Statistics Are a Girl's best Friend: Expanding the mechanistic Study Toolbox with Data Science

<u>Anat Milo</u>

Ben-Gurion University of the Negev, Beer-Sheva, 84105, Israel anatmilo@bgu.ac.il

The value of amassing and standardizing chemical data for improving the efficiency of chemical discovery is becoming increasingly clear. Machine learning analyses of these data are focused on finding correlations, trends and patterns to uncover needles of knowledge in the haystack of chemical reactions. However, in many cases, especially in academic settings, we do not have the means to produce large data sets, so by necessity we remain in the Small Data regime. In this talk, I will present our work in the field of organocatalysis focused on applying machine learning strategies to small data sets as a means to uncover underlying mechanisms. Specifically, our group has been involved in the *in-situ* modification and mechanistic interrogation of secondary-sphere interactions in organocatalysis. As proof-of-concept, we modified the secondary sphere of N-heterocyclic carbenes (NHCs) using boronic acids (BAs) under reaction conditions and tested their reactivity and selectivity in the benzoin reaction and probe their mechanism. Ultimately, we aim to show that whereas Big Data serves to identify hidden correlations, Small Data encourages the discovery of causation. In this sense, Small Data is not just a necessity, but is key to bridging the gap between human intuition and machine learning.



- [1] A. Milo, Democratizing synthesis by automation. <u>Science 2019</u>, 363, 122.
- [2] V. Dhayalan[‡], S. C. Gadekar[‡], Z. Alassad[‡], A. Milo, Unravelling mechanistic features of organocatalysis with *in situ* modifications at the secondary sphere. *Nat. Chem.*, **2019**, *11*, 543–551.
- [3] S. C. Gadekar[‡], V. Dhayalan[‡], A. Nandi, I. L. Zak, S. Barkai, M. Shema Mizrachi, S. Kozuch, A. Milo, Rerouting an organocatalytic reaction by intercepting its reactive intermediates. ChemRxiv. Preprint. <u>chemrxiv.13061717.v1</u> (manuscript accepted).