# Márcia Vagos

*Curriculum Vitae*

*Aarhus, Denmark*
*+45 93901197*
*marcia.vagos@gmail.com*
*marciavagos.github.io*
*linkedin.com/in/marcia-vagos*

## About me

I am a technical leader with a PhD in Scientific Computing and a decade of experience developing data-driven solutions across research and industrial projects. My background spans machine learning, signal processing, computational modeling, software engineering, and scalable computing systems. I specialize in turning complex, high-dimensional data into reliable solutions by combining algorithm development with robust software engineering.

I enjoy working at the intersection of scientific applications and engineering: understanding challenging problems, developing robust algorithms, and building efficient systems that deliver practical value. I am looking for roles where I can draw on my technical experience, from low-level computational solutions to high-level system architecture, to shape and deliver complex technical systems.

## Key Competencies

Scientific Software Development | Distributed Computing | Python Software Engineering | Numerical Simulation | Data Engineering | Machine Learning | MLOps | Accelerated Computing | Cloud Native Architecture | Signal Processing | Algorithmic Development | AI Infrastructure | Production Systems | Requirement Analysis | Stakeholder Coordination | Technical Leadership

## Education

2016–2020 **Ph.D. in Computational Biology**, *University of Oslo, Department of Informatics*.
Early Stage Researcher within the Marie Curie AfibTrain-Net project

2008–2013 **Integrated Master's in Bioengineering, spec. Biomedical Eng.**, *University of Porto*, Portugal.

## Professional Experience

03/2026–Present **Freelance Data Scientist**, Marsian Consulting, Aarhus, Denmark.
**Client:** Terma Group
Full-time, 100% position

- Led the development of machine learning and signal-processing solutions for sensor-based maritime situational awareness systems.
- Developed algorithms to extract meaningful features from large-scale Distributed Acoustic Sensing (DAS) data and correlate acoustic patterns with maritime activity.
- Designed and optimized data-processing pipelines for high-volume sensor datasets, enabling efficient analysis through scalable computing approaches.
- Deployed and maintained edge-computing infrastructure for continuous sensor monitoring and real-time data processing.
- Supported field deployment, installation, and calibration of advanced sensing systems.
- Collaborated with domain experts and technical stakeholders to translate operational requirements into robust software and data-processing solutions..

11/2024– 02/2026 **Senior Data Scientist**, Systematic A/S, Aarhus, Denmark.
AI Engineer for applications in defense
Full-time, 100% position

- Designed and deployed production-grade machine learning systems, including computer vision and AI applications for complex real-world datasets.
- Built end-to-end ML pipelines for satellite imagery applications, covering data ingestion, preprocessing, model development, evaluation, and monitoring.
- Implemented land cover classification solutions using transformer-based models for satellite image analysis.
- Developed scalable AI services and microservice architectures to integrate machine learning capabilities into enterprise systems.
- Designed evaluation frameworks for LLM-based systems, combining automated assessments, human feedback, and quantitative evaluation methods.
- Integrated LLMs with external tools and APIs to build reliable AI assistants and improve system capabilities.
- Developed signal-processing and data-analysis pipelines for large-scale Distributed Acoustic Sensing (DAS) datasets used in underwater activity monitoring.

01/2021– 10/2024 **Data Scientist**, Semcon Norge AS, Kongsberg, Norway.
Consultancy within data science, software development and R&D
Full-time, 100% position

- Developed and deployed machine learning models for medical imaging, energy forecasting, and sensor data analysis.
- Built deep learning models (TensorFlow, PyTorch) for image enhancement and analysis of clinical datasets.
- Designed ML solutions integrated into healthcare workflows, supporting clinical decision-making.
- Developed cloud-based ML systems using AWS (EC2, S3, SageMaker).
- Established CI/CD pipelines and software engineering best practices for ML development.
- Led product development initiatives and mentored junior engineers.
- Collaborated with stakeholders across engineering, product, and clinical domains.
- Leveraged Large Language Models (LLMs) with the Retrieval-Augmented Generation (RAG) framework to efficiently query and extract relevant information from textual data and document databases.

03/2016– 09/2020 **Ph.D. Fellow**, Simula Research Laboratory, Oslo, Norway.
Doctoral degree obtained from the University of Oslo, Department of Informatics
Full-time, 100% position

- Developed novel computational models with applications in numerical simulation of cardiac electrophysiology.
- Implemented large-scale simulations on computer clusters by employing distributed computing strategies to ensure efficient and scalable processing.
- Performed advanced exploratory data analyses and extracted meaningful features from simulated cardiac electrophysiology data, contributing to the understanding of the onset mechanisms of atrial arrhythmias.
- Established a framework for uncertainty quantification and sensitivity analysis using a combination of Bayesian and multivariate regression methods, leading to several publications in scientific journals.
- Developed methods to efficiently model parameters by assimilating experimental data to enhance model reliability and predictive capability.

08/2017–01/2018 **Research Visitor**, School for Cardiovascular Diseases, Maastricht University, Maastricht, Netherlands.
Research collaboration
Full-time, 100% position

- Developed numerical models for electrophysiological simulations, validated against experimental data in collaboration with researchers at the Maastricht University Hospital.

02/2014–07/2015 **Research Assistant**, Forschungszentrum Informatik, Karlsruhe, Germany.
Research work in collaboration with the Karlsruhe Institute of Technology
Full-time, 100% position

- Built a state-of-the-art 3D model of the human eye to simulate light-tissue interactions and performed simulations of thermal damage in the retina induced by laser irradiation using the Finite Volume Method.

02–07/2012 **Research Intern**, Aalto University, Espoo, Finland.
Exchange student with the Erasmus Program in the Neuroinformatics Group
Full-time, 100% position

- Applied advanced statistical modeling techniques, such as ICA and sparse PCA, to brain fMRI data to decorrelate activation patterns based on stimuli information.

## Honors & Awards

2024 Granted patent (in force): "Radiotherapy System Using Graph Network To Account For Patient Movement". Inventors: Per Håvard Kleven, Yapi Donatien Achou, Márcia Vagos, Mari Engebretsen Ryttervold, Karsten Rydén-Eilertsen. Patent No. SE545891 C2, 2024. (link)

2016 Marie Curie Early Stage Researcher Fellowship within the AfibTrain-Net project

## Professional Service

2015–2020 View my publications on ResearchGate (link) and Google Scholar (link).

2021–2023 Grand Award Judge at the International Science and Engineering Fair.

## Technical Skills

**Programming languages:** Python, C++, Matlab, Bash

**Machine Learning:** PyTorch, TensorFlow, Scikit-learn

**Data Science:** NumPy, CuPy, Pandas, cuDF, SciPy, Numba, OpenCV, Scikit-image

**MLOps & Deployment:** Docker, Kubernetes, FastAPI, Git, CI/CD

**Distributed & cloud computing:** PySpark, Dask, Slurm, AWS, GCP

**Data Engineering:** PySpark, Dask

**Databases:** SQLite, MongoDB, InfluxDB

**Data visualization & UI:** Matplotlib, Plotly, Streamlit, Dash, PyQt, ParaView

## Regulatory & QA

- Completed training in IEC 62304 (Medical Device Software Lifecycle Processes)
- Experience developing ML solutions aligned with healthcare and clinical requirements
- Experience with software testing, validation, and delivering documentation requirements in regulated environments
- Experience supporting reproducibility, traceability, and auditability of ML systems

## Leadership & Communication

- Led a team of data scientists and developers at a MedTech startup developing a software platform for medical image analysis and augmentation, while coordinating clinical testing in collaboration with Radiumhospitalet in Oslo.
- Led R&D initiatives across multiple projects, including developing and deploying an LLM evaluation framework adopted across the department, implementing a cloud-based analytics pipeline for predicting in-house energy consumption, and establishing internal data governance practices.
- Trained and mentored junior developers.
- Published scientific papers in high-impact journals and regularly deliver technical reports of projects and product specifications.
- Delivered presentations and technical talks at various venues, including international conferences, local events, and corporate seminars.

## Courses

| | |
|---|---|
| Sep 2025 | Artificial Intelligence Academy, EITCI Institute (ongoing) |
| Nov 2025 | Docker & Kubernetes (instructor led), Lund&Bendsen (online, 3 days) |
| Jul 2024 | Amazon Elastic Compute Cloud (Amazon EC2) - Automate Your Infrastructure (link), AWS training and certification |
| Jun 2024 | Build with Amazon EC2 (link), AWS training and certification |
| Apr 2024 | Fundamentals of Analytics on AWS - Part 1 & 2 (link), AWS training and certification |
| Mar 2024 | Enhancing data science outcomes with efficient workflows (link), Nvidia Deep Learning Institute |
| Mar 2024 | Fundamentals of accelerated computing with CUDA Python (link), Nvidia Deep Learning Institute |
| Jan 2024 | Fundamentals of Analytics on AWS - Part 1 (link), AWS training and certification (online) |
| Nov 2023 | Google Cloud: Big Data and Machine Learning Fundamentals (link), Google (online, 9 h) |
| Nov 2023 | Google Cloud: Launching into Machine Learning (link), Google (online, 14 hours) |
| Sep 2022 | IEC 62304 Medical Device Software, Software Life Cycle Processes (link), SoftComply (online, 2 days) |
| Aug 2022 | PSI Summer School in Particle Physics (link), Zuoz, Switzerland (1 week) |
| May 2021 | Data Modeling with MongoDB (link), MongoDB University (online, 1 week) |
| May 2017 | RegML: Regularization Methods for Machine Learning (link), Simula Research Laboratory, Norway (1 week) |
| Jan 2017 | Geilo Winter School in Machine Learning (link), Geilo, Norway (1 week) |
| Jul 2016 | Suurph Summer School in Computational Physiology (link), Simula Research Laboratory, UCSD, San Diego (4 weeks) |
| Jun 2016 | Image-based Biomedical Modeling Summer School (link), Utah State University (2 weeks) |

## Languages

Portuguese: *Mother tongue* • English: *Proficient* • Spanish: *Advanced* • Norwegian: *Intermediate* • German: *Basic* • Japanese: *Basic*