

VITOR DE OLIVEIRA MAFRA

+5531999669820 | vitoromafra@gmail.com

vitormafra.com | linkedin.com/in/vitor-mafra

Belo Horizonte, Minas Gerais - Brazil

EDUCATION

Universidade Federal de Minas Gerais

2024 - Expected May 2026

Master of Science in Computer Science; Advisor: Fabrício Benevenuto, Ph.D.

Department of Computer Science

Universidade Federal de Minas Gerais

2018 - 2023

Bachelor of Science in Information Systems; Advisor: Fabrício Benevenuto, Ph.D.

Department of Computer Science

Final Paper: Rumble at first glance: a quantitative look at rumble.com, the “free speech” online video platform

EXPERIENCE

Quaest

October 2023 - Present

Data Scientist - Research and Innovation team

- Built an LLM-based synthetic focus group product, with personas representing sociopolitical clusters of Brazilians. Led the work end-to-end, from QLoRA fine-tuning — trained on millions of question-answer pairs derived from real surveys — to evaluating how well responses adhered to each group’s profile.
- Developed an MCP server that processes raw data and generates research reports, cutting the intelligence team’s turnaround from days to hours and enabling even directors to edit charts and run new data cross-tabulations using natural language via WhatsApp.
- Designed and implemented a RAG architecture for Quaest’s surveys, enabling fast, contextualized queries over large volumes of data spanning 10+ years of electoral research.
- Developed a statistical sampling validation tool that utilized Voronoi diagrams and time-series modeling of historical bipartisan data from the Lula vs. Bolsonaro political landscape, ensuring precise geographical representation and contributing to Quaest’s emergence as the most accurate pollster in the 2024 Brazilian elections.
- Created a machine learning-powered social media intelligence product (itau.com.br/imprensa/efake) adopted by Itaú, using NLP models to detect scams and brand impersonation across WhatsApp, Telegram, and Discord, enabling near real-time alerts and automated triage of risky content to protect millions of customers from fraud.

Eleições Sem Fake, DCC - UFMG

January 2021 - Present

Research Assistant (Advisor: Fabrício Benevenuto, Ph.D.)

- Collaborated with Brazil’s Superior Electoral Court (TSE) to design and implement strategies for combating misinformation during the 2022 Brazilian elections, contributing to national-scale democratic resilience.
- Developed a reporting tool using Python, Pandas, Matplotlib and L^AT_EX, that led to 30+ published articles in top-tier media outlets (e.g., The Guardian, BBC News, Folha de S. Paulo) by enabling journalists, researchers, and NGOs to uncover insights from WhatsApp and Telegram data.
- Engineered a scalable data collection pipeline using Python, Pandas and MongoDB that anonymized and stored hundreds of thousands of daily messages, images, audio, and videos from public WhatsApp groups.
- Built a computer vision pipeline using perceptual hashing to classify and differentiate TikTok and Kwai videos from our dataset, revealing that ~18% of WhatsApp videos originated from these platforms (and uncovering that Helo, ByteDance’s TikTok-like app for India, was also trending in Brazil!).
- Published 4 peer-reviewed research articles, earning 56 citations and achieving an h-index of 4.

COVID Data Analytics, DCC - UFMG

June 2020 - December 2020

Research Assistant (Advisors: Wagner Meira, Ph.D. and Evandro L.T.P. Cunha, Ph.D.)

- Developed an Instagram crawler using Python and Scrapy, to scrape, on a large scale, COVID-related data
- Managed a COVID-19 data scraping on YouTube brazilian videos, uncovering misinformation trends.

PROJECTS

Bot do Bandeirão

September 2019 - Present

Using Python, Selenium, BeautifulSoup and AWS, made a bot that scrapes my college's restaurants menus and publishes them on Twitter so that more than 6000 students could evaluate and decide where to eat. Its popularity and open-source code bootstrapped an ecosystem of similar projects around Brazil. These bots work in prominent universities such as USP, UFRJ, UFOP, and ITA, helping more than 40000 students daily.

SIC

February 2017 - December 2017

Developed a Model United Nations event at my school, fostering communication, leadership, and global issue awareness through geopolitical discussions. The event became an annual tradition and continues to thrive.

EXTRA-CURRICULAR

Fênix - Estudos Astronáuticos UFMG (Rocketry Team)

October 2019 - December 2020

Simulation team member

- Improved simulation accuracy by building a Python wind model that incorporated real-time, real-world data.
- Enhanced software engineering practices by introducing basic Git version control to the team.

Programação 101 (Coding 101)

July 2018 - July 2019

Volunteer Tutor

- Improved preparedness of ~150 first-year students by teaching C programming before classes started.

Colegium Geopolitics Studies (Geopolitics Study Group)

October 2016 - December 2017

Co-founder and team member

- Weekly meetings with the purpose of studying and discussing international politics.

YEX - Youth Experience

January 2016 - December 2022

IT team member and Volunteer Counselor

- Worked as a volunteer counselor developing leadership, English, and communication skills with young people.

YMCA of the USA

July 2015

Volunteer Camp Counselor

- Lived in Hartwell, GA with an American family and worked as a volunteer camp counselor during summer.

ACADEMIC ACHIEVEMENTS

Gold medal in XVII OBA and Silver medals in XVIII and XX OBA (Brazilian Astronomy and Astronautics Olympiad)

2014, 2015, 2017

Sociedade Astronômica Brasileira (Brazilian Astronomical Society), Agência Espacial Brasileira (Brazilian Space Agency), UFRJ (Federal University of Rio de Janeiro)

Finalist in Jornal na Escola (Essay contest organized by the largest newspaper in my state) 2015

Jornal Estado de Minas

SKILLS

Computer Languages

Python, SQL (strong), C++, C (experienced, but rusty), Go (hobbyist).

Tools

Git, Linux, Jupyter, Pandas, Numpy, TensorFlow, Pytorch, MLflow
Scikit-Learn, NLTK, SciPy, Scrapy, Playwright, Seaborn, AWS, GCP

Natural Languages

Portuguese (native), English (advanced), Spanish (interm.), Mandarin (basic)