment Design:** Currently implementing custom Gymnasium-compatible environments for long-horizon tasks, studying reward shaping methodologies, and experimenting with PPO/SAC algorithms for policy optimization. Building intuition through hands-on coding rather than purely theoretical study.
- **LLM Fine-tuning & Evaluation:** Actively learning parameter-efficient fine-tuning techniques (LoRA, QLoRA) and implementing custom evaluation frameworks. Studying RLHF pipelines and alignment research to understand how models learn from human feedback at scale.
- **World Models & Representation Learning:** Deep-diving into JEPA architectures and studying how models can build internal representations of environment dynamics. Exploring connections between neurosymbolic approaches and learned world models for robust reasoning.
- **Empirical ML Research Methodology:** Following Anthropic's research output (Constitutional AI, scaling laws papers) to understand how to design rigorous experiments that measure genuine capabilities rather than benchmark performance. Developing research taste through critical reading of recent ML literature.

RESEARCH INTERESTS

Designing training environments for long-horizon agentic tasks, evaluation methodologies that measure real-world capabilities, neurosymbolic AI architectures combining learned representations with structured reasoning, and empirical approaches to AI alignment and safety.

EDUCATION

Bachelor of Science in Systems Engineering

Universidad Nacional Pedro Ruiz Gallo, Lambayeque, Peru | October 2019 – September 2024

Relevant Coursework: Machine Learning, Distributed Systems, Computer Networks, Algorithms & Data Structures, Database Systems

CERTIFICATIONS & CONTINUOUS LEARNING

- Oracle Cloud Infrastructure 2023 AI Certified Foundations Associate
- Oracle Cloud Infrastructure 2023 Certified Foundations Associate
- Project Management Professional (PMP) PMBOK 7.0 Training
- Agile Project Management with Scrum Certification
- Generative AI: LLMs Project Management & Practice (Udemy, 2026)

ADDITIONAL INFORMATION

- **Languages:** Spanish (native), English (professional working proficiency)

- **Collaboration Style:** Active pair programmer with experience in remote-first team environments, and frequent contributor to technical discussions and cross-functional problem-solving sessions
- **Work Authorization:** Open to visa sponsorship for US positions.