Renan Souza, Ph.D.

contact@renansouza.org x9t36ewAAAAJ renan-souza renansouza1
0000-0002-1794-808X RenanSouza.org

Summary

Tech lead, sr. software engineer, and researcher of intelligent data and AI platforms to accelerate discovery. With 15+ years at IBM, ORNL, SLAC, and UFRJ, I translate domain expertise into production-grade systems that are easier to use and scale spanning edge, cloud, and leadership-class supercomputers. My work centers on highly scalable, low-latency, observable, provenance- and metadata-first architectures that integrate heterogeneous data systems to enable reliable, reproducible, and explainable large-scale agentic and AI workflows.

Areas of Expertise

AI/ML, LLM-driven, and Agentic workflows ◆ Edge-Cloud-HPC computing ◆ Provenance-driven data analysis, lineage, and observability ◆ Scalable data engineering (SQL, NoSQL, KGs, Streaming, Parallel Data Processing)

Education

Federal University of Rio de Janeiro, Brazil

Ph.D. in Computer Science | Sep 2015 — Dec 2019
M.Sc. in Computer Science | Jan 2013 — Jul 2015
B.Sc. in Computer Science | Jan 2009 — Dec 2012

Experience

Oak Ridge National Laboratory **Oct 2022 — Present**
Staff Scientist & Sr. Software Engineer, HPC Workflows, Data & AI **Knoxville, USA**

- Leading R&D on workflow provenance and observability for AI-driven science, focusing on transparency, reliability, and reproducibility in end-to-end workflows.
- Designing and developing provenance models and open source systems (e.g., Flowcept) to connect user intent, agent decisions, workflow executions, and downstream results in unified traces.
- Validated and applied these methods through high-profile projects in additive manufacturing, electron microscopy, and advanced biological analysis across Edge-Cloud-HPC environments.
- Published and presented results in HPC and eScience venues, and drove community engagement through tutorials and reference architectures.

IBM Research **Apr 2015 — Oct 2022**
Staff Scientist & Sr. Software Engineer, Cloud, Data & AI **Rio de Janeiro, Brazil**

- Led applied R&D on hybrid cloud and HPC data platforms for AI systems, advancing scalable architectures on Kubernetes and OpenShift for distributed, enterprise-grade workloads.
- Developed and validated knowledge graph-centric approaches for large-scale data integration, lineage, and governance across heterogeneous and distributed data stores and AI pipelines.
- Partnered closely with internal global teams and major external clients, particularly in the Energy sector, to translate research into deployable systems adopted in production.
- Produced sustained research impact through peer-reviewed publications and 10+ USPTO patents spanning provenance, polystores, AI lifecycle management, and hybrid cloud systems.

SLAC National Accelerator Laboratory **May 2013 — Dec 2014**
Research Software Engineering Intern **Menlo Park, USA**

- Applied semantic web and scalable data management methods to publish structured measurement data for broad community use.

Federal University of Rio de Janeiro **Jan 2010 — Sep 2014**
Software Engineer (Intern → Engineer) **Rio de Janeiro, Brazil**

- Led applied research on semantic web and linked open data systems, translating ontology-based models into production platforms for public-sector information access in user-facing systems.

- Developed data warehousing approaches for integrating structured and unstructured data to support big data analytics, reporting, and information discovery across heterogeneous sources.

Petrobras **May 2007 — May 2008**
IT Intern **Rio de Janeiro, Brazil**

- Early industry experience in software development and user support.

Selected Projects

Orchestrated Platform for Autonomous Laboratories (OPAL)
OPAL (FAMOUS) advances autonomous science across multiple laboratories using AI agents, robotics, and automation, enabling HPC-scale, human-in-the-loop discovery workflows.

American Science Cloud (AmSC) **2025 — Present**
AmSC is a core pillar of the DOE Genesis Mission, delivering a secure, federated platform for AI-driven science across national laboratories. At ORNL, work within the Intelligent Interface team focuses on shaping agentic AI workflows that integrate data, compute, and facilities for scalable, reusable, mission-aligned discovery.

Advanced Manufacturing into Leadership-class Supercomputers via AI Agents **2025 — Present**
A core AmSC use case demonstrating agentic AI integration between advanced manufacturing facilities and leadership-class supercomputers. I provide technical leadership in defining the end-to-end architecture and translating the scientific vision into an operational platform, including multi-agent communication, provenance-aware infrastructure, and dynamic steering across facilities.

Flowcept **2023 — Present**
Flowcept is a provenance platform that captures runtime data with low overhead and links tasks, lineage, telemetry, and AI-agent interactions into end-to-end traces for accountability and reproducibility. I created and lead the platform, which underpins multiple DOE initiatives, such as OPAL (BER/ASCR) and broader Autonomous Science (ASCR-ACT), and research work on provenance for agentic workflows.

Technical Knowledge

Programming Languages: Python, Java, C, C++, C#, Shell, NodeJS, Scala, Lua
Data Science & ML: PyTorch, MLFlow, Airflow, Pandas, Polars, Jupyter, Matplotlib, Plotly
Agentic AI: MCP, LangChain, CrewAI, Streamlit, Chainlit, RAG, LLM-based orchestration
Big Data, Streaming, and Messaging: Spark, Dask, Parsl, Kafka, Redis, RabbitMQ
Databases & Data Lakes: PostgreSQL, MySQL Cluster, MongoDB, Elasticsearch, HBase, Hive, Redis, LMDB; Object Storages, Polystores, Data lakes, Data warehouses, and Data Lakehouses
Knowledge Graphs: AllegroGraph, Jena, Virtuoso, RDF, SPARQL, OWL
Parallel & Distributed Programming: MPI, OpenMP, CUDA, PubSub
Cloud, HPC, DevOps: Kubernetes, OpenShift; Slurm, LSF; Nvidia/AMD GPU Profiling; Prometheus, Grafana

---

Selected Publications

For 50+ published papers and 10+ patents, visit https://renansouza.org/publications

- **R. Souza**, T. J. Skluzacek, S. R. Wilkinson, M. Ziatdinov, and R. F. da Silva, "Towards Lightweight Data Integration using Multi-workflow Provenance and Data Observability," in IEEE International Conference on e-Science, 2023.
- **R. Souza**, L. G. Azevedo, V. Lourenço, E. Soares, R. Thiago, et al., "Workflow Provenance in the Lifecycle of Scientific Machine Learning," Concurrency and Computation: Practice and Experience, 2021.
- **R. Souza**, A. Gueroudji, S. DeWitt, D. Rosendo, T. Ghosal, et al., "PROV-AGENT: Unified Provenance for Tracking AI Agent Interactions in Agentic Workflows," in 2025 IEEE International Conference on eScience (eScience), 2025.
- **R. Souza**, S. Caino-Lores, M. Coletti, T. J. Skluzacek, A. Costan, et al., "Workflow Provenance in the Computing Continuum for Responsible, Trustworthy, and Energy-Efficient AI," in IEEE International Conference on e-Science, 2024.