POLITECNICO DI MILANO

PhD in INGEGNERIA DELL'INFORMAZIONE / INFORMATION TECHNOLOGY - 38th cycle

Research Area n. 1 - Computer Science and Engineering

THEMATIC Research Field: KNOWLEDGE MANAGEMENT TECHNIQUES FOR EXPLORING DATA CORPUSES ON BIOMEDICAL RESEARCH

Monthly net income of PhDscholarship (max 36 months)
€ 1400.0
In case of a change of the welfare rates during the three-year period, the amount could be modified.

Context of the research activity	
Motivation and objectives of the research in this field	The latest advances in biomedical research have produced an enormous number of scientific datasets, including for example genomic sequences, clinical electronic records, and peer-reviewed papers that demonstrate scientific discoveries. Such information explosion has been even more evident during the COVID-19 pandemic, when several data corpuses were opened to the public. Such information needs effective methods that can handle its complexity, in the context of end-to-end pipelines that include the retrieval of cleaned data, the extraction of relevant information, and the processing, with the final aim to analyze heterogeneous elements for supporting the exploration of scientific hypotheses and subsequent decision making. The goal of this PhD is to explore this context, focusing on several use cases inspired by life sciences domain experts, possibly also producing user-end tools that can serve their practical needs.
Methods and techniques that will be developed and used to carry out the research	The student will use a data science approach, going through the phases of 1) *Data collection and preparation* (data will be identified, cleaned and integrated); 2) *Data exploration* (prepared data will be analyzed with unsupervised techniques to gain insight on relevant relationships among instances and features); 3) *Model*

POLITECNICO DI MILANO

	generation (several models will be designed for the relevant variables; for each algorithm, tuning will be performed); 4) *Decision making* (insights gained by all the analysis steps will be discussed with domain experts; particularly relevant in this stage is the interpretability of the obtained results).
Educational objectives	The student will learn how to compare, classify, and adapt available methods for the extraction, linking, and management of data and knowledge corpuses (e.g., https://allenai.org/data/cord-19) in life sciences-related fields, possibly with applications to genomics. The student will also learn how to master domain-specific querying of data for answering biomedical research questions.
Job opportunities	Knowledge management is an attractive research and professional area. Graduates with expertise in this area can easily find jobs both in academia and in private organizations. PhD candidates who recently graduated in our group are currently employed in important public and private institutions.
Composition of the research group	1 Full Professors 1 Associated Professors 2 Assistant Professors 4 PhD Students
Name of the research directors	Stefano Ceri

Contacts
stefano.ceri@polimi.it http://www.bioinformatics.deib.polimi.it/geco/?team

Additional support - Financial aid per PhD student per year (gross amount)	
Housing - Foreign Students	--
Housing - Out-of-town residents (more than 80Km out of Milano)	--

Scholarship Increase for a period abroad	
Amount monthly	700.0 €
By number of months	6

POLITECNICO DI MILANO

Additional information: educational activity, teaching assistantship, computer availability, desk availability, any other information
LIST OF UNIVERSITIES, COMPANIES, AGENCIES AND/OR NATIONAL OR INTERNATIONAL INSTITUTIONS THAT ARE COOPERATING IN THE RESEARCH: Universitat Politecnica de Valencia; UFL – One Health Center; NUS - Singapore; L3S - Hannover EDUCATIONAL ACTIVITIES (purchase of study books and material, including computers, funding for participation in courses, summer schools, workshops and conferences): financial aid per PhD student 5.707,13 Euro TEACHING ASSISTANTSHIP: (availability of funding in recognition of supporting teaching activities by the PhD student) There are various forms of financial aid for activities of support to the teaching practice. The PhD student is encouraged to take part in these activities, within the limits allowed by the regulations. COMPUTER AVAILABILITY: individual use DESK AVAILABILITY: individual use The student will use the Data Science and Bioinformatics lab at DEIB