

JOSÉ MATOS

Coimbra, Portugal | josemnmatos@gmail.com | [josemnmatos.github.io](https://github.com/josemnmatos) | [linkedin/josemnmatos](https://www.linkedin.com/in/josemnmatos) | [github/josemnmatos](https://github.com/josemnmatos)

RESEARCH INTERESTS

LLM Reasoning and Trustworthiness, NLP & Data Curation for Low-Resource Environments, LLM Evaluation

EDUCATION

University of Coimbra, Portugal

Master's in Informatics Engineering – Artificial Intelligence

Expected July 2026

- **Thesis:** *"Adapting an LLM for Explicit Reasoning in Portuguese"* – Adapting an LLM for verifiable scientific reasoning over European Portuguese documents. Comprises the development of synthetic data generation, model fine-tuning, and evaluation techniques for low-resource languages. Supervised by Prof. Hugo Gonçalo Oliveira and Prof. Catarina Silva.
- **Coursework:** Artificial Intelligence (19/20), Advanced Machine Learning (18/20), Experiment Design & Methods (19/20)

Bachelor's in Informatics Engineering

Graduated Jan 2024

- Includes an exchange semester at the University of Zagreb, Croatia under the Erasmus+ program.

RESEARCH EXPERIENCE

Research Fellow

Apr 2025 – Present

CISUC/LASI - Center for Informatics and Systems of the University of Coimbra

Coimbra, Portugal

- Integrated two funded projects: **AMALIA** (RE-C05-i08 – building a large language model for Portuguese) and **NextGenAI: Centre for Responsible AI** (PRR C645008882-00000055);
- Designed and developed a framework capable of systematically characterizing reasoning patterns in LLMs. Applied it to study the stress caused by distillation on transferred reasoning capabilities between teacher and student models;
- Contributed to the European Portuguese NLP ecosystem, through the creation of specialized LLMs, large corpora, reasoning datasets, and evaluation baselines.

Publications

- **José Matos**, C. Silva, H. G. Oliveira. *"Cognitive Flow: An LLM-Automated Framework for Quantifying Reasoning Distillation"* **INLG 2025**.
- T. Z. Kuhn, **José Matos**, B. Neves, et al. *"CorEGe-PT: Compiling a Large Corpus of Academic Texts in Portuguese"* **LREC 2026 (accepted for publication)**.
- **José Matos**, C. Silva, H. G. Oliveira. *"Benchmarking RAG for Scientific Knowledge QA in European Portuguese"* **NSLP 2026 (accepted for publication)**.

Awards & Presentations

- Honorable mention for Best MSc Publication, *SPARK Awards 2025* by Feedzai & Centre for Responsible AI.
- Presented works at **INLG 2025** (Hanoi), **Responsible AI Forum 2025** (Lisbon), **LASI Annual Meeting 2025** (Coimbra).

INDUSTRY EXPERIENCE

Software Engineer

Oct 2023 – Sept 2024

iToWIN

Coimbra, Portugal

- Built a prototype for football live entity detection and tracking using Python and deep learning frameworks.
- Contributed to the development of a cross-platform club management system using Flutter, delivering across macOS, iOS, and Web.

Software Engineering Intern

July – Sept 2023

Instituto Pedro Nunes, Laboratory of Informatics and Systems

Coimbra, Portugal

TECHNICAL SKILLS

Languages: Python, Java, SQL

ML & NLP: PyTorch, HF Transformers, scikit-learn, LLM fine-tuning (LoRA/QLoRA), RAG pipelines

Data & Analysis: Pandas, NumPy, Experimental Design

Tools: Git, Docker, FastAPI, Linux

LEADERSHIP & VOLUNTEERING

IT Team Lead

Sept 2024 – Mar 2025

20th National Biomedical Engineering Students' Meeting

- Led planning and development of the full IT system for a 4-day national conference, supporting hundreds of concurrent users and processing hundreds of thousands of requests.

LANGUAGES

Portuguese – Native

English – C1