

GLEICE CHAVES

+1 (747) 220-4099 | gleice.schaves@gmail.com | Los Angeles, CA, USA | gleicechaves.com | linkedin.com/in/gleiceschaves

EDUCATION

University of Southern California <i>Master's, Data Science</i> Database Systems, Machine Learning, Data Mining, Research Methods, Applied Natural Language Processing	August 2024 - December 2026
Faculdade de Informática e Administração Paulista <i>Bachelor's, Business Analytics</i> Data Modeling, Data Warehousing, NoSQL/SQL, BI, Cloud Services, Data Governance, Business Development, Python	January 2019 - December 2022 GPA: 3.7

SKILLS

Skills: Analytics Engineering & BI: SAAS Tabular, DAX, Semantic Modeling, Tableau, Power BI (DAX, DAX Studio, VertiPaq), data visualization Data Engineering & Platforms: ETL/ELT, Hive, PostgreSQL, Databricks, MongoDB, Spark, CI/CD (GitHub Actions/Azure DevOps), PySpark, C#, Azure (Fabric, Databricks), AWS(S3,Athena, Lambda), Hadoop, Oracle, PLSQL DS: LangChain, Statistical Analysis, Prompt Engineering, Tokenization/Embeddings, scikit-learn, PyTorch/TensorFlow, Natural Language Processing (NLP), Causal Inference, R, Python, RAG, Feature Engineering, Linear/ Logistic Regression, Vector database(Chroma, Pinecode), predictive models, statistical methods, NumPy, Pandas	
---	--

PROFESSIONAL EXPERIENCE

Keck Medicine of USC <i>Analytics Engineer</i> - Enhanced data insight accuracy by 35% through the implementation of an end-to-end NLP Q&A layer on a Power BI/SSAS tabular model. Exposes KPI/measure metadata and synonyms to Microsoft Copilot, optimizing response quality for 100-question benchmark. - Standardized KPI definitions and documentation across 300+ tables by owning the enterprise BI/tabular semantic model utilizing Power BI/SSAS, Tabular Editor, and DAX, while authoring hardened SQL Server views to encode business rules. - Improved query performance by diagnosing bottlenecks in the data model using DAX Studio, VertiPaq Analyzer, and SQL Profiler/Query Store to eliminate CROSSJOIN and bi-directional filter hotspots.	Los Angeles, CA <i>March 2025 - Present</i>
Diagnostics of America <i>Data Engineer & Analytics</i> - Reduced billing process time from 20 days to hours by designing and deploying a federated data lake and warehouse using AWS (S3/Athena), Databricks, and BigQuery for 350 hospitals and labs. - Enhanced data reliability and simplified analytics by developing pipelines with Alteryx and SQL, incorporating incremental loads, RI/outlier/duplicate checks, and lineage/audit logs. - Delivered decision-ready reporting for C-suite by creating and delivering Power BI dashboards for service-line and patient-level costing, optimizing DAX/SQL models and queries, and establishing daily refresh SLAs.	Sao Paulo, Brazil <i>June 2021 - December 2022</i>
Link School of Business <i>Business Analyst</i> - Enhanced grant allocation efficiency by 30% through the development of a scoring model, executing market research using Excel with Power Query and Pivot Tables to identify high-potential startups. - Maximized return on ad spend (ROAS) by reallocating marketing budget based on established ROI models, utilizing Excel for monthly forecast and pacing reports across channels including search, social, display, and email.	Sao Paulo, Brazil <i>December 2019 - May 2022</i>

PROJECTS & RESEARCH

GenAI Hackathon - Keck Medicine of USC <u>KeckStraction</u> <i>Hackathon Winner</i> - Developed a signature-governed, stateful LLM pipeline that converts unstructured clinical PDFs into validated, EMR-ready JSON. - Implemented dual-mode orchestration (bootstrap vs constrained) for schema discovery and governed generation, reducing re-runs and token cost - Integrated Docling OCR and JSON Schema validation with real-time UI telemetry using sockets/SSE - Deployed a FastAPI + Postgres backend supporting versioned signatures, observability, and audit-ready storage aligned with HIPAA compliance	Los Angeles, CA, USA <i>October 2025 - Present</i>
ChatDB: Natural-Language Database Query System - USC Built a Streamlit + LangChain text-to-SQL/Mongo app turns English into parameterized PostgreSQL SQL and MongoDB aggregations with schema introspection, LangSmith tracing/evaluation and optional Ollama local LLM	Los Angeles, CA, USA
Human-AI Trust Calibration - PI: Gale Lucas - USC Institute <i>Research Assistant</i> - Co-authored annotation codebooks; performed thematic coding and qualitative analysis; authored methods and documented inter-rater reliability (Cohen's k) to inform calibrated human-agent teaming - Human-Agent Interaction (HAT), AI Trust Calibration, Behavioral Experiment Design, Qualitative Coding (Atlas.ti/NVivo)	Los Angeles, CA, USA <i>2025 - Present</i>