

UDIT ASOPA

+358-403515447

Linkedin-uditasopa

uditasopa.netlify.app

udit.asopa@gmail.com

Espoo, Finland

GitHub: udit-asopa

PROFESSIONAL SUMMARY

Research Software Engineer and Geospatial Data Scientist with 5+ years of experience developing reproducible Python workflows, machine learning pipelines, and large-scale environmental data processing systems. Strong background in Earth Observation, scientific computing, and AI-driven analytics, with hands-on experience building validated and maintainable workflows in Linux environments. Skilled in translating complex scientific and operational problems into reliable, scalable software solutions. Passionate about applying scientific software engineering, machine learning, and scalable computing to climate and Earth system digital twins.

WORK EXPERIENCE

Geospatial Data Scientist @ Brockmann Geomatics AB, Espoo Finland

Jan 2026 - Present

- Design and maintain reproducible Python- and SQL-based workflows for multi-source environmental and infrastructure datasets, with built-in validation, monitoring, and traceability
- Perform large-scale time-series analysis, machine learning on Earth observation, geospatial datasets to support scientific and operational applications
- Contribute to structured software engineering practices, including experiment tracking, schema consistency, and documentation

Stack: Python (pandas, numpy, xarray, MLflow), SQL, GeoPandas, Git, Linux, Jupyter

Geospatial Data Scientist (official title: SAR Remote Sensing Engineer) @ ICEYE Oy, Espoo Finland

Dec 2021 - Jan 2026

- Designed and operated end-to-end Python workflows for large-scale satellite imagery and time-series datasets, covering ingestion, preprocessing, feature engineering, and validation
- Developed and evaluated machine learning models for classification, anomaly detection, and predictive analytics on Earth observation datasets
- Built automated data quality and monitoring frameworks, improving reproducibility and reducing time-to-insight by approximately 40%
- Conducted rapid analytical investigations to diagnose data and model issues, identify root causes, and quantify operational impact
- Collaborated with product managers, engineers, domain experts to translate scientific and operational requirements into production-ready solutions
- Worked extensively in Linux environments using Git and automated pipelines for reliable delivery

Stack: Python (pandas, numpy, xarray, scikit-learn), SQL, GDAL, Rasterio, Git, Linux, Jupyter

Geospatial Data Researcher @ Delft University of Technology, Delft Netherlands

Apr 2020 - Mar 2021

- Built reproducible Python workflows for satellite-based infrastructure monitoring and deformation analysis
- Conducted experiment-driven statistical analysis and validation on multi-temporal Earth observation datasets
- Documented methodologies and findings to support scientific interpretation and decision-making

Stack: Python, Jupyter, Git, QGIS, MATLAB

Junior Research Engineer @ Indian Institute of Technology, Mumbai, India

Sept 2019 - Feb 2020

- Supported data preprocessing, validation, and reporting for environmental monitoring applications **Stack:** Python, ArcGIS, QGIS

PROFESSIONAL SKILLS

- **Scientific Software Engineering:** Reproducible workflows, modular pipelines, automated validation, and maintainable research code
- **Earth Observation & Climate Analytics:** Satellite imagery processing, environmental time-series analysis, and geospatial modeling
- **Machine Learning & AI:** Classification, anomaly detection, feature engineering, model evaluation, and applied generative AI
- **Workflow Automation:** Python and Bash scripting, orchestration concepts, and scalable analytical pipelines
- **Data Engineering & Quality:** ETL workflows, schema design, monitoring, traceability, and data lineage
- **Containers & Dev Practices:** Docker familiarity, containerized workflows, Git-based development, CI/CD awareness, and reproducibility, best practices

- **High-Performance Computing:** Exposure to GPU-based processing and large-scale workflows through CSC training on LUMI

TECHNICAL SKILLS

- **Programming & Scripting:** Python, SQL, Bash, Jupyter Notebooks
- **Scientific Python Stack:** pandas, numpy, xarray, scikit-learn, MLflow
- **Geospatial & Scientific Computing:** GeoPandas, GDAL, Rasterio, QGIS
- **Machine Learning & AI:** Predictive modeling, anomaly detection, RAG pipelines, FAISS, LangChain, LlamaIndex
- **Workflow & Data Engineering:** ETL pipelines, validation frameworks, monitoring, reproducible workflows
- **Containers & Systems:** Docker (familiarity), Git, Linux
- **Cloud & Platforms:** AWS familiarity, Azure analytics, Databricks SQL

EDUCATION

Education:

- **M. Eng.:** Remote Sensing & GIS (*Indian Institute of Remote Sensing, 2019*)
- **Post Graduate Diploma:** Embedded Systems & Informatics (*Center for Development of Advanced Computing, 2017*)
- **B. Eng.:** Electronics, IoT, Embedded (*Rajasthan Technical University, 2016*)

Languages: English (Native), Finnish (A1, upskilling), Swedish (A2, upskilling)

SELECTED PROJECTS

- **Disaster GeoRAG (Hackathon, 2nd place):** Retrieval-augmented VLM pipeline for EO imagery triage, combining geospatial metadata, vector search, and structured model outputs ([LinkedIn Link](#))
- **Multimodal OCR CLI:** Privacy-aware OCR and document parsing pipeline with configurable local vs cloud inference ([Github link](#))
- **Customer Churn Prediction (BCG X Job Simulation):** Built and optimized a machine learning model to predict customer churn, delivering actionable recommendations for customer retention ([Github Link](#))
- **From Browsing to Booking: A Predictive Model for Airline Customers:** Built and evaluated a machine learning model to predict customer booking propensity and identify key drivers influencing purchasing behavior ([Github Link](#))
- **Data Analytics Simulation (Deloitte):** Developed dashboards and business recommendations based on exploratory data analysis

CERTIFICATIONS

- Practical Machine Learning with Spatial Data -- [CSC – IT Center for Science](#), 2025
- [McKinsey Forward](#) -- [McKinsey & Company](#), 2025
- SQL Analytics and BI on Databricks -- [Databricks](#), 2026
- AI Upskilling Technical Foundations -- [Qualcomm](#), 2026
- Modern Data Warehouse Analytics in Microsoft Azure -- [Microsoft](#), 2025
- Data Science Job Simulation -- [BCG X](#), 2025
- Data Analytics Job Simulation -- [Deloitte](#), 2025
- Project Management: Beyond planning and control -- [Politecnico Milano](#), 2025