Breno Matos

LinkedIn: linkedin.com/in/brenomatos

Email: smatosbreno@gmail.com
Homepage: brenomatos.github.io/

EDUCATION

Max Planck Institute for Informatics Saarbrücken, Germany
PhD in Computer Science **Advisors:** *Anja Feldmann and Savvas Zannettou* *August 2025 - Currently*

Federal University of Minas Gerais (UFMG) Belo Horizonte, MG, Brazil
M.Sc. in Computer Science. **Area:** *Natural Language Processing.* **Average Grade:** *97/100* *May 2021 - August 2024*
Thesis awarded a Google Latin America Research Award

Federal University of Minas Gerais (UFMG) Belo Horizonte, MG, Brazil
B.Sc. in Computer Science **Area:** *Fairness in Information Retrieval.* **Average Grade:** *81/100* *2017 - March 2021*

WORK EXPERIENCE

Kunumi Belo Horizonte, Brazil
Data Scientist *July 2024 - August 2025*

- Led the development of an in-house credit analysis specific LLM.
- Building automated corporate credit analysis tools using SOTA LLMs. Tasks include (but not limited to) semantic search using FAISS, feature extraction, summarization, implementing agents.

Public Prosecutor's Office of Minas Gerais (MPMG) at DCC/UFMG Belo Horizonte, Brazil
Research Development Software Engineer *January 2020 - July 2025*

- Worked with state-of-the-art LLMs (LLaMa and CodeLLaMa) to build a conversational text-to-SQL tool that allows users to query MPMG's databases seamlessly.
- Helped develop a crawler configuration software. Utilized expertise in web scraping to architect solutions for complex data extraction needs. The software allowed the configuration of more than 5000 crawlers to target websites.
- Developed social media monitoring tools that allowed public prosecutors to gather data for 100+ daily Instagram profiles and monitor 1000+ Whatsapp and Telegram groups. These tools are used to aid investigations and prosecutions by MPMG.

Brazilian Institute of Geography and Statistics (IBGE) at DCC/UFMG Belo Horizonte, Brazil
Full Stack Developer *October 2020 - March 2021*

- Led the website development for a novel platform that aided IBGE's analysts in calculating economic metrics such as inflation based on product availability online. The platform was built with Django and Reactjs, hosted information about 10,000+ products crawled online, and provided insights on trends and aggregated data analyses.

ACADEMIC EXPERIENCE

LATIN & LOCUS labs Belo Horizonte, Brazil
M.Sc. Researcher *May 2021 – October 2024*

- Awarded a Google Latin America Research Award
- Researched radical communities online and misinformation text patterns. Proposed a new misinformation detection task and provided the first data and baselines.
- Researched virality patterns of songs on TikTok, unraveling a re-popularization effect.
- Wrote 4 papers (3 accepted, 1 under submission)

DataCommons Brazil (partnership between Google and UFMG) Belo Horizonte, Brazil
Research Development Software Engineer *September 2023 – October 2024*

- Responsible for building crawlers with Playwright to gather data from Brazilian public agencies from 2,000+ sources. Data Commons is Google's open-data initiative to address major challenges like climate change and health epidemics through public data.

Data Science Lab (DLAB) at Swiss Federal Institute of Technology in Lausanne Lausanne, CH
Research Assistant - Summer@EPFL Intern *July 2022 - October 2022*

- Worked on a project to assess radicalization online. Built a framework to audit online recommendation systems through 100+ bots that mimicked real-user patterns. Conducted experiments to determine if platforms radicalized users through recommendations and our findings pointed to it not being the case.

Social Computing Lab (LOCUS) at DCC/UFMG — Belo Horizonte, Brazil
Research Assistant — *October 2017 - February 2021*

- Worked on early-pandemic misinformation, gathering data shared across WhatsApp and Telegram groups and generating weekly reports for journalists, who used the reports to debunk COVID misinformation.
- Worked on modeling through complex networks and helped develop a method to analyze governance using voting session data from the US House of Representatives. Also worked with backbone extraction algorithms for dense networks.

TEACHING AND SERVICES

ACM Transactions on Social Computing — Belo Horizonte, Brazil
Reviewer — *May 2024 – Present*

Computer Science Department (DCC) at UFMG — Belo Horizonte, Brazil
Head Teaching Assistant — *March 2022 - July 2022*

- Responsible for overseeing 80+ students' semester projects and weekly activities for the Introduction to Data Science course at UFMG. Work included mentoring students and answering questions related to the coursework.

SKILLS SUMMARY

- **Platforms and Programming**: Python, C/C++, JavaScript, SQL, Bash, Web

- **Tools and Frameworks**: PyTorch, Docker, Jenkins, CI/CD, Git, Elasticsearch, HuggingFace, Django, React.js, NetworkX, Jupyter, Transformers, Sklearn, Playwright

- **Languages**: Portuguese (Native), English (Fluent), French (Beginner)

SELECTED PUBLICATIONS GOOGLE SCHOLAR LINK)

- *Breno Matos*, Francisco Galuppo, Rennan C. Lima, Flávio Figueiredo **"I've Heard This Before: Initial Results on Tiktok's Impact On the Re-Popularization of Songs."** – 1st Latin American Music Information Retrieval (LAMIR) workshop, 2024.

- *Breno Matos*, Rennan C. Lima, Jussara M. Almeida, Marcos A. Gonçalves, Rodrygo Santos, **"Intellectual Dark Web, Alt-lite and Alt-right: Are They Really That Different? A Multi-Perspective Analysis of the Textual Content Produced by Contrarians"** – Social Network Analysis and Mining Journal, Springer.

- *Breno Matos*, Rennan Lima, Jussara M. Almeida, Rodrygo L.T. Santos, & Marcos Gonçalves - **"On the Presence of Abusive Language in Mis/Disinformation"**, The 13th International Conference on Social Informatics. SocInfo '22

- Carlos Henrique Gomes Ferreira, Fabricio Murai Ferreira, **Breno Mato**s, and Jussara Marques de Almeida. – **"Modeling Dynamic Ideological Behavior in Political Networks"**. The Journal of Web Science 1 (2019)

- Carlos Henrique Gomes Ferreira, *Breno Matos* & Jussara M. Almeida - **"Analyzing Dynamic Ideological Communities in Congressional Voting Networks"**. Social Informatics: 10th International Conference, SocInfo 2018

- Daniel Kansaon, Michele A. Brandão, Julio C.S. Reis, Matheus Barbosa, *Breno Matos*, Fabrício Benevenuto - **"Mining Portuguese Comparative Sentences in Online Reviews"**. XXVI Brazilian Symposium on Multimedia and the Web (2020)

- *Breno Matos*, Fabrício Benevenuto, Ana P.C. Da Silva - "A Nudez como Estereótipo de Beleza através das Lentes do Instagram" – **english: [Nudity as a Beauty Stereotype through the Lens of Instagram]**. X Brazilian Workshop on Social Network Analysis and Mining, (2021)

- *Breno Matos*, Carlos H. G. Ferreira, Jussara M. Almeida - "Analisando a governabilidade presidencial a partir de padrões de homofilia na Câmara dos Deputados: Estudos de Casos no Brasil e nos EUA" – **english: [Analyzing presidential governability based on homophily patterns in the Chamber of Deputies: Case Studies in Brazil and the USA]**. In: 7th Brazilian Workshop on Social Network Analysis and Mining (BraSNAM), 2018

AWARDS AND RECOGNITIONS

Google Latin America Research Award — Belo Horizonte, BR
Research grant structured as an unrestricted gift - 8,500 USD/42,000 BRL — *December 2023*

ICWSM-24 — Virtual
Virtual Attendance Scholarship Award — *May 2024*

Best Short Paper — Virtual
X Brazilian Workshop on Social Network Analysis and Mining, BraSNAM — *2021*

Best Full Paper Nominee — Saint Petersburg
Social Informatics: 10th International Conference, SocInfo — *2018*