

# Tiago da Silva

tdsh97@gmail.com | [github.com/tiagodsilva](https://github.com/tiagodsilva) | [tiagodsilva.github.io](https://tiagodsilva.github.io) | [linkedin.com/in/tiagodasilvah](https://linkedin.com/in/tiagodasilvah)

## Education

---

### PhD in Applied Mathematics

2024/03 – 2024/12

School of Applied Mathematics, Brazil. Advisor: [Prof. Diego Mesquita](#).

- Thesis title: Streaming, Distributed, and Asynchronous Amortized Inference.
- Eligible for fast-track graduation due to recognized academic excellence and scientific productivity.

### MSc in Applied Mathematics

2024/03 – 2024/12

School of Applied Mathematics, Brazil. Advisor: [Prof. Diego Mesquita](#).

- Dissertation title: Expert-Aided Discovery of Ancestral Graphs.

### BSc in Data Science

2020/03 – 2023/12

School of Applied Mathematics, Brazil

- GPA: 9.9/10.0. 1st in class and 1st in the entrance exam. Received a fully-funded scholarship.

## Selected Publications

---

- When do GFlowNets learn the right distribution?** ICLR 2025 (spotlight)
  - [da Silva, T.](#), Silva, E., Alves, R., Souza, A., Garg, V., Kaski, S., Mesquita, D.
  - TL;DR: We show that a GFlowNet's learning objective may have an unattainable global minimum, which is undetectable by standard diagnostics. To address this, we propose the first tractable metric for assessing GFlowNets.
- Generalization and Distributed Learning of GFlowNets?** ICLR 2025
  - [da Silva, T.](#), Souza, A., Rivasplata, O., Garg, V., Kaski, S., Mesquita D.
  - TL;DR: We devise the first PAC-Bayesian generalization bounds for GFlowNets. Inspired by them, we also introduce the first general-purpose distributed learning algorithm for GFlowNets, drastically accelerating training convergence.
- Streaming Bayes GFlowNets** NeurIPS 2024
  - [da Silva, T.](#), Souza, D., and Mesquita, D.
  - TL;DR: We design a method to update GFlowNets trained on a streaming Bayesian posterior. Experiments show a drastic reduction in training time when compared against learning from scratch a model based on the entire dataset.
- On Divergence Measures for Training GFlowNets** NeurIPS 2024
  - [da Silva, T.](#), Silva, E., and Mesquita, D.
  - TL;DR: We empirically show that the inefficacy of divergence-based objectives for GFlowNets is due to their large gradient variance. We then develop variance reduction techniques that significantly accelerate training convergence.
- Embarrassingly Parallel GFlowNets** ICML 2024
  - [da Silva, T.](#), Souza, A., Carvalho, L., Kaski, S., and Mesquita, D.
  - TL;DR: We propose a divide-and-conquer approach to train a log-pool of GFlowNets in an embarrassingly parallel fashion. Results show a significant speed up in learning when the unnormalized target is expensive to evaluate.
- Exploring scientific literature by textual and image content using DRIFT** Computer & Graphics 2022
  - Pocco, X., [da Silva, T.](#), Poco, J., Nonato, L. G., Gomez-Nieto, E.
  - TL;DR: We developed a text- and image-driven visualization-based search engine for scientific literature.

## Preprints & Workshops

---

- Human-aided Causal Discovery of Ancestral Graphs** LatinX @ NeurIPS 2024
  - [da Silva, T.](#), Silva, E., Góis, A., Heider, D., Kaski, S., Mesquita, D., Ribeiro, A.
  - TL;DR: We devise a Bayesian human-in-the-loop algorithm for causal discovery under latent confounding.

## Employment

---

### Rei do Pitaco

2025/01 – 2025/11

- Data Scientist.
  - Spearheaded the development of the company's bet builder, allowing customers to create customized bets. I designed, implemented, and maintained both the algorithm and the web API responsible for the real-time odds calculations.
  - Handling up to 1.2 million requests per day, the service accounted for 50% of the company's sportbook revenue.
  - Left to pursue and expand my academic interests overseas.

- Software Engineer.
  - Extended the company's API to support larger workloads via asynchronous processing and caching.
  - Led the adoption of Infrastructure as Code (IaC) using Terraform as a standard for the company's web scraping services.
  - Assisted in the migration of the company's infrastructure between cloud providers (AWS and Azure).

- Data Science intern.
  - Designed predictive models to define the opening lines of bets on the outcomes of sport events (bookmaking).
  - Deployed and upheld the created models within applications serving thousands of concurrent users.

## Honors & Awards

---

**Award for Academic Excellence**, Brazilian Society of Applied and Computational Mathematics. 2023

**First place**, School of Applied Mathematics entrance exam. 2020

I was awarded 19 prizes in scientific competitions during high school, including:

**William Glenn Whitley Prize** for achieving the highest score on the State Mathematical Olympiad. 2019

**Top score in the country**, Brazilian Mathematical Olympiad of Public Schools. 2019

**Top score in the country**, Brazilian Mathematical Olympiad of Public Schools. 2018

**Gold medal**, Brazilian Chemistry Olympiad. 2018

**Gold medal**, Brazilian Mathematical Olympiad of Public Schools. 2017

**Gold medals**, State Chemistry Olympiad. Highest score in 2019. 2016-2019

## Research Experience

---

**Mohamed bin Zayed University of Artificial Intelligence**, UAE 2025/10 – ongoing

Working on probabilistic machine learning and Bayesian inference.

**Green AI Lab**, Brazil 2022/08 – 2025/10

Worked on streaming, distributed, and asynchronous algorithms for probabilistic machine learning.

**Aalto University**, Finland 2024/07 – 2024/10

Visiting scholar at the PML group. Worked on the generalization and expressivity of amortized inference algorithms.

**Visual Data Science Lab**, Brazil 2020/08 – 2023/01

Worked on the development of an open-source library for reverse engineering visualizations ([REV](#)).

## Teaching

---

I worked as a teaching assistant for four years at the School of Applied Mathematics in courses on statistics and machine learning. I assisted the professors in designing and grading homeworks. Also, I held office hours to support students.

## Languages

---

Portuguese (Native), English

## Skills

---

Computer languages: Proficient with [Python](#) and [SQL](#). Competent with R and Stan and Rust. Familiar with C++.

Scientific computing frameworks: PyTorch, PyTorch Geometric, GPyTorch, NumPy, SciPy, Jax.

Technologies: Git, Linux, Docker, FastAPI, Datadog, Kubernetes, Argo, Tailscale, Kafka.

## References

---

Diego Mesquita [diego.mesquita@fgv.br](mailto:diego.mesquita@fgv.br)

Amauri Souza [amauri.souza@aalto.fi](mailto:amauri.souza@aalto.fi)