Andre Garon Rico

Financial Data Engineer & FP&A Systems Architect
Charlottesville, VA · ricoandre@hotmail.com · linkedin.com/in/alrico · Portfolio: andrerico.github.io/AndreRico

SUMMARY

I build the systems finance runs on. At Klabin, one of Brazil's largest pulp-and-paper companies, I spent a decade building the SAP data warehouse and then the FP&A area on top of it, cutting the close and reporting cycle from about 4 days to less than 1. Since moving to the US I have built large-scale data platforms and machine-learning pipelines at Penn State and the University of Pennsylvania. Registered accountant (Brazil), Anaplan Certified Model Builder, and hands-on engineer in Python and SQL.

CORE SKILLS

Planning & EPM: Anaplan · SAP BPC · SAP BW-IP · driver-based planning · financial modeling
Finance: FP&A · close & consolidation · budgeting & forecasting · CAPEX · intercompany & managerial costing
Data: SQL · SAP BW 7.4 on HANA · PostgreSQL · data warehousing · ETL (Pentaho) · Hadoop / Hive
Analytics, ML & AI: Python · statistical modeling · ML on HPC/GPU · LLM integration · Tableau · BusinessObjects
Software & cloud: Django · React · TypeScript · Docker · AWS · Linux · Agile

EXPERIENCE

Senior Application Developer — University of Pennsylvania, Perelman School of Medicine · 2024 – present
- Lead developer of Biofilter, a versioned knowledge platform that integrates more than a dozen public databases; its data powers the 13.8M-node knowledge graph behind DRIVE-KG (PSB 2026).
- Statistical genetics and machine-learning research on the Penn Medicine BioBank (260,000+ participants): pipelines over billions of genetic variants and multi-terabyte datasets on HPC and GPU clusters.
- Built MyNHANES, a platform that downloads, normalizes and harmonizes NHANES survey data across cycles for longitudinal analysis.
- PharmCAT team member: tested the pharmacogenomics tool on AnVIL, Terra, All of Us, UK Biobank, DNAnexus and AWS, and built pipeline templates for other researchers.

Senior Application Developer — Penn State University, College of Agricultural Sciences · 2022 – 2024
- Built IGEM, a gene–exposure knowledge base (25 ETL pipelines, PostgreSQL, FastAPI, public website); first-author paper at the Pacific Symposium on Biocomputing 2025.
- Extended and applied CLARITE, a framework for association and interaction analyses in population-scale datasets.
- Built a data integration system for heterogeneous clinical-study datasets with PII protection.

Senior Consultant, FP&A — Klabin S/A, São Paulo · 2018 – 2021
- Led FP&A full time: annual budget, variance analysis, product-mix and margin analysis, long-term projections and reports for executive leadership.
- Automated management calculations and chained results; reduced the time to publish results.
- Delivered an integrated CAPEX planning and depreciation model (2019) and a managerial cost and intercompany planning system for pricing and inventory valuation across entities (2020).

Senior Consultant, FP&A and BI — Klabin S/A, São Paulo · 2016 – 2018
- Built the FP&A area inside the enterprise data warehouse, integrating all financial information.
- Redesigned consolidation and management reporting, cutting the close and reporting cycle from about 4 days to less than 1; redesigned budgeting and planning on SAP BW-IP and BPC.
- Delivered value-chain cost and EBITDA attribution (2017) and an automated management reporting platform (2018); automated management books and earnings releases.

Business Intelligence Senior — Klabin S/A, São Paulo · 2008 – 2016
- Built and ran the SAP BI platform for finance, forestry, pulp and paper, and HR; led the migration from SAP BW 3.5 to BW 7.4 on HANA.
- Built a machine-learning talent-intelligence engine on SAP HCM: candidate suggestions, positions at risk, training recommendations and turnover reduction.
- Owned security, backups, load monitoring and development processes; led BI project teams.

Earlier roles · 2000s – 2008
SAP BW Analyst and infrastructure at Cimpor Cimentos de Portugal · Data Warehouse Administrator at Gillette do Brasil (P&G) · Network and Windows server administration at ACSP.

PROJECTS

Finance & data platforms
- **12true** (personal project, 2026 – present): household finance platform on a double-entry ledger, with bank sync (Plaid), receipt reading, US tax organizers and an LLM assistant. Django, PostgreSQL, React; ~1,700 automated tests, 46 ADRs.

Open books
- **SaaS Handbook** (2026): managing and operating SaaS for small teams; the repository is its own example. Quarto, GitHub Actions. garon-sys.github.io/saas-handbook
- **Bioinformatics Handbook** (2026): bioinformatics for human genetics, from molecular biology to GWAS, rare variants, pharmacogenomics and knowledge graphs. andrerico.github.io/Bioinformatics

Genomic data systems
- **Biofilter** (2025 – 2026): versioned knowledge platform over 17 public sources; frozen Parquet bundles queried with DuckDB, no database server. github.com/RitchieLab/biofilter
- **IGEM** (2022 – present): gene–exposure knowledge base with 25 ETL pipelines, FastAPI, Python client and public website. github.com/HallLab/IGEM
- **MyNHANES** (2024): ingestion, harmonization and rule-based transformation of NHANES survey data across cycles; Django, SQLite. github.com/HallLab/MyNHANES
- **PharmCAT** (2024 – present): cross-platform testing and pipeline templates for the open-source pharmacogenomics tool. pharmcat.clinpgx.org
- **Penn Medicine BioBank research** (2024 – present): rare-variant and ExWAS studies of hearing loss, ML for cardiovascular traits, and breast cancer research on 260,000+ participants.

EDUCATION & CERTIFICATIONS

Bachelor's, Accounting and Business/Management — Mackenzie Presbyterian University (2016) · **Postgraduate specialization in Business Administration (MBA program)** — FAAP (2007) · **Bachelor's, Data Science** — Universidade Cidade de São Paulo (2004)
Anaplan Certified Model Builder (2025) · **Registered accountant**, CRC-SP, Brazil (2016)

PUBLICATIONS

- Prado-Rico JM, Bargues-Carot A, …, **Rico ALG**, Hall MA, …, Huang X. Neuron-derived circulating miRNAs reveal lead (Pb) as a key component of metal mixtures exposure. *medRxiv* (preprint), 2026.
- Rajagopalan A, Nguyen TA, Guare LA, **Rico ALG**, …, Ritchie MD, Hall MA, Romano JD, Setia-Verma S. DRIVE-KG: Enhancing variant-phenotype association discovery in understudied complex diseases using heterogeneous knowledge graphs. *Pacific Symposium on Biocomputing*, 2026 (preprint: *medRxiv*, 2025).
- Bargues-Carot A, Prado-Rico JM, …, **Rico ALG**, Hall MA, …, Huang X. MicroRNA expression in asymptomatic welders: implications for biomarker discovery for environmentally-linked neurodegenerative disorders. *medRxiv* (preprint), 2025.
- **Rico ALG**, Palmiero N, Ritchie MD, Hall MA. Integrated exposomic analysis of lipid phenotypes: leveraging GE.db in environment by environment interaction studies. *Pacific Symposium on Biocomputing*, 2025 (first author).
- Zhou J, **Rico ALG**, Guare L, et al. Flexibly encoded genome-wide association study identifies novel nonadditive genetic risk variants for cardiometabolic traits. *medRxiv* (preprint), 2023.
- Zhou J, Guare L, **Rico ALG**, Gonzalez Zarzar T, Palmiero N, Assimes TL, Setia-Verma S, Hall MA. Flexibly encoded GWAS identifies novel nonadditive SNPs in individuals of African and European ancestry. *medRxiv* (preprint), 2023.

LANGUAGES

English (fluent) · Portuguese (native) · Spanish (reading and listening)