Pablo Rosado

Senior Data Scientist & Researcher

mail@pablorosado.com
08011 Barcelona +34 654273008

Currently doing AI for the EU.
10 years' experience in Data Science in academia and in industry.
Previously a researcher in astrophysics, member of Nobel Prize-winning collaboration.

+Programming (Python/R)
+Data analysis (Pandas/NumPy)
+Cloud computing (AWS/Azure/GCP)
+ETL & Databases (SQL/NoSQL)
+Statistics (frequentist/Bayesian)

+Machine Learning/Deep Learning (Sklearn/Keras)
+Data visualisation (Matplotlib/Plotly)
+Mathematical modelling (Mathematica)
+Research & Lecturing (Physics/AI/Data Science)
+Science communication (Esp/Eng/Deu)

LANGUAGES

Spanish (native)   **English** (fluent)   **German** (fluent)   **Catalan** (intermediate)   **Italian** (basic)

QUALIFICATIONS

- 2010-2013: Hannover (Germany). **PhD in Astrophysics** at the Max Planck Institute for Gravitational Physics (Albert Einstein Institute) and the University of Hannover. Thesis topic: Gravitational waves data analysis.

- 2008-2009: Münster (Germany). **Master of Science** at the Institute for Nuclear Physics, WWU Münster.

- 2005-2009: Seville (Spain). **Double degree in Physics** from the Universities of Seville and Münster. Graduated with **highest honours** (*20 matrículas de honor*). Average mark: 9.25 / 10.

- 2008-2013: I attended 12 summer schools and lecture series in physics and astronomy (in Australia, France, Germany, Poland, USA, UK, and Spain), with a special emphasis on **statistics and data analysis**.

DATA SCIENCE IN INDUSTRY

- 2021-present, Remote: Data Scientist at the European Union Intellectual Property Office.

  ➢ **Graph Machine Learning, Natural Language Processing**: Link prediction in a network of short texts.

  ➢ **Statistical analyses**: Exploring datasets of images and texts, extracting insights and checking for biases.

  ➢ **End-to-end Machine Learning**: Development of production-ready services and model selection pipelines.

- Summer 2019, London/Coventry (UK): Tech Lead at Alan Turing Institute & U. Warwick – Data Science for Social Good.

  ➢ **Lecturing Data Science & Machine Learning**: Teaching outstanding students, in collaboration with Imperial College & Warwick Business Schools, University of Chicago, Accenture & Microsoft.

  ➢ **Delivering fully functional data products**: Two projects which I led during the school:
    - Detecting fraudulent public procurements for the Government of Paraguay. GitHub repos.
    - Classifying medical reviews for Cochrane. GitHub repos.

  ➢ **Technical Mentoring**: Guiding fast-paced Data Science projects with a positive social impact, managing architecture (Azure), databases (Postgres), code reviewing and ensuring good practices with sensitive data.

  ➢ **Ethical and interpretable model selection**: Ensuring interpretability and reducing unethical biases in models.

- 2018-2021, Barcelona (Spain): Data Scientist at Holaluz – Renewable energy retail. My GitHub.

  ➢ **Time Series Forecasting**: Prediction of energy consumption, outlier detection.

  ➢ **Churn and Fraud Detection**: Classification and anomaly detection ensuring model interpretability.

  ➢ **Short Texts Classification**: Real-time categorisation and prioritisation of incoming emails.

- 2017-2018, Barcelona (Spain): Data scientist at GFT Technologies – Consultancy in banking.

  ➢ **Computer Vision**: Real-time object recognition and tracking in videos from a mobile phone camera.

  ➢ **Natural Language Processing**: Segmentation and classification of scanned documents, and entity extraction.

  ➢ **Presenting**: Showcasing our work at German and Spanish venues including Google, Deutsche Bank & Deutsche Börse.

RESEARCH & DATA SCIENCE IN ASTROPHYSICS

- 2016, Melbourne (Australia): Postdoctoral researcher at Monash Centre for Astrophysics, Monash University – Research and data science in theoretical and observational gravitational wave astrophysics.
  - ➢ **Bayesian Statistics**: I worked on a pioneer method to search for a stochastic background of millions of weak signals in noisy data as part of the LIGO collaboration, awarded the 2017 **Nobel Prize in Physics**.
- 2014-2016: Melbourne: Postdoctoral researcher at Centre for Astrophysics and Supercomputing, Swinburne University of Technology – Research and data science in radio astronomy.
  - ➢ **Time Series Analysis**: At the International Pulsar Timing Array I worked on the search for correlated deviations in arrival times of millisecond pulses in >10 years of time series of noisy data from radio telescopes.
  - ➢ **Mapping**: I worked on a real-time sky mapping tool, correlating data from >300 antennas of UTMOST telescope.
  - ➢ Other tasks: **Lecturing**, **mentoring** and evaluating PhD and Masters students. **Organising** weekly scientific colloquium, which involved planning schedules with speakers from around the world and chairing talks.
- 2014: Hannover: Postdoctoral researcher at the Max Planck Institute for Gravitational Physics – Research and data science in theoretical and observational gravitational wave astrophysics.
  - ➢ **Machine learning**, **Bayesian statistics** & **SQL databases**: I designed a pipeline to study patterns in millions of computer-simulated galaxies, and find similar patterns in a database of real galaxies.

DATA SCIENCE & DATA ENGINEERING SKILLS

- Software development:
  - ➢ **Python** (proficient), **Git** (proficient), **Docker** (and basic Kubernetes), **Unix Shell** scripting, **R, Jenkins**.
  - ➢ Development, deployment, testing, maintenance and documentation of data pipelines, Python packages, dashboards, and flask/FastAPI services, both on-premise and in the Cloud (**Google Cloud**, **Azure**, **AWS**).
- Artificial Intelligence:
  - ➢ Extensive knowledge of different **Machine Learning** and **Deep Learning** algorithms and tools, especially in NLP (Pandas, NumPy, Sklearn, TensorFlow, Keras, Gensim, Spacy, Shap, Jupyter, Dash).
  - ➢ Certified Stanford University course on **Machine Learning** & **Deep Learning Specialization** (Coursera).
- **Database** and **Big Data** tools: **PostgreSQL**, **MongoDB**, basic knowledge of Spark.
- **Bayesian** and **frequentist statistics** (used in most of my research publications).
- **Agile** software development methodologies and project management tools (Trello, Jira, Kanbanize, Confluence).
- Other software: **Mathematica**, **Matlab**, **Excel**, **WordPress**.

ADDITIONAL SKILLS

- **Expertise and reliability**.

  I have reviewed numerous scientific articles of the American Physical Society, and a USA research grant worth approx. $200,000 from the National Science Foundation. I have contributed to more than 30 scientific articles in **high impact journals**, *accumulating over 2000 citations to date.* *My Google Scholar profile*.

- **Science communication and public outreach**.

  I have given **talks** *worldwide to hundreds of academics and to the* *general public*. *I also have experience speaking to the media via TV, radio and in journals in Spain, Germany and Australia.* *My website*.

- **Music and video production**.

  I studied at Seville Professional Conservatory and write *my own music*. *I have arranged, co-produced and mixed* *an album* *and performed as a piano soloist and in bands. I have directed and edited documentaries, short films and* *music videos*, **winning international awards**, *and worked as a videographer for my university.*

- **Values**.

  I am an ardent **supporter of human rights, environmental protection, veganism and effective altruism**. *I have organised humanitarian events and volunteered as a videographer for an NGO.* *More about me*.