JOSE SABATER

Senior Software Engineer

+44 7587 193575
+34 639 762351
jsm@jsabater.info
Edinburgh, UK
Almería, Spain

Senior Software Engineer with a long track in complex problem solving, data engineering, and data analysis and visualisation. Seventeen years programming in Python and committed to stay up to date and follow best practices. Strongly interested in finding new and better ways to obtain knowledge and insight from data.

SKILLS

Professional Skills

Problem solving
Project management
Effective communication
Team coordination
Time management
Collaboration
Decision-making

Technical Skills

Data analysis
Python (17 years)
Data engineering
Data visualisation
Cloud and distributed infrastructures
DevOps workflow
Agile development (certified SAFe practitioner)
Machine learning
Databases and SQL
Reproducibility

EDUCATION

PhD in Astrophysics
University of Granada (Spain), 2009

Degree in Physics (Master's level)
University of Granada (Spain), 1997 - 2003

WORK EXPERIENCE

Senior Software Engineer
UK Astronomy & Technology Centre | since *Sep. 2020*

Development of observatory and simulation tools for the Square Kilometre Array (SKA), a data-intensive infrastructure that will produce about 1 Exabyte of data per year. Cloud native architecture in an Agile environment following best practices.

- Led the architectural definition of the Sensitivity Calculator (library and application) and the future simulations framework.
- Definition and implementation of the REST API of the Sensitivity Calculator application (Open API, JSON Schema, Flask, swagger).
- Development of the DevOps workflow for the Sensitivity Calculator (Gitlab CI, Docker, Kubernetes, Helm).
- Selection and implementation of the end to end testing framework for the Sensitivity Calculator (Cypress).
- Development of front-end user interfaces, logic, and tests for the Sensitivity Calculator and the Observation Design Tool (Typescript, Angular).
- Participation in the design of the Observatory Science Operations Data Archive.

Core developer of the Atacama Large Millimetre Array (ALMA) data calibration pipeline.

- Refactoring of the pipeline legacy code to follow best practices and avoid future errors. Implementation of tests.
- Understand the requirements requested by stakeholders and transform them in new features for the pipeline.
- Participation in the definition of the new pipeline architecture.

Supervision of PhD project: *Automatic identification and classification of radio sources*.

- Development of machine learning and artificial intelligence pipelines to allow the classification of data (a task that is impossible to perform manually given the amount of data). Deep learning, self organised maps, random forests, boosted trees (pytorch, Keras, xgboost, scikits-learn, SnakeMake, TPOT).

JOSE SABATER

Senior Software Engineer

LANGUAGES

English

Spanish

CODE & SOCIAL

Github
https://github.com/nudomarinero

Gitlab
https://gitlab.com/nudomarinero

OTHER

Honorary fellow of the University of Edinburgh

Tutor
Django Girls Edinburgh
Fourier Analysis
Statistics

WORK EXPERIENCE (continued)

Post-Doctoral Research Associate
University of Edinburgh (UK) | *Oct. 2010 – Aug. 2020*

Scientific research and commissioning of state-of-the-art data-intensive infrastructures for radio-astronomy. I moved into leading roles in the technical developments required to deliver the science.

- Development of open source data analysis and pipeline tools. For example, the Python package wquantiles that has been downloaded more than 400K times (astropy, necpp, etc).
- Principal Investigator and manager of the AWS-SKA *Astrocompute in the Cloud* grant to calibrate radio-astronomy data (2 PB) in the AWS cloud (EC2, S3, Route 53, Ansible).
- Research of the use of distributed infrastructures for the calibration of massive astronomy data in collaboration with AWS, the EGI federated cloud and the Rutherford Appleton Laboratory (Ceph, OpenStack).
- Curator of the Amazon Web Services Open Data LOFAR ELAIS-N1 public dataset: https://registry.opendata.aws/lofar-elais-n1/
- Led the effort that produced the deepest image of the sky at low frequencies. Coordination of a diverse team of researchers and developers across 4 continents. Core role in the efforts to deliver the new science based on LOFAR surveys.
- Work presented in 75 refereed research articles, more than 60 conferences, and technical documents.
- Design and direction of 4 Master's projects, 5 senior honours projects, and 3 Career Development Summer Scholarship projects. Assisted in the supervision of 3 PhD projects.
- Development of reproducible data flows for the calibration, analysis and visualisation of research data (numpy, Jupyter notebooks, matplotlib, pandas, Singularity).
- Some of the developed data flows made heavy use of complex SQL queries that needed to be efficient (PostgreSQL and PL/pgSQL including server management, MySQL, SQL Server).

PhD and Post-Doctoral Researcher
IAA - CSIC (Spain) | *Jan. 2004 – Sep. 2010*

Scientific work: Study of the effect of galactic environment and interactions on the feeding of central supermassive black holes.

- Development of reproducible scientific workflows and participation in the implementation of data standards.
- Principal Investigator and developer of a grant to adapt astrophysical software to the distributed GRID computing infrastructure.