

Christian Vega

Data & AI Technical Lead — Production ML/LLM · RAG · Delivery in regulated healthcare & payer data

christian.vega.data@gmail.com · 939-253-1147 · [linkedin.com/in/christian-vega-183a2b263](https://www.linkedin.com/in/christian-vega-183a2b263) · vizlogic.tech

Puerto Rico (U.S.) — authorized to work in the U.S., remote-ready · Bilingual EN/ES

SUMMARY

Six years in regulated healthcare, payer, and government data (Optum, Humana, PRASA), delivered under HIPAA and Medicare C&D. Leads the AI/LLM automation workstream at Optum: owns scope, technical direction, and delivery, and develops the people around the work, staying technical enough to architect the systems and earn engineers' trust. Has driven three data & AI products from raw data to live production — RAG over ~1.24M government contracts, an ML pricing model with 80% prediction intervals validated at 79.6% empirical coverage, and a business platform on AWS.

Regulated-data posture: source code and client/employer data are kept private by design (HIPAA & client confidentiality). Live products to inspect — nidopr.app · licitapr.com · vizlogic.tech. Architecture & evaluation walkthroughs available on request.

PROFESSIONAL EXPERIENCE

Senior Data Analyst — Optum (UnitedHealth Group)

Mar 2025 – Present

Python · SQL · Snowflake · Azure Data Factory · LLM integration

- Built and shipped an in-house LLM tool into HIPAA-bound production at a UnitedHealth company — automating first-draft Business Requirements Documents and data-dictionary / column-definition context (~47 documents generated since launch), work that previously took analysts hours per document and freed business-analyst capacity for higher-value work.
- Leads the AI/LLM automation workstream end-to-end — owns scope, technical direction, and delivery; sets architecture and LLM/prompt standards; coordinates cross-functional stakeholders from design through production.
- Designed and taught a data-analytics enablement curriculum, upskilling non-technical supervisors and managers to work with data directly.
- Directs the migration of legacy systems onto a modern, automated, ML/LLM-driven data platform.

Founder & Principal — Vizlogic — vizlogic.tech (concurrent)

Jan 2024 – Present

Python · FastAPI · PostgreSQL/pgvector (HNSW) · LightGBM · Nuxt 3 · TypeScript · Docker · AWS (Lightsail / ECS Fargate) · Cloudflare

- Founder and principal of an independent, self-funded data & AI practice — owns strategy, architecture, model evaluation, and delivery for three live production products.
- Built a production RAG system over ~1.24M government contracts across six source systems (~90% corpus coverage) — Splink entity resolution de-duplicating 422K !' ~345K records, hybrid pgvector/HNSW retrieval, gated by a 62-case evaluation suite at 100% pass. ([LicitaPR](https://licitapr.com), live at licitapr.com)
- Owns a production ML pricing system end-to-end, feature pipeline to monitored inference — 80% prediction intervals validated at 79.6% empirical coverage and model bias corrected from +44% to +5%. ([Nidopr](https://nidopr.app), live at nidopr.app — ~64,000 listings indexed nightly from 10+ sources)
- Operates for reliability: 34 scheduled nightly jobs — including an LLM enrichment pass that parses and normalizes raw listing descriptions into clean, structured records — under a 40-check monitoring layer (phone/Discord/email alerting) and automated QA, plus an autonomous data-selected publisher with OCR photo-safety screening.
- FT-OS — a complete small-business operating system (visual builder, POS, kitchen display, inventory, admin) live on AWS, shipped behind a 148-check QA gate.

Data Analyst — TrueNorth Corporation — client: PRASA (PR Aqueduct & Sewer Authority)

Jan 2023 – Feb 2025

MSSQL Server · SSIS · ETL · Python · Power BI · SSRS · REST APIs

- Led the technical delivery of PRASA's water-balance automation initiative — set direction for a delivery team (an analyst and two developers, with a project manager) and coordinated ~20 stakeholders across multiple departments.
- Replaced a recurring ~2-month, multi-person manual compilation — pulled by a business analyst from multiple data silos with developer and cross-department support — with a single automated ETL pipeline (SSIS + Python).
- Designed and shipped end-to-end data pipelines and BI dashboards for PRASA, moving operational reporting from manual handling to automated, near-real-time delivery.
- Engineered REST API integrations to pull and process external sources into the warehouse, improving dashboard data coverage and accuracy.
- Automated reporting (MSSQL + Python + SSRS), cutting report-generation time ~30%.
- Led data-governance initiatives across PRASA analytics — integrity, security, and compliance.

Quality Analyst — Next Level Solutions

Jun 2021 – Dec 2022

SQL · Postman · Snowflake · SSIS · SSMS · Azure DevOps

- Built automated test frameworks (SQL + API validation via Postman) that cut manual testing time ~30%.
- Ran Snowflake data-model query testing and validation (SSIS) to certify accurate reporting structures.
- Planned deployments and test cases with cross-functional teams (~20% faster releases) and streamlined API testing to reduce unresolved bugs ~15%.

Compliance Leader — Humana (Medicare Advantage)

Apr 2019 – May 2021

SQL · Power Automate · SSIS · IBM Cognos · Tableau

- Led compliance initiatives — automated audit workflows and data validation (SQL, Power Automate) — reaching a 95% compliance rate.
- Conducted data-integrity audits (SSIS, IBM Cognos) that reduced regulatory reporting errors ~10%, and drove ~25% faster claims/inquiry resolution via automated Tableau dashboards.
- Enforced HIPAA and Medicare Parts C & D adherence through data-governance policies and encryption.

SKILLS

Leadership & Delivery: Data & AI strategy, Project & delivery leadership, Cross-functional stakeholder management, Technical direction & architecture review, Roadmaps & delivery ownership, Data governance & compliance programs, Team enablement & mentoring

Machine Learning & AI: Machine Learning (LightGBM / gradient boosting), Conformal prediction / calibrated prediction intervals, Model evaluation & back-testing (Diebold-Mariano), Retrieval-Augmented Generation (RAG), Vector search (pgvector / HNSW), Embeddings & hybrid retrieval, LLM integration (Azure OpenAI, Snowflake Cortex, Ollama), Agentic / multi-agent orchestration, Entity resolution (Splink), OCR (Tesseract / EasyOCR), Model deployment & monitoring (MLOps / LLMOps)

Data Engineering & Pipelines: Python, SQL (advanced), ETL / ELT, Data pipelines (production-grade), Azure Data Factory, SSIS, dbt, REST API integration, Resumable / idempotent crawlers, Web scraping (Playwright / Patchright), Data modeling, Data quality

Data Platforms & Warehousing: Snowflake, Data warehousing, PostgreSQL (pgvector, FTS), MSSQL Server, MySQL, Google BigQuery, AWS Data Lake, SQLite

Cloud, MLOps & Production Ops: AWS (Lightsail; ECS Fargate + EFS), Azure (Data Factory, DevOps), Docker, CI/CD (git push-to-deploy), cron / systemd scheduling, Custom Python health checks & alerting, Cloudflare (Tunnel, Pages, DNS), Production monitoring / observability, Incident root-cause analysis (RCA), Restore-verified backups (restic)

BI, Governance & Compliance: Power BI, Tableau, Looker, SSRS, ECharts, HIPAA, Medicare Parts C & D, Data governance, Metadata management, Data security, Automated testing & validation

Languages & Domain: Python, SQL, TypeScript / JavaScript (Nuxt 3 / Vue), PowerShell, Bash, Healthcare / payer data (Optum, Humana), Regulated / high-stakes data, Bilingual EN/ES

CERTIFICATIONS

- Microsoft Certified: Azure AI Engineer Associate (AI-102) — In progress
- IBM Data Science Specialization
- Python Developer — Zero to Mastery

EDUCATION

- Associate's Degree in Biotechnology — 2016 (a hypothesis-driven scientific foundation carried into data and ML work)