Job Reference

520_23_LS_NLPBIA_RE2

Position

Research Engineer - NLP for Biomedical Information Analysis (RE2)

Data de tancament

Diumenge, 31 Desembre, 2023
Reference: 520_23_LS_NLPBIA_RE2
Job title: Research Engineer - NLP for Biomedical Information Analysis (RE2)

About BSC

The Barcelona Supercomputing Center - Centro Nacional de Supercomputación (BSC-CNS) is the leading supercomputing center in Spain. It houses MareNostrum, one of the most powerful supercomputers in Europe, was a founding and hosting member of the former European HPC infrastructure PRACE (Partnership for Advanced Computing in Europe), and is now hosting entity for EuroHPC JU, the Joint Undertaking that leads large-scale investments and HPC provision in Europe. The mission of BSC is to research, develop and manage information technologies in order to facilitate scientific progress. BSC combines HPC service provision and R&D into both computer and computational science (life, earth and engineering sciences) under one roof, and currently has over 900 staff from 55 countries.

Look at the BSC experience:
BSC-CNS YouTube Channel
Let's stay connected with BSC Folks!

We are particularly interested for this role in the strengths and lived experiences of women and underrepresented groups to help us avoid perpetuating biases and oversights in science and IT research.

Context And Mission

The Natural Language Processing for Biomedical Information Analysis (NLP4BIA) group at BSC is an internationally renowned research group working on the development of NLP, language technology, and text-mining solutions applied primarily to biomedical and clinical data. It is a highly interdisciplinary team, funded through competitive European and National projects requiring the implementation of natural language processing and advanced AI solutions making use of diverse technologies, including Transformers and recent advances in Large Language Models (LLM) to improve healthcare data analysis.
The NLP4BIA-BSC is looking for a Research Engineer with experience in Language Technologies and Deep Learning. The candidate will be involved in technical work related to international projects, being part of a team of researchers working on topics related to multilingual information extraction in the clinical field, including Named-Entity Recognition, Entity Linking and Language Modeling. The candidate will have the opportunity to advance the state of the art of cross-lingual biomedical NLP methods by working in a multidisciplinary environment alongside linguists, medical experts, and other engineers.

**Key Duties**

- Development of multilingual information extraction models in the biomedical field, including mention extraction and linking of terms to controlled terminologies. Pre-training of cross-lingual large language models for healthcare.
- Technical project coordination: Coordinate technical contributions from different partners in technological projects.
- Documentation and Reporting: Create technical reports and project documentation in both English and Spanish.
- Scientific writing: Collaborate in drafting technical research proposals and writing scientific papers.

**Requirements**

- **Education**
  - Ph.D. in Computer Science, Computational Linguistics, or a related field. Candidates with a minimum of a master's degree will be considered.

- **Essential Knowledge and Professional Experience**
  - Demonstrated experience in Natural Language Processing technologies (publication track record or projects in the field of Deep Learning/NLP)
  - Ability to design, implement, and evaluate existing NLP models.
  - Solid knowledge of information extraction techniques, including Named Entity Recognition (NER) or Entity Linking.
  - Experience in developing and training models using transformer architectures.
  - Practical experience with deep learning libraries (e.g. Pytorch, TensorFlow, Spacy, Transformers...)
  - Knowledge of deep learning methods for fine-tuning large language models, including contrastive learning, siamese networks, etc.
  - Advanced programming skills in Python.
  - Experience in software development resources (Git)

- **Additional Knowledge and Professional Experience**
Strong programming skills in at least one of the following languages: Java, Python, C++, Scala, R.
Development and integration of software tools.
Demonstrable use of source code repositories (e.g. Github, BitBucket, Sourceforge).
Experience with software tools such as Galaxy, GATE, UIMA, NLTK, cTakes.
Experience in/with cloud computing, docker, docker swarms and VMs.
Skills and knowledge related to data Warehouses, NoSQL (e.g. MongoDB, Elasticsearch), SQL (e.g. PostGreSQL, MySQL, Oracle, MS-SQL).
Experience with biomedical or clinical textual data (literature, EHRs, social media).
Fluent in English and Spanish.

- Competences
  - Excellent communication and presentation skills.
  - Strong technical writing skills.
  - In terms of personality, we are looking for someone who enjoys working in a wide range of areas and adapts quickly to new situations.
  - A proactive mindset, with creativity to design new solutions.
  - Ability to work effectively both autonomously and as part of an interdisciplinary team.
  - Comfortable working under pressure and meeting strict deadlines.
  - Proactive working style.

Conditions

- The position will be located at BSC within the Life Sciences Department
- We offer a full-time contract (37.5h/week), a good working environment, a highly stimulating environment with state-of-the-art infrastructure, flexible working hours, extensive training plan, restaurant tickets, private health insurance, support to the relocation procedures
- Duration: Open-ended contract due to technical and scientific activities linked to the project and budget duration
- Holidays: 23 paid vacation days plus 24th and 31st of December per our collective agreement
- Salary: we offer a competitive salary commensurate with the qualifications and experience of the candidate and according to the cost of living in Barcelona
- Starting date: 01/01/2024
Applications procedure and process

All applications must be made through BSC website and contain:

- A full CV in English including contact details
- A Cover Letter with a statement of interest in English, including two contacts for further references

Applications without this document will not be considered.

In accordance with the OTM-R principles, a gender-balanced recruitment panel is formed for every vacancy at the beginning of the process. After reviewing the content of the applications, the panel will start the interviews, with at least one technical and one administrative interview. A profile questionnaire as well as a technical exercise may be required during the process.

The panel will make a final decision and all candidates who had contacts with them will receive a feedback with details on the acceptance or rejection of their profile.

At BSC we are seeking continuous improvement in our recruitment processes, for any suggestions or feedback/complaints about our Recruitment Processes, please contact recruitment [at] bsc [dot] es.

For more information follow this link

Deadline

The vacancy will remain open until a suitable candidate has been hired. Applications will be regularly reviewed and potential candidates will be contacted.

OTM-R principles for selection processes

BSC-CNS is committed to the principles of the Code of Conduct for the Recruitment of Researchers of the European Commission and the Open, Transparent and Merit-based Recruitment principles (OTM-R). This is applied for any potential candidate in all our processes, for example by creating gender-balanced recruitment panels and recognizing career breaks etc.

BSC-CNS is an equal opportunity employer committed to diversity and inclusion. We are pleased to consider all qualified applicants for employment without regard to race, color, religion, sex, sexual orientation, gender identity, national origin, age, disability or any other basis protected by applicable state or local law.

For more information follow this link

Barcelona Supercomputing Center - Centro Nacional de Supercomputación