Date of publication of the job offer: April19th 2021

Job Offer: Joint industrial PhD position **on applied quantum mechanics and machine learning for molecular simulations of reactions related to drug discovery**

Description of the work to perform during the PhD:

The next frontier in molecular simulations is to be able to perform reactive simulations using fast machine learning potentials. Accurate prediction of the outcomes of an organic reaction is still an unsolved task and only experienced chemists can make reliable predictions based on underlying mechanistic and quantum chemical intuition. In this research project, the PhD candidate will develop a new methodology for the fast simulation of reactive molecules based on fast machine-learning-quantum computation. The major expected outcomes are:

1. Selection of a set of relevant and simple systems to produce a database of accurate quantum mechanical data. The data will be generated using GPUGRID.net
2. Training models and neural network potentials to be used for the simulation of the chemical reaction. Validation of the models comparing the yield predicted and experimental values in collaboration with other PhD students participating in the project.
3. Further validation on internal synthesis data at Bayer and expansion of the applicability domain to a larger set of synthesis routes.

The PhD candidate will perform research 1.5 years at UPF (Barcelona, Spain) and 1.5 years at Bayer (Berlin, Germany).

About the AIDD project

Machine learning is changing our society, as exemplified by speech and image recognition applications. Also the life sciences change rapidly through the use of artificial intelligence, and it is expected that fields like drug development can take advantage of machine learning. The main goal of the AIDD project is to train and prepare the next generation of scientists who need to have skills in both machine learning and drug discovery and will, after graduating, be able to helping speeding up the drug development process. The European Marie Skłodowska-Curie Innovative Training Network funds the AIDD project that brings together twelve academic partners (Helmholtz Zentrum München (coordinator), Germany; Aalto University, Finland; Freie Universität Berlin, Germany; Katholieke Universiteit Leuven, Belgium; Johannes Kepler Universität Linz, Austria; The Swiss AI Lab IDSIA, Switzerland; TU Dortmund, Germany; Universiteit Leiden, Netherlands; Université du Luxembourg, Luxembourg; University of Vienna, Austria; Universitat Pompeu Fabra, Spain and Vancouver Prostate Center, University of British Columbia, Canada) as well as four industrial partners (AstraZeneca, Sweden; Bayer Aktiengesellschaft, Germany; Janssen Pharmaceutica NV , Belgium and Enamine Limited Liability Company, Ukraine).

The AIDD network offers 15 PhD fellowships (referred under the programme as ESR, Early Stage Researcher position). The employed fellows will be supervised by academics who have strong technical expertise and have contributed to some of the fundamental AI algorithms which are used billions of times each day in the world, and by machine learning scientists working at pharmaceutical companies. The developed methods by the fellows will contribute to an integrated "One Chemistry" model that can predict outcomes ranging from different properties to molecule generation and synthesis. The network will offer comprehensive, structured training through a well-elaborated Curriculum, online courses, and six schools.

About the research partners:

Universitat Pompeu Fabra (academic Partner) is a public, international and research-intensive university that ranks 10th in the 2020 Times Higher Education Young University Ranking and the 1st of the Spanish Universities in the QS World University ranking 50 under 50 (position 28 worldwide and 7th in Europe). It is the 1st Spanish

University (IUNE Report 2019) in yearly scientific output per professor, citations per professor, projects within the Spanish National R&D Plan (per 100 lecturers), in projects within the EU Framework Programme (per 100 lectures). It has about 12,000 enrolled students (nearly 1.300 of them in one of its 9 PhD programs), 1,500 teaching and research staff, and 700 administrative and service staff. The UPF is strategically located in the PRBB, one of the leading South European biomedical research hubs, that hosts 1300 international staff and scientists, core scientific facilities and multiple research institutes. The student participating in this project will be enrolled in the PhD studies in BioMedicine and will join the Computational Science Laboratory (lead by Prof. Gianni De Fabritiis) whose interests are the application of computation to solve real world problems, defining intelligence as a form of computation. The research group develops machine learning models with intelligent, useful behavior using reinforcement learning and deep learning, for specific environments. Biomedicine is one environment where physics-based simulations and machine learning provide novel, innovative approaches. The group leads GPUGRID.net , one of the top distributed computing projects worldwide for running molecular simulations on GPUs and the open platform PlayMolecule.org that has around a thousand registered scientists. The group and its spin-off company Acellera have collaborated with major industries worldwide like Sony, Nvidia, HTC mobile, UCB, Pfizer, Biogen and Novartis.

2. **Bayer (industrial Partner)** is a **global enterprise with core competencies in the Life Science fields of healthcare and agriculture.** Its products and services are designed to benefit people and improve their quality of life. At the same time, the Group aims to create value through innovation, growth and high earning power. Bayer is committed to the principles of sustainable development and to its social and ethical responsibilities as a corporate citizen. In fiscal 2018, the Group employed around 117,000 people and had sales of EUR 39.59 billion. Capital expenditures amounted to EUR 1.5 billion, R&D expenses to EUR 5.2 billion.

Project and Institution that finance the contract: This project is funded by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 956832.

Official number reference: PREUR01121 - H2020-MSCA-ITN-2020-AIDD-956832-G.DeFabritiis

Essential Skills and Experience

- Master's degree in computer science, cheminformatics, bioinformatics or equivalent subject
- Courses in machine learning
- Courses in programming.

Desired skills:

- Experience of software engineering
- Proven experience of Python programming
- Experience of deep learning libraries for instance TensorFlow and/or PyTorch)
- Experience with libraries such as RDKit or scikit-learn would be of advantage
- Good command of modern software development tools, such as git
- Courses in drug development

The successful candidate will also demonstrate a passion for driving scientific questions with a positive and problem-solving attitude and the willingness to undertake challenging analysis tasks in a timely fashion. Excellent English is required, both spoken and written, and the ability to work effectively both independently and in cross-functional teams. We also believe that you enjoy teamwork, have a collaborative nature and will be an encouraging colleague to all. Female researchers and candidates are particularly encouraged to apply.

Note that due to the requirements of the Marie Skłodowska-Curie Innovative Training Network,candidates must

not have resided or carried out their main activity (work, studies, etc.) in Spain for more than 12 months in the 3 years immediately prior to the recruitment date.

Benefits of the opening:

- Marie Skłodowska-Curie funding offers very competitive salaries allowance for the period hired in Spain: 3.119,58 EUR/month; allowance for the period hired in Germany: 3.171,9 EUR/month. Note that net salary is subject to country-specific deductions.
- Mobility and family allowances are also offered.

Information on the application process:

Selection process: Prior to applying please check the programme eligibility criteria: https://ai-dd.eu/esr-positions#eligibility.

To apply send a CV, motivation letter and proof of your educational degree (or expected time until you obtain your degree) to recruit@ai-dd.eu. Make sure to provide sufficient details about your educational background and work experience.

VERY IMPORTANT indicate both in the subject of the email and the motivation letter the code "ESR5".

Deadline to apply 28/04/2021 (the screening will start immediately; do not wait until the deadline to submit your application).