Materials Informatics for green energy harvesting and storage

Paul Clarke,[a] Omer Kaspi,[a] Hadar Binyamin,[a] Abraham Yosipof,[b] Hanoch Senderowitz[a]

[a]*Department of Chemistry, Bar Ilan University, Ramat Gan, 5290002, Israel*
[b]*Department of Information Systems, College of Law & Business, Ramat-Gan, P.O.Box 852 Bnei-Brak 5110801, Israel*

Materials informatics is a rapidly developing field. This development is fueled by the continuous increase of available experimental and computational data, as well as by new computational methods to model them, in particular Machine Learning (ML) workflows. Like its "sister field", chemoinformatics, materials informatics can help bridge the so-called data-knowledge gap by offering smart ways to navigate the enormous materials space in search of new materials with favorable properties.

This talk will focus on the application of ML methods in the field of green energy harvesting and storage. A common theme to the energy devices we chose to study is their reliance on molecular "triggers", namely, small molecules which are key to the operation of the device. For energy harvesting, we focus on solar cells. Such cells hold the potential to meet the growing worldwide demand for clean and sustainable energy. Today most solar cells are based on silicon, yet new alternatives are continuously emerging including organic photovoltaic cells, and dye sensitized solar cells (DSSCs). Previously we have demonstrated how pharmacophore models can be used to identify new dyes with favorable (predicted) electronic properties for DSSCs. More recently, we have developed the first global ML model for the Dye Sensitized Solar Cell Database (DSSCDB) which leverages the use of a device fingerprint (DFP) to resolve duplicate dye entries and Artificial Neural Networks for the prediction of overall Power Conversion Efficiency (PCE). This global model exhibits high predictive ability on an external test set and has a broad applicability domain, including organic compounds of various scaffolds, metal organic compounds and dye mixes.

For energy storage we focus on organic redox flaw batteries. These are rechargeable batteries that employ organic molecules as the active materials in their redox reactions. Organic redox flow batteries offer potential advantages over traditional metal-based redox flow batteries such as low-cost, high-energy density, and environmental friendliness. From among the multiple parameters responsible for the operation of these devices, the oxidation potential of the organic molecules is critical for their efficiency and durability. Thus, we have secured from the literature a dataset of ~30,000 QM-calculated oxidation potentials of organic molecules and set out to develop corresponding ML models. Our preliminary efforts highlight the difficulties in modeling the oxidation potentials of organic compounds as well as suggesting ways to overcome these difficulties.

As a final take home message, we emphasize the need to conduct research in the field of materials informatics in general and energy devices in particular, in close collaboration with experimentalists, to provide insight into the observed trends and to capitalize on the results.