Nazareno Andrade

nazareno@gmail.com +55 83 98150 5501
nazareno.github.io Google Scholar profile

Short Summary

- 12 years of applied data science
- Large experience leading teams of data scientists, developers and researchers
- Designed and implemented data-centric products (software, reports, dashboards) for NGOs, companies, and government agencies.
- Proficient in R, Python and JS for Data Science, Machine Learning, and Visualization
- Experience in various research methods, inc. qualitative research and A/B testing

Employment

2009-present: **Universidade Federal de Campina Grande (Brazil)**, Professor
- Teach Data Science, Data Visualization, Software Design, and Research Methods to Computer Science grad and undergrad students.
- Research on Data Science/Vis, Machine Learning, Civic Tech, and Social Computing.
- Develop data-centric software and analysis projects for various partners and contexts.
- Coordinate Lab Analytics, a Civic Data Science group.

2009-2011: **TUDelft (Netherlands)**, Post-doctoral researcher
- Applied mixed methods to study peer-production online communities
- Supervised PhD students
- Collaborated on a multidisciplinary team (computer scientists, economists, physicists, sociologists, and mathematicians).

2008 (6 months): **University of British Columbia (Canada)**, Visiting Schoiar
- Research on characterizing user behavior in peer production systems

2004 (9 months) **Hewlett Packard Labs (UK)**, Research intern
- Research on Online Communities
- Developed parallel computing software that ran on a 100-node datacenter

Research

I have worked with controlled experiments, characterization of large datasets, conception of theoretical frameworks, semi-structured interviews, focal groups, and ethnography. I have:

- Created a trajectory-based model for how listeners' attention changes and fixates on artists. The model was learnt from 200M listens from Last.fm. (ISMIR paper, code)

- Systematized how practitioners design Anthropographics - Data Visualizations to promote compassion - (IEEE TCGV paper, open science, catalog) and conducted A/B tests to assess current practices' efficacy (ACM CHI paper, open science).
- Leveraged mixed-methods to show how the design of Stack Overflow is perceived by users with varying cultural backgrounds (ACM CSCW paper) and to examine user experience with a situated physicalization (Pacificvis paper)
- Characterized multifaceted contributor profiles in five large Q&A sites from the StackExchange platform using cross-sectional and longitudinal views (CSCW, paper)
- Applied a model of disruption and consolidation in innovation to find patterns in a network of artist influences from AllMusic (ISMIR paper, code) and latent-space modelling to study a dynamic network of jazz artists from discogs (ISMIR paper, code)
- Conducted a controlled experiment with five variants of a social jukebox to examine how different designs mediate conflicts (ISMIR paper). with quant. and quali. results.
- Organized the Recys 2020 Challenge for predicting user engagement with tweets.

This research has been published in venues such as ACM CHI (Human Factors in Computing Syst.), ACM CSCW (Computer-Supported Cooperative Work and Social Computing), and IEEE Transactions on Visualization and Computer Graphics. Collaborators include colleagues from UFMG (Brazil), TUDelft (Netherlands), UBC (Canada), Univ. of Washington (US), INRIA (France), IBM Research (Brazil, US), and Sorbonne Univ. (France).

Data and Technology projects

I coordinate since 2012 the Lab Analytics, a Civic Data Science group I co-founded at UFCG. The lab usually has 20-30 researchers and software engineers and developed projects funded by Hewlett Packard, government agencies, startups, and NGOs.

These are some of the projects resulting in products in operation:

- Parlametria is a platform that uses NLP and predictive models to inform the advocacy actions of a network of more than 100 NGOs.
- Datapedia Eleições and Olho N'Água are visualization platforms/dashboards to explore large historical databases about elections and water reservoirs, respectively.
- Tá de Pé weekly crunches ~100GB to help citizens analyze governmental purchases of public school meals and goods to fight the pandemic in Brazil.
- Data analysis reports for varied audiences, including an analysis of campaign donations received by members of the Brazilian Congress, a retrospective of environmental law bills in the Congress in the last 2 years, and an exploration of patterns in calls to a support line for Cancer patients.

In all projects, I was the main responsible for communication with stakeholders, software architecture, scientific/technical lead of data analysis and modelling, and participated in coding.

Hacktivism

I conceived, helped organize, and facilitated four editions of the Hackfest for Civic Participation, promoted by governmental auditing agencies, civil society and Academia. The largest Hackfest had 200 participants in a hackathon and 1500 attendants in talks and workshops. I also facilitated two replications of the event in Brazil and one in Colombia.

My team won the 1st prize in a national competition for Apps to Confront Corruption by the Brazilian Federal Government. I was a full-stack developer and data scientist in the project.

House of Cunha is a project that combines data visualization and statistics to unveil coalitions in the lower house in Brazil during 2014. The site got extensive media coverage and received more than 100K visitors. I talk about it in this TEDx talk.

As a consequence of these projects, I have spoken in national events organized by Government agencies (e.g. Data Analysis for the Public Administration Seminar) and represented hacktivism in boards related to social participation and transparency.

Education

2008 - PhD in Electrical Engineering (on a Computer Science topic) from UFCG.
2004 - MSc in Computer Science from UFCG, magna cum laude
2002 - B.Tech in Telematics from IFPB.

Relevant Technology Skills

Extensive experience programming in R, Python, Javascript (including D3.js), Java and SQL. Some experience with Spark and Google Big Query. Experience with dashboard design and storytelling in Tableau and Flourish. Examples of training material (pt-br) are online on this site, code notebooks, and youtube playlist.

Other interests

Playing musical instruments, djing, vinyl collecting, urban cycling, and dogs. Visualizing music-related data for fun