THÉOPHILE L. MOUTON

SENIOR DATA SCIENTIST



CONTACT

theophile.mouton92@gmail.com

Montpellier, France

www.theophile-mouton.com

Théophile L. Mouton

TheophileMt92

SKILLS

- R (tidyverse, ggplot2)
- Python
- Quarto reports and slides
- Interactive web apps (Shiny)
- Statistics, modelling and ML
- Geospatial analysis (sf, terra, leaflet)
- LLMs (Claude API, extraction)
- Web (Astro, HTML/CSS, JS)
- Git/GitHub, Positron
- Client communication
- Presentations and training delivery
- Project and team coordination
- Stakeholder engagement

EDUCATION

2020 - 2023

University Claude Bernard Lyon I

PhD in Macroecology

2014 - 2016

University of Pau and Pays de l'Adour

MSc in Ecology

PROFILE

Data scientist with 10+ years of commercial and research data science in R and Python, building dashboards, reproducible reporting systems, and analytical pipelines for research institutions, NGOs, and government clients. Delivers projects end to end: from client requirements to handover. Combines PhD-level statistics and machine learning with hands-on engineering in R and Python on a git-based reproducible stack. Delivers training; quick to adopt new tools (Positron, LLMs). Trilingual (English, French, Spanish).

WORK EXPERIENCE

DataSphere Analytics

Founder and Data Consultant — 2024 - PRESENT

- Design, build and ship interactive Shiny applications end to end for research institutions, NGOs and government clients, from scoping through deployment and handover.
- Built FINS and SAFE (Pimiento Research Group, University of Zurich) and DateBack (IRD/CNRS): Shiny applications with data validation, interactive mapping, filtering and export.
- Delivered an automated Quarto reporting pipeline for the Government of Vanuatu, integrating multi-sector data into a parameterised damage assessment driven by external configuration files.
- Built and published a reusable branded Quarto report template, documented and deployed on GitHub Pages.
- Authored R Graph Gallery tutorials; led a Global Fishing Watch Data Lab working group (Rome, 2026): live Quarto report and auto-updating slides.

IRD (Institut de Recherche pour le Développement)

Data scientist / Postdoctoral researcher — Jan 2026 - Jul 2026

- Architected an AI-assisted extraction pipeline in Python using the Anthropic Claude API, automatically pulling 27 structured fields from a 3,000+ document corpus.
- Built a validation framework to verify extraction accuracy at scale: inter-rater agreement statistics (Gwet's AC1), cluster bootstrap confidence intervals, and McNemar's test on a stratified sample.
- Engineered a reproducible, fully documented screening and analysis system, delivered as a public GitHub repository structured for reuse and handover by the client team.
- Scoped and led the project independently, from defining the methodology through to delivery.

LANGUAGES

- English (Fluent, British)
- French (Mother tongue)
- Spanish (Fluent)

PUBLICATIONS

- Authored 12 peer-reviewed journal articles, including 5 as first author in leading scientific journals including *Nature Communications* and *Science Advances* (invited to resubmit).
- Contributed to the writing of several scientific reports for IUCN, and Government agencies.

SIDE PROJECT

- Multilingual international events directory: designed and deployed a static site in Astro (front-end development, structured data, UX, deployment). view the project

WORK EXPERIENCE

University of Zurich

Data scientist / Postdoctoral researcher — Jan 2025 - Dec 2025

- Led a global modelling project applying a Python reinforcement-learning algorithm (CAPTAIN) to optimise prioritisation across ~1,000 entities under multiple weighted criteria.
- Built geospatial data pipelines over large-scale satellite datasets: a random forest model predicting values for unobserved cells, and an aggregation workflow across 89 raw data files normalised by satellite overpass frequency.
- Produced a manuscript in revision at Science Advances, with fully reproducible analysis code and figures generated directly from the pipeline.
- Archived the analysis codebase on Zenodo with a citable DOI.

IUCN (International Union for Conservation of Nature)

Postdoctoral researcher — 2023 - 2024

- Coordinated a distributed team of 60+ contributors across international institutions, managing contributions, review cycles and delivery of technical documentation to deadline.
- Designed and delivered two training workshops for 250+ participants, in Greece and South Africa.
- Applied text mining, statistical modelling and spatial analysis in R, resulting in a peer-reviewed publication.

University of Montpellier / Université Claude Bernard Lyon I

PhD in Macroecology (Statistical modelling) — 2020 - 2023

- Compiled, cleaned and modelled a 30-year dataset spanning 70 sites and 100+ variables, building the full processing and analysis workflow in R.
- Assembled and managed a team of 15 scientists across three continents; led three first-author publications.

REFERENCES

Yan Holtz
Dataviz Universe
Contact details on request

Prof. Fabien Leprieur
University of Montpellier
Contact details on request