

Job offer - Computer Science Department - Universitat Politècnica de Catalunya

Software Engineer - Data Curation

We are looking for a Software Engineer to design and develop a data management tool that will assist our Linguistic Team to create, curate, and maintain the training and test datasets that are used for fine tuning of neural models.

The Project:

Automation of business processes is a crucial step in the digital transition of organizations. This automation is typically carried out by expert engineers that design and implement the required software solutions, resulting in high costs for companies and organizations going through this transition.

We are aiming at a system that will rely on existing no-code/low-code platforms to allow non-technical users to create their own solutions by simply describing the process steps using natural language. The system exploits existing deep learning technologies, and fine-tunes them to understand and structure user utterances related to typical actions in a business process.—

Job Responsibilities:

- Design and develop a data curation tool that will allow the Linguistic Team to create, maintain, and exploit datasets to train language-understanding deep learning models.
- Obtain tool requirements from the expected users (Linguistic Team)
- Determining mechanisms for automatic data monitoring and validation
- Designing and setting data pipelines across different sources and systems
- Ensuring data quality, consistency and interoperability

We offer:

- Exciting work environment, in close collaboration with a start-up converting cutting edge research in everyday technology.
- Possibility of publishing relevant results
- Possibility of presenting the project as a Master Thesis at UPC.
- Possibility of remote working

About You

The successful candidate must have:

- Degree in Computer Science or Software Engineering
- Familiarity in software development in Linux environments
- Familiarity in GUI development
- Familiarity with Git versioning systems.
- High programming skills, preferably in Python and/or Clojure
- Capacity to interact with domain experts to gather their needs on data resources
- Effectiveness in the transfer of technical expertise to both colleagues and non-specialists
- Strong written and verbal communication skills
- Good team player
- Interested in carrying out R&D work in contact with a start-up

Desirable:

- Master courses / Master degree in Software Engineering
- Previous exposure to managing linguistic data
- Some familiarity with the field of Natural Language Processing

Contact: Lluís Padró (padro -at- cs.upc.edu)