



KARASIEWICZ Stéphane

Senior Data Scientist

27 September 1989

French/British

3 rue du Verdun, 62230 Outreau, France

(+33) 0671675613

KARASIEWICZ Stéphane

KARASIEWICZ Stéphane

skaraz.science@gmail.com

About me

Senior Data Scientist with 7+ years delivering ML solutions from research to production. Specialized in environmental analytics, time-series modeling, and scalable pipelines. Published author of 2 CRAN packages and 6 peer-reviewed publications. Expert in Python/R and translating complex problems into actionable insights.

Skills

- Python (Expert)
- R (Expert)
- SQL
- Bash/Shell scripting
- scikit-learn
- TensorFlow
- XGBoost
- PyTorch
- Supervised/unsupervised learning
- Time-series forecasting
- Clustering
- Pandas
- NumPy
- Dask
- ETL pipelines
- Data lake architecture
- NetCDF/Zarr

Experiences

- 02/24-01/26 Senior Data Scientist Home based
 Raincoat LLC

 - ML infrastructure for peril risk assessment
 - Engineered end-to-end ML pipeline
 - Deployed 5 interactive dashboards with country level maps
 - Automated data management system for 2TB+ satellite data
 - Analysis reports translating model outputs to insights
 - Tools: Python, scikit-learn, Docker, PostgreSQL
- 07/23-02/24 Freelance Data Scientist Home based
 Self-employed - Delivered 12+ ML projects

 - Customer churn model (XGBoost)
 - Price forecasting (LSTM)
 - Custom analytics solutions with Python and R
 - Full ML lifecycle: preprocessing, feature engineering, tuning, deployment
 - Tools: Python, scikit-learn, TensorFlow, SQL
- 08/22-07/23 Data Scientist Certification OpenClassrooms (Online)
 400+ hour program: Python, ML deployment, cloud

 - 10 projects: regression, classification, clustering, NLP, CV
 - Deployed 3 ML models via REST APIs with CI/CD
 - Recommendation system handling 100K+ interactions
 - Interactive dashboards with Streamlit and Dash
 - Projects: Segmentation — CNN — NLP — Docker/AWS
- 10/21-03/22 Research Engineer IFREMER, Boulogne-sur-mer, France
 R&D data science for marine research

 - Automated 500GB+ data extraction and preprocessing
 - Reduced manual effort by 50 hours/month
 - Optimized R scripts: 70% computation time reduction
 - Automated Rmarkdown reporting (12 annual reports)
 - Mentored 4 researchers: R, Git, reproducible research
 - Tools: R, Rmarkdown, SQL, Git, ggplot2
- 09/20-09/21 Postdoctoral Researcher IFREMER, Boulogne-sur-mer, France
 Algal bloom detection with satellite imagery & ML

 - Bloom detection algorithm: 10+ years satellite data
 - Processed 50,000+ km² coastal waters (MODIS, Sentinel)
 - Published BDAalgo R package (CRAN)
 - Multivariate analysis: 8 environmental drivers identified
 - 2 peer-reviewed publications, 1 dataset
 - Tools: R, satellite sensing, spatial stats, ML
- 09/19-09/20 Freelance Data Scientist Freelance, France
 Data analysis & translation for 15+ clients

 - Statistical consulting and data visualization
 - Custom R tutorials for Excel-to-R transitions
 - Script optimization: 40-60% performance gains
 - Translated 200+ pages technical docs (ENG-FR)
 - Tools: R, Python, statistics, visualization
- 02/18-07/19 Postdoctoral Researcher IFREMER, Centre BREST, France
 Marine ecology with satellite data & stats

 - Analyzed 15+ years satellite data (1TB+)
 - Phytoplankton community dynamics study
 - Multivariate techniques (PCA, niche modeling)
 - Identified niches of 20+ harmful algae species
 - Mentored 2 Master's students
 - 2 first-author papers, 1 dataset
 - Tools: R, satellite analysis, multivariate stats










Skills (cont.)

- Satellite data processing
- Docker
- Git
- CI/CD
- Linux system admin
- Dash
- Streamlit
- Plotly
- ggplot2
- Shiny
- API development
- Geospatial analysis
- Satellite imagery processing
- Environmental modeling
- Statistical analysis (multivariate, PCA, spatial-temporal)

Education

- 2014-2017 Ph.D. in Marine Ecology University Lille, France
Phytoplankton evolution under global changes
- Novel statistical framework (WitOMI) for niche shifts
 - Analyzed 20+ years monitoring data (10M+ records)
 - Advanced multivariate techniques
 - 3 first-author papers, 2 CRAN packages
 - Teaching: Biostatistics (87h), Ecosystemic Ecology (16h), Marine sampling (24h)
 - Focus: Multivariate analysis, niche theory, biodiversity, time-series
- 2012-2014 M.Sc. Marine Ecology Stockholm University, Sweden
Marine community ecology, ecosystem modeling, statistical analysis
- 2010-2012 B.Sc. (Open University) Open University, England
Ecosystem Science, Oceanography, Evolution, Environmental Change
- 2008-2010 Foundation Degree in Animal Science Bicton College, England
Animal Welfare, Behaviour and Management
- 2007 BAC S Option SVT Lycée Marc Chagall, Reims, France
mention européenne anglais

Open Source & Publications

-  BDAIgo - CRAN Package. Oct 2023 — 3,000+ downloads — 12+ citations. R package for bloom detection in time-series. Extracts 22 phenological variables. <https://cran.r-project.org/web/packages/BDAIgo/>
-  Subniche - CRAN Package. Jul 2022 — 5,000+ downloads — 15+ citations. Statistical package for ecological niche modeling using OMI framework. <https://cran.r-project.org/web/packages/subniche/>
-  WitOMI Tutorial: Documentation for Within Outlying Mean Indexes. <https://github.com/KarasiewiczStephan/WitOMI>
-  Karasiewicz S., Lefebvre A. (2022). Environmental impact on harmful species. *J. Marine Science & Eng.*, 10(2):174. <https://doi.org/10.3390/jmse10020174>
-  Karasiewicz S., Lefebvre A. (2021). Environmental impact on harmful species: Dataset. SEANOE. <https://doi.org/10.17882/85097>
-  Karasiewicz S., Chapelle A., Bacher C., Soudant D. (2020). Harmful algae niche responses. *Harmful Algae*, 98:101785. <https://doi.org/10.1016/j.hal.2020.101785>
-  Karasiewicz S., Chapelle A., Bacher C., Soudant D. (2019). Harmful algae niche responses: Dataset. SEANOE. <https://doi.org/10.17882/69703>
-  Karasiewicz S., et al. (2018). Realized niche analysis of phytoplankton communities. *Harmful Algae*, 72:16-30. <https://doi.org/10.1016/j.hal.2017.12.005>
-  Karasiewicz S., Dolédec S., Lefebvre S. (2017). Within outlying mean indexes. *PeerJ*, 5:e3364. <https://doi.org/10.7717/peerj.3364>

Languages

French: Native — English: Native — Spanish: Basic