

**Research Theme: Biomedical Data Science/Drug Discovery/Machine learning/Computational Biology**

**MSc Research Project Title: Integrating knowledge with artificial intelligence for natural product drug discovery**

**Principal Investigator/Supervisor: Wilson Goh**

**Co-supervisor/ Collaborator(s) (if any):**

### **Project Description**

#### **a) Background:**

*Natural products (NPs) are a huge reserve of bioactive compounds and will continue to play crucial roles in drug discovery. However, high biodiversity impedes identification of useful bioactive compounds. Recent advances in omics technologies have opened the gateway towards realizing and harnessing the potential of NPs in areas such as functional analysis of therapeutic effects and drug development. The emergence of Artificial Intelligence (AI) will further accelerate NP-based drug discovery and development.*

#### **b) Proposed work:**

With the rapid development of computers and AI, virtual screening has become a very popular method for drug discovery. AI models can help identify potential drug-target interactions (DTIs) and thus shorten the overall drug development cycle without the need for large amounts of money. Usually in AI models, data with known biological activity are used as training data and NPs are the test data, so that the biological activity of NPs is predicted based on the known biological activity data. Therefore, researchers expect to obtain high-quality NP datasets that provide accurate information on molecular structure, physicochemical properties, biological activity, etc. One previous review describes the process of dataset preparation and said that the lack of large quantities of high-quality bioactivity data is a great challenge confronted by ligand-based AI models. In the field of NPs, the situation is even worse due to difficulties in the exploration of biological activity. Besides, there is no database that integrates both bioactive data and NPs. We need a database that contains both NPs and bioactive data and ideally knows their interrelationships such as similarity.

#### **c) Preferred skills:**

Highly motivated candidates with strong computing/mathematical/statistical skills and analytical chemistry skills are strongly encouraged to apply.

### **Supervisor contact:**

**If you have questions regarding this project, please email the Principal Investigator:**

wilsongoh@ntu.edu.sg

**SBS contact and how to apply:**

Associate Chair-Biological Sciences (Graduate Studies) : [AC-SBS-GS@ntu.edu.sg](mailto:AC-SBS-GS@ntu.edu.sg)

Please apply at the following:

**Application portal:**

<https://venus.wis.ntu.edu.sg/GOAL/OnlineApplicationModule/frmOnlineApplication.ASPX>