<p style="text-align:center">PhD offer in synthetic data validation. Preliminar call</p>

Doctoral Program: Statistics and Operations Research, UPC
Catedras ENIA: ENIA Chair Sustainable AI (Provisional resolution)

Deadline Application: 10th February 2024 at 8pm

Initial temptative date: 1st March 2024
Duration: 34 months (possible extension)
Type of contract: PSR Group 3
Gross salary: 1.340 €/month

Project Title: Development and Evaluation of Metrics for Assessing Synthetic Data Quality

Project Description:

The generation of synthetic data has emerged as a valuable technique for preserving the privacy of original data and expanding the size of training samples in scenarios where obtaining real data is costly, challenging, or infeasible. Validating the quality and the accuracy of synthetic data is a research area at the intersection of AI and statistics that is still ripe for exploration.

This PhD opportunity aims to define a set of metrics for evaluating the quality of synthetic data. These metrics will be selected based on their ability to validate key statistical properties, such as right parametric distribution and comparable variance and correlations among variables. This thesis aims to develop a framework to compare synthetic data with real-world data, leveraging statistical techniques like hypothesis testing, clustering, and regression analysis to quantify the degree of similarity.

This new methodology will be a tool for researchers and professionals, to ensure that the quality of synthetic data generated is sufficient for use in predictive models and further statistical analysis. The goal is to enhance the precision and reliability of decision-making processes.

Requirements:
- Master's degree in a relevant field (statistics , computer science, artificial intelligence, data science, or similar).
- Strong background in statistical inference and methods and in machine learning.
- Proficiency in programming skills in R.
- Interest in interdisciplinary research at the intersection of AI and statistics.
- Proficiency in English

Prior research experience in related domains will be considered an asset

Benefits
- Engage in cutting-edge AI and statistical research.
- Contribute to innovative solutions in synthetic data validation.
- Integration into the GRBIO group (https://grbio.upc.edu/en)

The PhD will be based in Barcelona, with plenty of opportunities for placements in other research centres in Spain and Europe

- **How to Apply:** Interested candidates should submit the following documents to Guadalupe Gómez Melis (lupe.gomez@upc.edu and cc to ignacio.perez.blasco@upc.edu)

Documentation

- CV (free format)
- Academic transcripts
- Relevant certificates
- Research interests and motivation for applying
- Contact information for two academic references
- Other supporting materials such as reference letters