

# 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

- Thesis title: Streaming, Distributed, and Asynchronous Amortized Inference.
- Advisor: [Prof. Diego Mesquita](#).

**BSc in Data Science** 2020/03 – 2023/12

School of Applied Mathematics, Brazil

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

## Selected Publications

1. **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.

2. **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.

3. **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.

## Research Projects

1. **When do GFlowNets (not) learn the right distribution?**

GFlowNets have demonstrated exceptional performance in, e.g., NLP and combinatorial optimization. However, an understanding of the limitations of GFlowNets and a sound procedure for assessing the closeness of a trained model to its learning objective are notably absent from the literature. In this project, we aim to address questions such as: given a parametric model, which distributions can a GFlowNet learn? How to measure the closeness of the sampling distribution to the target?

2. **Do GFlowNets generalize?**

Generalization is at the core of GFlowNet learning: during training, only a portion of the state space is explored and can be used for risk minimization. In this context, we ask: can we obtain (the first) non-vacuous statistical certificates for GFlowNets? Also, which algorithmic changes would (provably) boost the generalization performance of GFlowNets?

## Research Experience

**Green AI Lab, Brazil** 2022/08 – ongoing

Working with probabilistic ML. My recent efforts were mostly directed towards leveraging GFlowNets for asynchronous and approximate Bayesian inference. I have also worked on geometric deep learning, learning theory, variational autoencoders, diffusion probabilistic models, and PINNs. Our current research led to publications at [ICML](#) and [NeurIPS](#).

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

I was a visiting scholar on the Probabilistic Machine Learning group under the supervision of Prof. Vikas Garg and Prof. Sami Kaski. I worked on developing (non-vacuous) statistical guarantees for GFlowNets and on geometric deep learning.

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

Research assistant supervised by Prof. Jorge Poco. I assisted the development of a framework for reverse engineering of visualizations (see the open-source library [REV](#)) and of a platform for image-based literature search (see our [C&G paper](#)).

<sup>1</sup>Lowest-passing grade of 6.0.

## Honors & Awards

---

<b>Award for Academic Excellence</b> , Brazilian Society of Applied and Computational Mathematics.	2023
<b>First place</b> , School of Applied Mathematics entrance exam.	2020
I was awarded <u>19 prizes in scientific competitions</u> during high school, including:	
<b>William Glenn Whitley Prize</b> for achieving the highest score on the State Mathematical Olympiad.	2019
<b>Top score in the country</b> , Brazilian Mathematical Olympiad of Public Schools.	2019
<b>Top score in the country</b> , Brazilian Mathematical Olympiad of Public Schools.	2018
<b>Gold medal</b> , Brazilian Chemistry Olympiad.	2018
<b>Gold medal</b> , Brazilian Mathematical Olympiad of Public Schools.	2017
<b>Gold medals</b> , State Chemistry Olympiad. Highest score in 2019.	2016-2019

## Employment

---

<b>Rei do Pitaco</b> (largest fantasy sports company in Brazil)	2023/01 – 2023/07
<ul style="list-style-type: none"><li>• Data Science intern.<ul style="list-style-type: none"><li>▸ Designed predictive models to define the opening lines of bets on the outcomes of sport events (bookmaking).</li><li>▸ Deployed and upheld the created models within applications serving thousands of concurrent users.</li></ul></li></ul>	

## Teaching

---

I have worked as a teaching assistant (TA) for over three years in the School of Applied Mathematics. I was a TA in the courses of Exploratory Data Analysis (2021.1), Linear Algebra (2021.2), Probability (2022.1), Statistical Inference (2022.2), Machine Learning (2023.1), Time Series (2023.2, 2024.2), and in the graduate-level course of Machine Learning (2024.1). I assisted the professors with preparing and grading both homework and exams, and held office hours to support students.

## Languages

---

Portuguese (Native), English

## Skills

---

Computer languages: Proficient with Python and SQL. Competent with R and Stan. Familiar with C++ and JavaScript.

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

Technologies: Git, Linux.

Data visualization: Matplotlib, Altair, Vega-lite, D3.

Computer vision libraries: OpenCV, YOLOv5, SAM.

## References

---

Diego Mesquita	<a href="mailto:diego.mesquita@fgv.br">diego.mesquita@fgv.br</a>
Amauri Souza	<a href="mailto:amauri.souza@aalto.fi">amauri.souza@aalto.fi</a>