# Explainable Artificial Intelligence in the Life Sciences

Andrea Mastropietro[1-3]

*[1]Department of Life Science Informatics and Data Science, B-IT, LIMES Program Unit Chemical Biology and Medicinal Chemistry, University of Bonn, Friedrich-Hirzebruch-Allee 6, D-53115 Bonn, Germany*

*[2]Lamarr Institute for Machine Learning and Artificial Intelligence, University of Bonn, Friedrich-Hirzebruch-Allee 6, D-53115 Bonn, Germany*

*[3]Data Science Center, Nara Institute of Science and Technology, 8916-5 Takayama-cho, Ikoma, 630-0192 Nara, Japan*

Deep learning has become a widely used tool in chemoinformatics and bioinformatics. Graph-based models, such as graph neural networks, can be applied to molecular graphs to predict chemical properties with high accuracy [1], and diffusion models can be employed to generate valid molecules [2]. However, deep learning lacks transparency, which is undesirable in applications such as drug design, where model outputs must be interpretable to be trusted. To address this limitation, explainable artificial intelligence strategies have been developed. This talk will discuss how neural networks can be used and explained in the context of life sciences and biomedicine.

From the point of view of predictive artificial intelligence, we will examine the learning characteristics of graph neural networks for compound activity and potency prediction [3,4]. Moreover, we will dive into generative modeling, explaining the generations of diffusion models for linker design.

We will analyze whether such deep learning-based strategies are capable of extracting and learning meaningful biochemical knowledge from data, or whether they mainly rely on memorizing statistical patterns, investigating their applicability in biomedical tasks such as drug design.

Bibliography:

[1] Chuang, K. V., Gunsalus, L. M., & Keiser, M. J. Learning Molecular Representations for Medicinal Chemistry (2020). Journal of Medicinal Chemistry, 63, 8705–8722.
[2] Hoogeboom, E., Satorras, V. G., Vignac, C., & Welling, M. Equivariant diffusion for molecule generation in 3D (2022). In International Conference on Machine Learning, 8867-8887.
[3] Mastropietro, A., Pasculli, G., Feldmann, C., Rodríguez-Pérez, R. & Bajorath, J. EdgeSHAPer: Bond-Centric Shapley Value-Based Explanation Method for Graph Neural Networks (2022). iScience 25, 105043.
[4] Mastropietro, A., Pasculli, G., & Bajorath, J. Learning characteristics of graph neural networks predicting protein–ligand affinities (2023). Nature Machine Intelligence, 5 (12), 1427-1436.