Bryant Mejia
Software Engineer | New York, NY
Bryantmejia722@outlook.com | Phone: 15707901185
Portfolio | Github | Linkedin

Professional Profile
Senior Software Engineer with 6+ years of experience building production platforms, AI-powered applications, and configuring enterprise-grade cloud infrastructure. Designed model evaluation workflows, retrieval systems, and tools for LLM applications. Expert ability to translate business requirements into scalable solutions across cloud infrastructure, data systems, and application platforms.

Professional Experience
AI Platforms Engineer, Independent Contributor, San Francisco, CA
Dec 2024 – Apr 2026
- **Engineered an autonomous deep research agent** for large-scale document analysis, incorporating deterministic retrieval pipelines, traceability, and human review checkpoints to improve output reliability.
- **Designed deterministic agent workflows** for retrieval, research, reasoning, and document synthesis across 100k+ token document corpora, improving reproducibility and inspection of multi-stage AI workflows.
- **Formal MLOps education**, earned the AWS Machine Learning Engineer Associate certification (July 2025).

Senior Software Engineer, Platform & AI Integration, Gras, Inc, New York, NY
Feb 2023 – Dec 2024
- **Architected secure full-stack platforms** handling sensitive business data, emphasizing reliability, auditability, and operational visibility.
- **Built event-driven data pipelines** processing 50k+ events/day, generating operational telemetry used for monitoring application behavior and critical business workflows.

Senior Software Engineer, Cloud & Data Platforms, Capgemini USA, Remote
Mar 2021 – Feb 2023
- **Led a global team of 12** offshore and domestic engineers, establishing engineering standards and driving the adoption of reuse-first architectural patterns.
- **Successfully migrated 30+ customer-facing applications**, introducing deployment safeguards, monitoring standards, and operational controls for cloud-native environments.
- **Established observability tools** using Prometheus and ELK, defining service-level objectives and operational metrics that reduced incidents by 30% and improved system health visibility.

Featured Projects
Multimodal Search System View Project
May 2026
OpenSearch, OpenCV, PySceneDetect, Gemini Embedding 2, Python
- **Multi-vector-based indexing and retrieval pipelines** supporting semantic search and ranking across video, image, and text embeddings of heterogeneous content sources.

Report Agent – Full-Stack Deep Research Agent View Project
Feb 2025 – Jan 2026
RAG, Langgraph, LangSmith, Google Gemini, Claude 3.5 Sonnet, Python

- **Core Logic & Performance:** Engineered a plan-and-execute multi-stage LLM reasoning pipeline using Python, LangGraph, and RAG. Synthesized hundreds of documents into structured markdown reports (10–20 sections, 50+ citations) with optimized frontend rendering pipelines.
- **Frontend & UI Architecture:** Designed and delivered a highly interactive, responsive web interface featuring a human-in-the-loop review system that enables users to approve, refine, or restructure multi-stage AI research plans in real time.
- **Agent Observability:** Integrated LangSmith tracing and evaluation tooling to monitor parallel tool execution, latency, token usage, and workflow behavior across complex multi-stage research pipelines.

Technical Skills

Frontend Development:
React.js, HTML5, CSS3, Next.js, Responsive Design, Cross-browser Compatibility, State Management, Web Performance, Accessibility (WCAG), CSS-in-JS, Tailwind CSS, Webpack, Vite, Babel, Component Libraries, Progressive Web Apps (PWAs)

Backend Development:
Python (Django, Flask, FastAPI), TypeScript, JavaScript (ES3 - ES16), Node.js, REST, Microservices, gRPC, WebSockets, GraphQL

LLM Integration:
LangChain, LangGraph, LlamaIndex, Prompt Engineering, Model Monitoring, Function Calling, Human-in-the-Loop Review, Response Evaluations, Retrieval Systems, Grounded Generation, Fault Tolerance

Cloud & DevOps:
GCP (Cloud Run, BigQuery, Compute Engine), AWS (Lambda, ECS, S3, CloudWatch), Firebase, Docker, Kubernetes, Terraform, CI/CD (GitHub Actions, Jenkins, Gitlab), Agile, Git

Data & Databases:
PostgreSQL, Redis, DynamoDB, Kafka, RabbitMQ

Observability:
Prometheus, Datadog, Grafana, OpenTelemetry, LangSmith, Logging, Metrics, Alerting

Management:
Cross-functional Stakeholder Collaboration, Global Team Management (12+ engineers), Engineering Standards & Governance, Incident Response

Modernization:
Monolith-to-Microservices Migration (30+ systems), Service Decomposition, Domain-Driven Design (DDD), Strangler Pattern, Service Catalog Design, Event-Driven Architecture (EDA), API-First Design (REST, GraphQL), Multi-region Scalability

Education

Certified Machine Learning Engineer Associate, AWS, July 2025
Generative AI Fundamentals,Databricks, 2024
Designing Big Data Analytics Solutions, Great Learning, 2024
Associate Degree – Computer Science, Business Applications Development, Harrisburg Area Community College, 2020
C Programming & Computer Science Boot Camp, 42 Silicon Valley, 2018