



Master project 2021-2022

Personal Information

Supervisor	Osnat Hakimi
Email	ohakimi@uic.es
Institution	UIC/BSC
Website	https://projectdebbie.github.io/
Group	DEBBIE research group

Project

Web development & bioinformatic tools

Project Title:

Information extraction from biomaterials text using machine learning

Keywords:

text mining, machine learning, biomaterials, information extraction, knowledge representation

Summary:

The Database of Experimental Biomaterials and their Biological Effect (DEBBIE) is an EU Horizon 2020 funded project to develop an open access database of biomaterials automatically curated from the scientific literature. DEBBIE was designed to facilitate a more efficient access to the large literature in the field, generate a comprehensive map of research activity and findings, and enable evidence-based selection of materials for medical applications. Beyond the development of database, the project is dedicated to the creation, adaptation and optimization of text mining tools for the biomaterials domain. These tools, including the DEBBIE retrieval and annotation pipeline (see illustration below), are openly available for download and use through the project's GitHub repositories. In the project proposed here, the student will help to design additional extraction tools directed at discovering and organizing ambiguous and non- explicit information for biomaterials text. The student will participate in the evolution of the DEBBIE pipeline and it's associated tools.

References:

<https://onlinelibrary.wiley.com/doi/10.1002/adfm.201909910> <https://www.nature.com/articles/s41578-020-0215-z>

Expected skills::

Python- must, desired: NLP, Java, HTML

Possibility of funding::

No

Possible continuity with PhD: :

To be discussed

Comments:

The project will be co-hosted by UIC and the Barcelona supercomputing Center
