

## BORSA DE CURRÍCULUMS VITAE

---

### PER A Natural Language Processing Engineer

Ref. 2020-FURLONG

#### S'OFEREIX:

##### Context and Mission

The successful candidate will be involved in applied research projects aiming at extracting structured biomedical knowledge from textual contents. She/he will participate in **the design and development of Text Mining and Natural Language Processing tools and workflows** tailored to analyze different types of texts in the context of different ongoing projects in the group. The selected candidate will contribute to the activities related to textual data collection as well as the implementation, tuning and evaluation of Natural Language Processing algorithms / Machine Learning models.

The Integrative Biomedical Informatics (IBI - <http://grib.imim.es/research/integrative-biomedical-informatics/>) group led by Laura I. Furlong and Ferran Sanz at GRIB (IMIM-UPF) in Barcelona is an international research group characterized by a unique experience in the fields of *Network Medicine* and *Biomedical Text Mining*. Text Mining activities at IBI focus on the extraction and aggregation of biomedical and clinical knowledge from several types of textual sources including **biomedical scientific literature, clinical notes** and **social media**. Some relevant examples of application scenarios include: (i) *mining genetic information on diseases from scientific articles*, thus contributing to the population of the DisGeNET - <https://www.disgenet.org/> - and PsyGeNET - <http://www.psygenet.org/> - knowledge platforms; (ii) *extracting toxicology information from legacy reports*; (iii) *analyzing social media posts and user behavior to detect mental disorders*; (iv) *processing clinical notes to support and automate healthcare decision making and planning*. Part of Text Mining research lines at IBI are carried out in the context of national and international research projects including eTRANSafe, FAIRplus, EMIF, Open PHACTS, MedBioinformatics, iPiE, eTOX.

**Requisits dels candidats:****Required qualifications and skills**

- University degree in informatics, language engineering, bioinformatics or related area
- Master or PhD degree is a plus
- Strong programming skills in Python and/or Java
- Experience in configuring and querying Database Systems, both SQL (e.g. MySQL) and NoSQL (e.g. MongoDB, Elasticsearch) and in Unix
- Experience in working with source code repositories (e.g. Github, BitBucket, etc.)
- Good English communication skills
- Strong problem-solving skills applied to real world analytical problems and ability to communicate outcomes of analyses in clear and precise manner
- Strong interpersonal skills
- Ability to work in a multidisciplinary environment

**Desirable qualifications and skills**

- Experience in one or more of the following areas: entity/relation extraction, information extraction, document classification, semantics, summarisation, question answering, knowledge graph

**Resum del projecte:**

The **Enhancing TRANslational SAFETY Assessment through Integrative Knowledge Management (eTRANSafe)** project develops an integrative data infrastructure and innovative computational methods and tools that aim to drastically improve the feasibility and reliability of translational safety assessment during the drug development process. This infrastructure will be underpinned by development of open standards and robust policies widely accepted by stakeholders, including regulatory agencies and international organisations.

**Per a més informació i presentació de currículum vitae i expedient acadèmic, contactar amb:**

Please send a CV, letter of interest and the contact information for 2-3 referees to [carina.oliver@upf.edu](mailto:carina.oliver@upf.edu) with Ref: IBI

**Data límit de presentació de CV:** 28 de febrer de 2020