

RENATO J. QUILICHE ALTAMIRANO

Río de Janeiro, Brasil · renato.quiliche.pe@gmail.com · +55 (21) 981702387

[LinkedIn](#) · [GitHub](#)

Summary

Data scientist Ph.D. candidate in PUC-Rio Production Engineering program. Specialized in complex systems, causality science, machine learning, generative AI, ontologies, public policies, and process automation. Experienced in taking research-grade models into production-oriented, distributed implementations with PySpark, Spark MLlib and Databricks, and in translating business problems into analytical solutions end to end. Has experience in risk and disaster management, software for customer service, finance, economics, healthcare, and education. Seeking for job opportunities in both academia and private sectors.

Education

Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio)

Doctor of Philosophy in Production Engineering

Río de Janeiro, Brasil

Jun 2023 – Present

Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio)

Master of Science in Production Engineering

Río de Janeiro, Brasil

Jun 2021 – Jun 2023

Pontifícia Universidad Católica del Perú (PUCP)

Bachelor of Science in Economics

Lima, Perú

Jun 2014 – Jun 2019

Experience

Stone

Brazil

Data scientist

Nov 2025 – Apr 2026

- Developed automated experimental pipelines in Python integrating Large Language Models, Small Language Models, and formal Ontologies. Leveraged Google Colab's GPU infrastructure to optimize model inference and large-scale data processing.
- Developed formal ontologies with handcrafted and automated methods, and Retrieval Augmented Generation, applied to the case of Customer Service Chatbot

Springer Nature

Remote

Peer Reviewer

Jun 2024 – Present

- Methodology-oriented critical evaluation, focusing on the adequacy and rigor of the applied mathematical and statistical methods.
- Reviewed for: Humanities and Social Sciences Communications (IF: 3.6, Q1), Scientific Reports (IF: 3.9, Q1/Q2), Discover Sustainability (IF: 3.0, Q2), and Modeling Earth Systems and Environment (Q3).

Freelance

Remote

Developer

2023 – Present

- Developed custom PowerBI based interactive dashboards for academic domain, tracking main indicators of researchers/students performance and production
- Developed custom Streamlit interactive dashboards using Python to bridge data insights between academic research and industrial applications.
- Developed automated information collection, analysis and interpretation system using Google Workspaces
- Developed simulation models of system dynamics to support long-run policymaking

United Nations Development Program, Government of Brazil

Brazil

Consultant

Jun 2023 – Jun 2024

- Supported the collaborative design of Plano Nacional de Proteção e Defesa Civil
- Design and implementation of policy surveys targeting municipal, state and federal public officials to assess the quality of risk and disaster management policies

- Descriptive data analysis of surveys, and Natural Language Processing of interviews using Claude 3.5 Sonnet

Pontificia Universidade Católica do Rio de Janeiro

Brazil

Teaching assistance

Jun 2023 – Jun 2024

- Supported the collaborative design of Plano Nacional de Proteção e Defesa Civil
- Design and implementation of policy surveys targeting municipal, state and federal public officials to assess the quality of risk and disaster management policies
- Descriptive data analysis of surveys, and Natural Language Processing of interviews using Claude 3.5 Sonnet

Peruvian Ministry of Health

Peru

Consultant

Jun 2021

- Designed public health policies related to public insurance in post COVID-19 context
- Technical products included data analysis, visualizations, and econometric analysis of SIS-induced medical care-seeking

International Development Bank

Remote

Consultant

2018

- Supported the elaboration of trade policy guidelines for Peruvian case study
- Developed computable general equilibrium models, processed input-output matrices, applied cross-sectional and panel econometrics

Selected Publications

Identifying final demand points for aid in the aftermath of sudden-onset climate-related disasters in Peru: A supervised learning approach **2025**

Quiliche Altamirano, R. J., et al. · International Journal of Disaster Risk Reduction

DOI: 10.1016/j.ijdr.2025.105593

- Focus: multi-hazard assessment, policymaking, machine learning, hyperparameter optimization, statistics

A predictive assessment of households' risk against disasters caused by cold waves using machine learning **2023**

Quiliche Altamirano, R. J., et al. · International Journal of Disaster Risk Reduction

DOI: 10.1016/j.ijdr.2023.104109

- Focus: geoprocessing of disaster data, supervised machine learning pipelines, statistics

Using spatial patterns of COVID-19 to build a framework for economic reactivation **2021**

Quiliche Altamirano, R. J., et al. · MDPI Sustainability

DOI: 10.3390/su131810092

- Focus: geoprocessing of COVID-19 deaths and affected people, spatial econometrics, statistics, machine learning

Coding projects

Distributed Disaster-Risk Classification on Databricks **2026**

PySpark, Spark MLlib, Spark SQL, Databricks, Unity Catalog Volumes

- Re-engineered the M.Sc. dissertation modeling stack as a distributed pipeline on Databricks, migrating from single-node Python to Spark MLlib over 101K household records and 112 features (72.9K train / 28.2K test).
- Built an end-to-end Spark ML Pipeline (VectorAssembler → RobustScaler → GBTCClassifier) with cost-sensitive class weighting ($scale_pos_weight \approx 70$) to handle a 1.4% positive rate, and tuned hyperparameters with TrainValidationSplit over a ParamGridBuilder search space.
- Implemented imbalance-aware evaluation with distributed Spark SQL aggregations (MCC, G-Mean, balanced accuracy, $GIBA_\alpha$), reaching AUC 0.81 with 0.67 sensitivity and 0.83 specificity on held-out data.

- Delivered interpretability at scale through Spark feature importances and grouped risk profiling, and persisted the fitted `PipelineModel` to Unity Catalog Volumes for reproducible reload and batch scoring.

OntoUML-Stone-Public Evaluation Framework **2026**

Formal Ontologies, Python, [GitHub](#)

- Maintained and documented an open-source experimental automated evaluation framework for LLM-based customer service chatbot with ontologies

Data Science Course – PUC-Rio **2023 – 2024**

Python, Machine Learning, Data Analysis, [GitHub](#)

- In charge of the practical learning Tasks 1 and 2 to train ML models

M.Sc. Dissertation Codebase – PUC-Rio **2021 – 2023**

Python, Supervised Learning, Geoprocessing, [GitHub](#)

- Engineered end-to-end machine learning pipelines and spatial data processing scripts to model and predict household disaster risk
- Later ported to a distributed PySpark/Databricks implementation (see project above)

Honors & Awards

Bolsa doutorado nota 10 / Full PhD Scholarship **2025**

Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro - FAPERJ

- Awarded for outstanding academic performance and research production in Ph.D.

Bolsa doutorado / Full PhD Scholarship **2023**

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES

- Awarded for winner research project in Ph.D., causal discovery and system dynamics

Best B.Sc. Thesis 2021-2022 **2023**

Pontificia Universidad Católica del Perú - PUCP

- Awarded for best thesis in agricultural and development economics

Bolsa mestrado / Full MSc Scholarship **2021**

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES

- Awarded for winner research project in M. Sc., supervised learning automated classifiers

Financial grant for best essay **2019**

United Nations Educational, Scientific and Cultural Organization - UNESCO

- Awarded for best essay on Gender Female Labor discrimination in STEM fields

Technical skills

Programming: Python, SQL, R, Julia, LaTeX, Git/GitHub

Data & ML: Pandas, NumPy, Scikit-learn, PySpark, Spark MLlib, Databricks, PowerBI, Streamlit

Statistics & Modeling: Predictive modeling, supervised learning, hyperparameter optimization, imbalanced classification, econometrics, spatial analysis, system dynamics

Generative AI: GPT-5, GeminiAI, Claude, Retrieval Augmented Generation, Formal Ontologies

Languages: Spanish (Native), Portuguese (Fluent), English (Fluent)