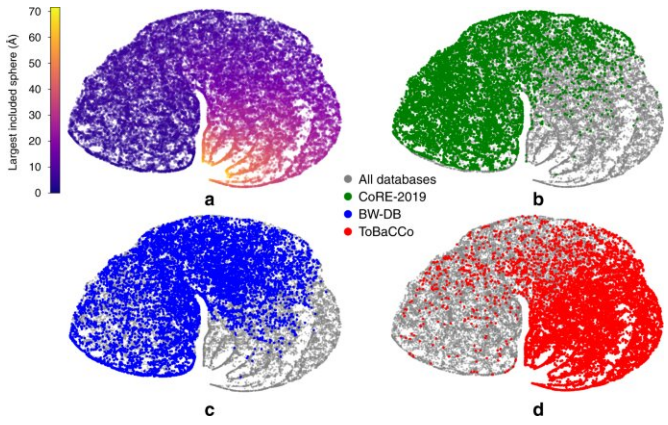
Machine-learning helps sort out massive materials' databases

14 September 2020



EPFL and MIT scientists have used machine-learning to organize the chemical diversity found in the ever-growing databases for the popular metal-organic framework materials.

Metal-organic frameworks (MOFs) are a class of materials that contain nano-sized pores. These pores give MOFs record-breaking internal surface areas, which can measure up to 7,800 m^2 in a single gram of material. As a result, MOFs are extremely versatile and find multiple uses: separating petrochemicals and gasses, mimicking DNA, producing hydrogen, and removing heavy metals, fluoride anions, and even gold from water are just a few examples.

Because of their popularity, material scientists have been rapidly developing, synthesizing, studying, and cataloging MOFs. Currently, there are over 90,000 MOFs published, and the number grows every day. Though exciting, the sheer number of MOFs is actually creating a problem: "If we now propose to synthesize a new MOF, how can we know if it is truly a new structure and not some minor variation of a structure that has already been synthesized?" asks Professor Berend Smit at EPFL Valais-Wallis, which houses a major chemistry department.

To address the issue, Smit teamed up with Professor Heather J. Kulik at MIT, and used machine learning to develop a "language" for comparing two materials and quantifying the differences between them. The study is published in *Nature Communications*.

Armed with their new "language," the researchers set off to explore the chemical diversity in MOF databases. "Before, the focus was on the number of structures," says Smit. "But now, we discovered that the major databases have all kinds of bias towards particular structures. There is no point in carrying out expensive screening studies on similar structures. One is better off in carefully selecting a set of very diverse structures, which will give much better results with far fewer structures."

Another interesting application is "scientific archeology": The researchers used their machine-learning system to identify the MOF structures that, at the time of the study, were published as very different from the ones that are already known.

"So we now have a very simple tool that can tell an experimental group how different their novel MOF is compared to the 90,000 other structures already reported," says Smit.

More information: Seyed Mohamad Moosavi et al. Understanding the diversity of the metal-organic framework ecosystem, *Nature Communications* (2020). DOI: 10.1038/s41467-020-17755-8

Provided by Ecole Polytechnique Federale de Lausanne

APA citation: Machine-learning helps sort out massive materials' databases (2020, September 14) retrieved 1 October 2020 from https://techxplore.com/news/2020-09-machine-learning-massive-materials-databases.html

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.