

Harnessing the knowledge of metal-organic frameworks

Peter Andrew Wood¹, Peyman Moghadam², Aurelia Li², Seth Wiggin¹, Andi Tao², Andrew Maloney¹, Suzanna Ward¹, David Fairen-Jimenez²

¹Ccdc, Cambridge, United Kingdom, ²Adsorption & Advanced Materials Laboratory (AAML), Cambridge, United Kingdom
E-mail: wood@ccdc.cam.ac.uk

Research into the rational design and synthesis of extended materials has grown considerably over the last 20 years, with these materials finding applications in areas such as gas storage, catalysis and drug delivery. The Cambridge Structural Database (CSD) contains over 875,000 small molecule crystal structures, including tens of thousands of metal-organic frameworks (MOFs) and other extended materials. The CSD is therefore an excellent resource for both identifying existing structures with promising characteristics and for guiding future research developments. A major issue in this area though is the huge diversity in the composition of extended materials, and the definition of such materials, which makes collating crystal data for all the extended materials in the CSD a challenging prospect.

Here we will discuss recent research to identify all the extended metal-organic systems in the CSD and crucially to develop a maintainable system that will continue to keep this metal-organic framework subset of the CSD up-to-date in the future. This approach used an iterative method based on chemical substructures to identify a subset of over 75,000 metal-organic frameworks within the CSD. We will review the nature and properties of the metal-organic frameworks in the CSD, as well as mention some future directions for MOF research that the approach has highlighted.

Keywords: [metal-organic framework](#), [data mining](#), [materials design](#)