**ПРОЕКТИРОВАНИЕ ГРАФА ЗНАНИЙ ДЛЯ ИССЛЕДОВАНИЯ
ВЗАИМОДЕЙСТВИЙ БИОПОЛИМЕРОВ И МАЛЫХ МОЛЕКУЛ****Широкий Н. И. (ИТМО)****Научный руководитель – кандидат химических наук Серов Н. С.
(ИТМО)**

Введение. Естественнонаучные исследования являются источником колоссальных объемов информации. В последние годы, благодаря активному развитию инструментов обработки, анализа и хранения данных, а также технологий искусственного интеллекта замечается тенденция к исследованию накопленных данных для выявления скрытых паттернов, закономерностей и взаимосвязей, позволяющих получить более полные и достоверные представления исследуемых объектов. Актуальность исследования взаимодействия биополимеров и малых молекул объясняется тем, что посредством химической модификации возможна корректировка свойств биополимеров, либо даже создание полимеров с абсолютно новыми свойствами. Разработанные таким образом биополимеры могут быть использованы в медицине, для разработки сенсорных систем различного назначения, и т. д.

Основная часть. Создание графа знаний взаимодействий биополимеров и малых молекул решает следующие задачи:

- 1) Агрегация разнородных источников данных, описывающих различные свойства одних и тех же сущностей. При сборе данных о биополимерах и малых молекулах и их формировании в граф свойств (property graph) предполагается максимально возможная очистка данных от дубликатов и группировка свойств биополимеров и малых молекул, относящихся к одним и тем же сущностям, но исходно находящихся в различных источниках данных. [1]
- 2) Формирование обобщенной карты взаимосвязей между биополимерами и малыми молекулами различных типов с описанием характеристик и свойств данных связей, характеризующих силу связывания (константа диссоциации, энергия связи). [2]
- 3) Исследование топологии графа знаний и возможности прогнозирования потенциальных связей в графе при помощи алгоритмов машинного обучения для работы с графами. [3]

Выводы. Главным продуктом предлагаемой реализации графа знаний является создание уникальной базы данных, обеспечивающей быстрый доступ к данным о связывании биополимеров и малых молекул, а также возможность прогнозировать связывания для кандидатов при добавлении их в базу.

Список использованных источников:

1. Liang, X., Li, D., Song, M., Madden, A.D., Ding, Y. et al. (2019). Predicting biomedical relationships using the knowledge and graph embedding cascade model. *PLoS ONE*, 14.
2. Timón-Reina, S., Rincón, M., & Martínez-Tomás, R. (2021). An overview of graph databases and their applications in the biomedical domain. *Database: The Journal of Biological Databases and Curation*, 2021.
3. Dogan, T., Atas, H., Joshi, V., Atakan, A., Rifaioglu, A.S., Nalbat, E., Nightingale, A., Saidi, R., Volynkin, V., Zellner, H., Cetin-Atalay, R., Martin, M.J., & Atalay, V. (2021). CROssBAR: comprehensive resource of biomedical relations with knowledge graph representations. *Nucleic Acids Research*, 49, e96 - e96.