

Department of Biotechnology

Institute of Bioprocess Sciences and Engineering

Univ. Prof. Dr. rer. nat. Dr.-Ing. Johannes Felix Buyel

To interested Master students
in Biology, Biotechnology, (Bio)Engineering
and (Bio)Informatics

Vienna, 2022-10-14

DataExtractor

Research internship and **Master thesis**
at the **Institute of Bioprocess Sciences and Engineering**
in cooperation with Fraunhofer Austria

Dear interested Students,

Abstract: The amount of data available in life sciences, especially biotechnology is ever increasing and more rapidly so in the last decades. Whereas data science is therefore gaining substantial interest, for example for the meta-analysis of clinical trial data, the necessary tools for such kind of work are often not rolled out in breadth yet. In the current project the Institute of Bioprocess Sciences and Engineering in cooperation with Fraunhofer Austria will develop a text mining solution, preferably based on open-source code, to rapidly screen large publication databases for relevant content. The added value of such tool is the substantially increased number of data compared to (biased) searches by individual researchers, i.e. thousands vs. dozens of manuscripts evaluated. The project will use the development of recombinant protein expression levels achieved in plants over time as a show case.

Tasks:

1. Define and refine the text mining query to ensure suitable search outcomes
2. Build dictionaries and taxonomies to contextualize redundancy and ambiguity in terminology
3. In a rapid prototyping approach test the code on a well-defined training data set
4. Apply the refined code in a public data base search (e.g. on pubmed)
5. Write a glorious thesis and publication 😊

Duration: The initial internship will take 3 months and the master thesis will be 6 months. In the course of the project, weekly meetings with your supervisor as well as flexible on-demand meetings will ensure the success of your work.

Requirements: The student successfully applying for this project has good basic knowledge in (plant) biotechnology and bioinformatics and is willing to gain more insights into data science and programming, e.g. using Python. S/He is skilled in written and spoken English to familiarize herself/himself with the relevant protocols and to fluently communicate within the international environment at IBSE.

Contact: For further questions and applications, please contact Johannes Buyel

Muthgasse 18, 1190 Vienna
T +43 1 476 54-79083
F +43 1 476 54-79009
johannes.buyel@boku.ac.at
www.boku.ac.at/dbt/ibse