

Alejandro Fontal, PhD

DATA SCIENCE | MACHINE LEARNING | SOFTWARE DEVELOPMENT

Linked  /AIFontal  GitHub: @AIFontal  ORCID: 0000-0003-1138-2158

✉ alejandro.fontal.92@gmail.com 📞 (+34) 623 107 939 🌐 alfontal.dev 📍 Barcelona, Spain

Computational scientist and research software engineer working across computational biology, data science, applied machine learning, and scientific software. I build robust analytical pipelines, health data workflows, and research tools for real-world biological and public health problems.

WORKING EXPERIENCE

Climate & Health Program @ ISGlobal

Postdoctoral Researcher | Data & Software Lead

📅 Oct. 2024 - Present 📍 Barcelona, ES

- Lead software and data engineering for the group, building internal packages, web tools, and GitHub-based workflows that support reproducible research and team-wide development.
- Build reproducible pipelines for long-read aerobiome metagenomics and link results to health outcomes.
- Repurposed a laser-based pollen detector for bacterial discrimination, combining fluorescence, light scattering, and supervised ML for near-real-time identification.

PhD Fellow | Early Stage Researcher (ESR)

📅 Oct. 2019 - Oct. 2024

Marie Skłodowska-Curie Actions ESR within HELICAL ITN.

- Developed time-series methods to quantify environmental signals in epidemics, including transient and lagged associations, multi-scale correlations, and phase-shift analyses.
- Integrated nationwide case data with climate reanalysis, remote sensing, GIS layers, and air-mass back-trajectories to model drivers of Kawasaki Disease, COVID-19, and influenza.
- Coordinated field sampling campaigns in Japan to characterize the biological and chemical composition of air masses.

Protein Engineering @ DuPont Industrial Biosciences

Data Scientist

📅 Sep. 2018 - Aug. 2019 📍 Leiden, NL

- Pioneered R&D efforts to integrate data-driven solutions into the protein engineering workflow, with a strong focus on ML applications.
- Built and maintained reproducible bioinformatics and data pipelines.
- Partnered with wet-lab teams to design experiments and close the model-experiment loop, accelerating workflow automation.

Graduate Trainee in Data Science

📅 Mar. 2018 - Sep. 2018

MSc Internship project: ML for Empirical Protein Design

- Built, trained and benchmarked deep learning models to predict enzyme performance from sequence and structure.

Educational Staff Development @ Wageningen University

Assistant in MOOCs Development

📅 Sep. 2016 - Dec. 2017 📍 Wageningen, NL

- Developed assignments and content for 10+ edX MOOCs.
- Provided technical support to students and educational instructors.

EDUCATION

PhD in Biotechnology

📅 2019-2024

Universitat de Barcelona

- PhD focused on environmental determinants of disease onset, spatiotemporal modeling, and long-read aerobiome metagenomics.

MSc in Bioinformatics

📅 2016-2018

Wageningen University & Research

- Minor in Data Science, cum laude distinction.
- Thesis on interpretable deep learning for protein subcellular location prediction

BSc in Biotechnology

📅 2011-2015

Universitat de Barcelona

- Thesis at the VHIR Bioinformatics Unit on meta-analysis of transcriptomics tools.

Languages: Spanish & Catalan (native), English (C2)

DEVELOPMENT SKILLS

Python: Expert (9+ years); primary language. Package development, testing and maintenance. End-to-end DS/ML work: data wrangling, visualization, statistical modeling, machine learning, and app deployment (pandas, polars, statsmodels, scikit-learn, tensorflow, pytorch, plotnine, fastapi).

R: Proficient with tidyverse and Bioconductor; strong R/Python interoperability in workflows.

Shell and CLI: Daily Linux user; scripting and HPC workflows (SLURM, Nextflow).

Git & DevOps: Collaborative development, code review, and automated testing/deployment (GH Actions, GitLab CI/CD).

Web / APIs: Working knowledge of modern web development; REST services, deployment and hosting.

AI/Agentic Development: Practical experience with agentic coding and automation workflows: tool orchestration, structured outputs and multi-step codebase tasks (Codex CLI, Claude Code).

Systems/Infrastructure: Linux-based deployment for data applications, including containerized services, automation, and secure remote access.