

GILBERT IGWE

Senior Data and AI Leader • Applied ML, NLP and GenAI • Real Estate and Financial Services

Atlanta, GA • consult@gilbertigwe.com • 404-908-9506 • [linkedin.com/in/gilbertigwe](https://www.linkedin.com/in/gilbertigwe)

SUMMARY

Senior data and AI leader with 11+ years translating ambiguous business problems into production ML, NLP, time series, and decision intelligence systems for Fortune 100 enterprises and government. Currently Principal Scientist at the State of Connecticut, leading enterprise AI, data, and analytics initiatives across 100+ agencies serving 3.7 million citizens. Co-founded Solution Cabin LLC in 2015, a specialized AI/ML practice that delivered embedded transformation engagements at United Airlines, Cisco Systems, Spotify, UHG Optum, Verizon, and Wellstar Health System. Active real estate investor and operator in the Atlanta metro. Hands-on builder who codes, models, and ships, with a track record of moving executive-level decisions through data.

CORE SKILLS

Strategy and Leadership: Data and AI strategy, ML roadmap and governance, executive stakeholder communication, transformational consulting, technical mentorship across distributed offices, vendor and partner management

ML and Statistics: Supervised and unsupervised learning, time series forecasting (Prophet, ARIMA, gradient-boosted), NLP and text mining, anomaly detection, Bayesian inference, causal inference and quasi-experimental design, A/B and switchback testing, optimization (Gurobi, OR-Tools)

GenAI and LLM: RAG (hybrid retrieval, cross-encoder re-ranking, citation grounding), LLM evaluation (faithfulness, hallucination, answer relevance via RAGAS and TruLens), agent orchestration, fine-tuning, OpenAI / Anthropic / Bedrock, vector stores (Pinecone, Weaviate, FAISS, pgvector, Qdrant)

Engineering and Languages: Python, SQL, TypeScript, Bash; FastAPI, Docker, Kubernetes (EKS, AKS, GKE), MLflow, Weights and Biases, Airflow, Prefect, Celery, Ray Serve; CI / CD with GitHub Actions; REST and gRPC API design

Data Platforms: Spark, PySpark, Databricks, dbt, Snowflake, BigQuery, Delta Lake, Kafka, Postgres, Redis

Cloud: AWS (SageMaker, Bedrock, Lambda, ECS, EKS), GCP (Vertex AI, BigQuery), Azure (ML, OpenAI, AKS), Oracle Cloud (OCI)

Visualization, BI and Geospatial: Tableau (including geographic and map layers), Power BI, advanced Excel (financial modeling, what-if analysis, scenario planning), Plotly, PostGIS, raster and satellite signal integration

Industry Coverage: Real estate (active investor and operator), financial services, telecom, transportation, healthcare, government services, technology

PROFESSIONAL EXPERIENCE

State of Connecticut, Department of Administrative Services | Hartford, CT (Remote) 2023 to Present

Principal Scientist, Analytics, AI and Data

- Lead enterprise AI, data, and analytics initiatives across 100+ state agencies, delivering measurable impact for 3.7 million citizens. Define the data strategy, ML roadmap, and governance posture that informs senior leadership and drives agency-level operating decisions.
- Architect the analytics foundation for the statewide Web Modernization Program: event taxonomy, KPI framework, cohort definitions, and quasi-experimental evaluation patterns adopted as the standard of practice across all participating agencies.
- Led ALPHA CT, a forensic analysis of the Connecticut Health eligibility funnel covering 3,075 sessions. Surfaced a 29.2% abandonment rate concentrated in the first six steps, isolated two distinct user archetypes, and quantified a 4% completer-to-appointment drop. Findings shaped the live redesign roadmap currently in execution.
- Standing up an enterprise LLM platform for document triage, policy lookup, and citizen-facing assistance. Defined the evaluation rubric covering faithfulness, citation accuracy, and refusal correctness, plus the human-in-the-loop review workflow required before any agency-facing rollout.
- Designed the executive measurement layer (Python plus Power BI) that consolidated 17 modernization workstreams into a single dashboard now used in monthly leadership reviews.
- Mentor data scientists and analytics engineers across multiple offices and contractor teams. Drive code review standards, model card practices, and experimental design rigor.

Solution Cabin LLC | Atlanta, GA 2015 to 2023

Managing Partner and Principal AI / ML Engineer

- Founded and led a specialized AI / ML consulting practice. Embedded senior data scientists and applied ML engineers inside Fortune 100 enterprises to deliver production ML, NLP, anomaly detection, fraud detection, content moderation, recommender systems, and enterprise analytics. Built the talent bench, engineering delivery model, and governance framework (code review standards, deployment pipelines, monitoring contracts, and retraining cadences) that allowed

engagements to ship with internal-team rigor. Selected client engagements include United Airlines, Cisco Systems, Spotify, UHG Optum, Verizon, and Wellstar Health System.

Spotify | New York, NY 2022 to 2023

Senior Data Scientist (embedded engagement)

- Partnered with Personalization and Marketplace teams on listener growth experiments. Designed A/B and switchback tests for podcast and audiobook discovery surfaces, including power analysis under heavy-tailed engagement distributions.
- Built feature pipelines in PySpark on Databricks ingesting hundreds of millions of daily events. Productionized cohort scoring jobs orchestrated in Airflow with Delta Lake as the source of truth.
- Prototyped an embedding-based content similarity service for cold-start podcast recommendations. Shipped a FastAPI inference endpoint backed by FAISS and benchmarked latency and recall against the incumbent collaborative filter.

Cisco Systems | San Jose, CA 2021 to 2022

Senior Data Scientist (embedded engagement)

- Embedded with the Network Assurance group. Built unsupervised anomaly detection on streaming router telemetry using isolation forests and seasonal-hybrid ESD over Kafka topics.
- Designed a Bayesian change-point layer over device KPIs that reduced false-positive alert volume by an estimated 40% in pilot deployments.
- Built a capacity-planning forecasting service combining Prophet and LightGBM, deployed to AWS via SageMaker endpoints behind an internal gRPC API. Owned CI / CD through Docker and GitHub Actions.

United Airlines | Chicago, IL 2018 to 2021

Senior Data Scientist (embedded engagement)

- Member of an advanced analytics team supporting fleet operations and MRO. Built predictive maintenance models for line-replaceable units across narrow-body and wide-body fleets, informing parts provisioning and turnaround scheduling at major hubs.
- Designed a digital-twin-style simulation of the parts and MRO supply chain in SimPy and PySpark. Quantified bottleneck sensitivity to lead-time variance across 60+ Tier-1 suppliers and informed buffer policies adopted by the program office.
- Built a Gurobi integer-programming model for crew and gate optimization scenarios under irregular operations, packaged behind a Python service used in disruption tabletop exercises.

SELECTED INDEPENDENT BUILDS (PRIVATE REPOSITORIES, DEMOS AVAILABLE ON REQUEST)

DocSensei | Document intelligence and RAG platform

2023 to Present

- End-to-end RAG system for high-volume document workflows in insurance and legal. Hybrid retrieval (BM25 plus dense embeddings), cross-encoder re-ranking, page-faithful citations, and a hallucination guardrail scoring each generated span against retrieved evidence. Architecture is directly applicable to commercial real estate loan files: rent rolls, operating statements, leases, intercreditor agreements, and trustee reports. Currently in beta at a U.S. insurance brokerage under a perpetual source-code license.

ORCA | Asset-level risk and liability modeling

2024

- Asset-level liability and exposure model for a national operator. Combined satellite-derived signals, regulatory schedules, and operator-submitted data into a single risk index. Quantified expected exposure under alternative regulatory regimes and surfaced under-reporting patterns. Geospatial layer integrates raster signals with operator records using PostGIS and Tableau map layers.

Setup Score | Quantitative equity and options signal research

2022 to Present

- Multi-factor signal generator (Setup Score v6 and v7) blending momentum, volatility regime, options flow, and earnings drift. Pine Script and Python research stack with live tracking across selected positions. Hands-on financial markets and instrument-level research informing personal trading capital.

EDUCATION

Master of Business Administration (Candidate) | Executive MBA Program, Ross School of Business, University of Michigan

Master of Science, Data Science | Georgia State University, Atlanta, GA

Bachelor of Management Studies | University of Mumbai, India

First Class with Distinction; Class Valedictorian.