

Data Engineer

Caio Donalisio Peres Nogueira

Brazil, São Paulo, SP

Phone: +55 (11) 98922-1303

Email: caio.donalisio@gmail.com

LinkedIn: <https://www.linkedin.com/in/caio-donalisio-54705614a/>

Experience

07/2021 - Current

Data Engineer (Founding Team) – *Inspira Legal*

- > Development and operation of web crawlers for 70+ publicly available legal sources, updated daily.
- > Development of processing pipelines for 75M+ legal documents.
- > Led migration to Google Spanner database.
- > Management and operation of Elasticsearch database with text search < 1s.
- > Development and deployment of generative AI applications for processing structured legal document sets with 10,000+ pages.

> **Main Tools: Python, Google Cloud (GCP), GenAI.**

10/2019 – 12/2020

Data Analyst - *Intelligir Serviços e Tecnologia LTDA*

- > Analysis of legal process data, data exploration, development of predictive ML models and NLP techniques to optimize the triage of legal processes.

> **Main Tools: Microsoft Office, Python.**

08/2018 – 07/2019

Economic Analyst – *Macrosector Consultores. Economia & Negócios LTDA*

- > Market data research (focusing on the agricultural sector), research of national and international economic indicators, preparation of economic presentations and reports for clients, development of economic projections.

> **Main Tools: Microsoft Office, Excel VBA.**

Education

> **Bachelor of Materials Engineering.**
Escola Politécnica – Universidade de São Paulo
2017

> **Bachelor of Economic Sciences.**
FEA – Universidade de São Paulo
2024

Skills

> **Cloud & Infra:** Google Cloud (Cloud Run, Spanner, Vertex AI, GKS, BigQuery, GCS), AWS, Docker, Kubernetes.

> **Data Engineering:**
Python, Spark, Apache Beam, Web Scraping. SQL.

> **AI & Analytics:** GenAI, RAG, GraphRAG, Langchain, Machine Learning, NLP, Looker Studio. R, Office, VBA, Pandas.

Languages

> **Portuguese**
N - Native

> **English**
C2 – Fluent

> **Italian, Spanish**
B2 – Intermediate

> **French, Japanese**
B1 - Intermediate

Others

> Volunteer in the Poli Cidadã digital inclusion program (2014).

> Participant in the Challenger AI 2024 acceleration program. A program of workshops and technical mentoring in the creation and management of startups and Artificial Intelligence products, promoted and subsidized by the Estonian government (duration 30h).