Report for newly appointed faculty startup

1. Name of project leader： **Sidorov Pavel**
2. Project title: **Discovery of novel reactions by Artificial Intelligence**
3. Report

Our research project is focused on the development of approaches of novel reaction generation by Machine learning (ML) and Artificial Intelligence (AI). Recently, the interest of chemoinformatics shifted from building models to predict some specific properties, towards the prediction of structures possessing a desired property. We have stepped further with this challenge and developed a system to predict reactions of a given type. In addition to that, we have designed a number of "chemical filters" which help with the decision if a reaction is new and feasible. Using reactions extracted from US Patent database (>2.5 million reactions, public data) we have generated novel, chemically correct Suzuki coupling reactions. To ensure the feasibility of these reactions, fast DFT calculations have been performed. Thus, the benefits of the approach we propose have been demonstrated.

To carry out the project, we have purchased three high-end calculation workstations equipped with GPU for neural network training. As a team with an overseas PI, we collaborate extensively with his team in the University of Strasbourg (France). In addition, we work together with the team in Kazan Federal University (Russia) – the leading developers of reaction data management tools. We have invited Dr Nugmanov from that team for a short-term stay for the joint development, which further reinforces our bonds with Kazan Federal University.

On the other hand, the project also motivated us to find collaborations within ICReDD. Prof Ito has shared with us his experimental expertise and has suggested to further focus on Buchwald–Hartwig amination, an important reaction for organic synthesis. Prof Maeda will help with further investigation of discovered reactions with the GRRM and AFIR methodologies. Prof Takigawa as an expert in Machine learning develops new types of neural networks to increase the success rate of the model. From now on, the further developments will be advanced in this collaboration.

4. Research achievement

Bort, William; Baskin, Igor I.; Sidorov, Pavel; Marcou, Gilles; Horvath, Dragos; Madzhidov, Timur; et al. (2020): Discovery of Novel Chemical Reactions by Deep Generative Recurrent Neural Network. ChemRxiv. https://doi.org/10.26434/chemrxiv.11635929.v1 (under review in Nature Communications)