Ivan Ogasawara

Open Source Developer / Research Software Engineer / Data Scientist

Summary

With over 22 years of experience in software development, I am deeply passionate about compilers and data science. My expertise spans multiple programming languages including Python, C++, and JavaScript, applied across backend, frontend, DevOps, and packaging. I have also contributed as a Research Software Engineer (RSE) to scientific projects in transportation engineering and public health. I have contributed to renowned open-source projects such as PyTorch, ibis-framework, and the Jupyter ecosystem. I have developed and maintained diverse tools and libraries for DevOps, compilers, and automation, using advanced technologies like LLVM, PostgreSQL, Docker, OpenAI ChatGPT API, hugging face, etc. As a leader of Open Science Labs, I foster a community dedicated to open science and open source software. I have also been working as a teaching assistant in Data Science and Programming courses.

Education

2015–2016	**Transportation Engineering**, *Universidade Federal de Santa Catarina*, Florianópolis/Brazil, *Master's degree (interrupted)*.
2004–2005	**Information Technology**, *Faculdade Eniac*, Guarulhos/Brazil, *Lato Sensu Post-Grad*.
2003–2004	**Database Technology**, *Faculdade Eniac*, Guarulhos/Brazil, *Associate's degree*.

Experience (last 10 years)

2024–Present   **Teaching Assistant**, *WorldQuant University*, Part-time.
Teaching Assistant in the course 'Applied Data Science Lab', Database Management, Data Cleaning and Preprocessing, Regression and Classification Modeling, Data Visualization, Ethics in Machine Learning, Business Insight and Intelligence.

2024–Present   **Teaching Assistant**, *The GRAPH Courses*, Part-time.
Teaching Assistant in the course 'Python Basics and Beyond: Intro to Data Analysis with Python', focusing on topics such as Fundamentals of Python, Pandas, Plotly, Quarto Dashboards, and GitHub.

2015–Present   **Executive Director**, *Open Science labs (OSL)*, Part-time.
At Open Science Labs (OSL), my role spans devops, community management, and program coordination. I established the Incubator and Open Source Internship Programs, enhancing our outreach and significantly impacting the open-source community. OSL has thrived successfully participating in the Google Summer of Code through NumFOCUS, with four students completing their projects to date. The Internship Program has also flourished, aiding seven individuals in completing their projects. I currently manage more than 10 projects in incubation and three affiliated projects. My efforts include securing three grants from the Python Software Foundation, providing mentorship to recipients, and fostering various internal mentorship initiatives. Currently, I am serving as a tech leader, developer, and mentor in open-source projects such as Makim, Sugar, ArxLang, SciCookie, Envers, Anamnesis.ai, etc.

2023–Present   **Research Software Engineer**, *The GRAPH Network/LiteRev*, Part-time.
In the LiteRev project, I've been acting as a Tech and Team Leader, Software Architect, and DevOps engineer. In this project, we have been using technologies such as: Python, Django, scikit-learn, pacmap, optuna, hdbscan, bokeh, celery, docker-compose, CI/CD, etc

2022–2023   **Research Software Engineer**, *The GRAPH Network*, Part-time.
As a server administrator, I was responsible for ensuring the servers are operating efficiently and implementing new services for internal usage. Additionally, I provided internal training to the team on topics such as git, Python tools, and best practices. I was also responsible for defining and implementing best practices for open source projects and designing the architecture for new implementations. My expertise included working with technologies such as Python, R, CI/CD, Docker, PostgreSQL, Ansible, Nginx, Apache Superset, and packaging with conda.

2022–2023 **Software Developer**, *Inlyse*, Part-time.
I acted as a project leader of a homemade platform for customers where customers can manage their license keys (for Inlyse products) and manage the payments for their license through a stripe integration. I have worked on some tools for feeding the internal pipeline for ML training models. Some of the technologies used: Python, C++, Bash, JavaScript, Ansible, Django, Docker, CI/CD, Stripe, etc.

2018–2021 **Software Developer**, *Quansight/OSBIG*, Part-time.
Contributed to Open Source projects such as PyTorch, Ibis-Framework, Jupyter ecosystem, OmniSciDB. Worked on internal and client projects with technologies such as Django, Django-Rest Framework, asyncio, Bluetooth BR/EDR communication, Docker, GraphQL, Argo Workflow, etc. As extra activities, I was a member of the DEI committee, where we've organized internal events to promote and discuss topics related to diversity, equity, and inclusion, and I have also organized and facilitated weekly meetings about compilers.

2017–2017 **Software Packager**, *Anaconda, Inc*, Contract.
Acted as a software packager with conda for libraries in Python, R, and Java.

2016–2017 **Research Software Engineer**, *Fundação para o Desenvimento Científico e Tecnológico em Saúde - Fiotec*, Contract.
Contributed to the development of portals for visualization of influenza and dengue incidence and estimation in Brazil. Development using technologies such as Python, Javascript, pandas, bokeh, django, flask, fiona, rasterio, PostgreSQL, MapServer, Docker, etc.

2016–2016 **Research Software Engineer**, *Fapeu - Fundação de Amparo a Pesquisa e Extensão Universitária*, Contract.
Contributed to a web system for analyzing the visco-elastic characteristics of pavements in laboratory tests.

2013–2016 **Research Software Engineer**, *Fapeu - Fundação de Amparo a Pesquisa e Extensão Universitária*, Florianópolis/Brazil, Full-time.
Analyzed international literature to investigate methods in the transportation engineering area, such as high-speed Weigh-in-Motion (WIM) and fatigue and complex modulus tests. Developed solutions using Digital Signal Processing, Machine Learning, and High Performance Computing techniques.

Volunteer Experience

2020–2023 **Software Reviewer, Review Editor**, *PyOpenSci*.
Served as a software package reviewer, review editor, and advisor for pyOpenSci.

2015–2016 **Chair**, *SciPy Latin America*.
SciPyLA Conference 2016, held between May 16 and 20, 2016, aimed to promote: science, the adoption of information technology as a tool for scientific study and, the use of Python as the main computing tool in science (https://www.scipy.lat/conf/2016/).

Languages

Portuguese: Native, Spanish: Fluent, English: Advanced, Italian: Basic

Certificados

- Compilers. edX/Stanford (2021)
- Leading With Effective Communication, Inclusive Leadership Training. edX/Catalyst (2021)
- Parallel Computing with Dask. DataCamp (2018)
- Deep Learning in Python. DataCamp (2018)
- Machine Learning with Python. DataCamp (2018)
- Machine Learning with the Experts: School Budgets. DataCamp (2018)
- Unsupervised Learning in Python. DataCamp (2018)
- Importing & Managing Financial Data in Python. DataCamp (2018)
- Supervised Learning with scikit-learn. DataCamp (2018)
- Introduction to Applied Biostatistics: Statistics for Medical Research. Osaka University / EDX (2017)
- Intro to Hadoop and MapReduce. Udacity (2016)
- Intro to Data Science. Udacity (2014)
- Machine Learning: Supervised Learning. Udacity (2014)
- Intro to Descriptive Statistics. Udacity (2014)
- Design of Computer Programs. Udacity (2012)