



TD&K Student Position on Utilizing Advances in Databases and Text Analytics for Medical Sciences

Technion Data & Knowledge Laboratory (TD&K Lab)

<https://tdk.cs.technion.ac.il/>

Contact information:

Dean Light <light.skep@gmail.com>

The objective of this position is to build software that utilizes recent database approaches to text analysis, namely the framework of “document spanners,” to optimize core processes in bioinformatics and medical sciences. Specifically, we target the important operation of oligo design. Oligos (Oligonucleotides) are short DNA or RNA molecules that are synthetically generated for a wide range of applications, from basic medical research to forensic uses like COVID-19 testing. Oligo design takes place before the oligos are created in the lab, when oligos are planned on the computer by algorithms that analyze and simulate potential interactions between the oligo and different DNA molecules in different cells.

We propose to address these challenges by adopting the framework of document spanners that views text analysis, typically around Natural Language Processing, as centered around two core subtasks: (a) the extraction of basic relations from text, and (b) the relational manipulation of the extracted relations. This approach enables us to adopt and adapt decades of advances on database optimization to the task at hand. Within this framework we plan to develop novel algorithms that will enable researchers to design oligos more quickly and accurately. With these algorithms, we aim to improve the performance of core tasks in oligo design, and to devise efficient algorithms for new (and unsolved) ones. Moreover, we aim not only to facilitate existing practice of oligo design, but also to enable new medical research and clinical tools that require even more sophisticated oligo design.

At the Technion Data & Knowledge Laboratory (TD&K Lab), Led by Prof. Benny Kimelfeld, we are looking for talented and highly motivated undergraduate students to help us develop novel oligo design software, while improving their skills in software design, programming and algorithmic design. These projects are led and managed by a consultant in the lab and are done in collaboration with the group of Prof. Naama Geva Zatorski in the department of Medicine, Technion.

Job Requirements:

- Undergraduate student in the CS faculty.
- At least one and a half years left till graduation.

Job Responsibilities:

Under the guidance and training of the project manager, we expect the students to eventually be able to:

- Write high quality maintainable code effectively, including
 - Writing Full-Stack applications
 - Writing computational libraries
 - Writing advanced algorithms
 - Conducting computational experiments and documenting them
 - Split code across different modules and technologies intelligently
 - Package and deliver code for external consumers, including
 - Writing and designing user facing interfaces
 - Writing documentation
 - Interact with open source code projects
 - Efficiently interact with domain knowledge in software projects
 - Reason about user behavior and design and implement effective user/app flows
 - Collaborate as a team via
 - dependencies analysis of software projects
 - version management software (git + github)
 - de-risking strategies via design and code reviews
 - Pass code reviews and reason about code health and development loop optimization
-