Fernando Cruz

Data Scientist

fernandocruz184@gmail.com +351 918509672 Viana do Castelo, Portugal
https://cruz-f.github.io/ https://www.linkedin.com/in/fernando-cruz-ba2b65142/

Senior Data Scientist with a **Ph.D. in Biomedical Engineering and 5+ years** of experience leading the design, development, and deployment of AI systems. Main developer of Cegid Pulse OS, currently delivering **multi-agentic AI workflows** for Cegid's **core products**. Experience in **data engineering** (Databricks & ETL) and **full project delivery for major clients**. Academic experience includes lecturing on **Data Science** at the university level. Proven success in fast-paced environments, including winning the **8th Startup Braga Acceleration Program**.

EXPERIENCE

Cegid, Portugal — *Data Scientist*

02/2024 - PRESENT

Main developer of Cegid Pulse OS, a production-grade platform that currently offers multi-agentic AI workflows for two Cegid products (small businesses & finance) in production and being beta tested with several others.

Led the implementation of Cegid Pulse OS's main features, including **AI workflows**, **context engineering**, and **agentic orchestration**. Enabled agents to take actions in the product using either tools (**APIs' proxies)** or **MCP**.

Installed **monitoring, observability, and evaluation** frameworks to ensure platform performance and quality in production.

Contributed to organizational growth by **leading small squads**, **mentoring interns,** and **guiding technical interviews** for the AI department.

Presented and showcased Cegid Pulse OS to internal teams, driving adoption of Pulse OS and Generative AI across Cegid.

OmniumAI, Portugal — *Software Engineer*

07/2021 - 12/2023

Co-founded the CibusAI platform, securing victory in the **8th Startup Braga Acceleration Program** among a competitive pool of startups (1). The platform is currently piloting with clients in the food science industry.

Led the **end-to-end development of a genomics data platform for a major client**, Chr. Hansen, including requirements gathering, solution design, technical solution development (full-stack web application and data pipelines), and deployment on the client's infrastructure.

Main developer of Omnia, an **automated machine learning framework** for biological and chemistry predictive models, accelerating model generation and deployment for client projects. Omnia was used to accelerate CibusAI development.

Engineered **end-to-end data pipelines** in **Databricks using ETL tools** to process raw biological and chemistry data into golden datasets. Created, tuned, and deployed **machine learning models** for CibusAI.

HARD SKILLS

LLM/AI Tools: Agents, Workflows, RAG, Azure AI

ML Frameworks: PyTorch, Scikit-Learn, Pandas, Numpy, Spark MLlib, MLflow

Data Engineering: Databricks, PySpark, Delta Lake

Software Engineering & Architecture: Object-Oriented Programming, Concurrency (Async), Domain-Driven Design, Clean Architecture, RESTful API, Pub/Sub & Message Broker

Languages: Python, SQL, JavaScript & TypeScript

Cloud/DevOps: Azure, Kubernetes, Docker, Git

Web/Backend: FastAPI, Django, React.js, HTML & CSS

Databases: PostgreSQL, MongoDB

Agile: Scrum, Kanban

SOFT SKILLS

Leadership & Entrepreneurship driven by entrepreneurial experience

Critical thinking driven by academic background

Agile &problem-solving driven by the implementation of novel platforms in world-leading companies

Universidade do Minho, Portugal — *Invited Assistant Professor*

09/2022 - 02/2023

Lectured on the Data Science course in both Master's of Bioinformatics and Biomedical Engineering, covering core principles and applications for biology and chemistry. Assessed and evaluated student projects to measure technical proficiency and understanding.

Lectured on the Software Engineering Labs course in the Master's of Bioinformatics, covering Unix, Docker, and data pipelines for practical skill development.

Centre of Biological Engineering, Portugal — *PH.D. Researcher*

10/2018 - 09/2022

Awarded the Academic Merit Scholarship (2020) for outstanding academic performance and successfully **authored and co-authored 5 peer-reviewed articles** in top scientific journals.

Created a **computational framework** for modeling and simulating biological systems.

End-to-end data pipelines on-prem infrastructure **using ETL tools** to process raw biological data into a data warehouse.

Implemented a full-stack web application and reporting tools for data visualization and analysis.

ITQB-UNL, Portugal — *Scientific Researcher*

02/2018 - 09/2018

Responsible for a novel **microscopy imaging analysis technique** to process complex biological data (7).

Chr. Hansen A/S, Denmark — *Intern*

02/2017 - 07/2017

Developed a novel computational model for **modeling and simulating** biological systems in **a world-leading bioscience company**.

EDUCATION

Universidade do Minho, Portugal — *PH.D. in Biomedical Engineering*

2018 - 2022

Universidade do Minho, Portugal — *Master's Degree in Bioinformatics*

2015 - 2017

Universidade do Minho, Portugal — *Bachelor's Degree in Applied Biology*

2012 - 2015

LANGUAGES

Portuguese - Native (C2)

English - Proficient (C1)