Subject: Re: Converting COD2020 to SDF
Posted by thomas on Fri, 29 Jan 2021 08:36:13 GMT
View Forum Message <> Reply to Message

Some years ago, when we needed statistical data from crystallographic data for our conformer generation and forcefield algorithms in OpenChemLib, and after getting a veto from Colin Groom to use the CSD for that purpose, we looked for alternatives. As a result we invited one of the COD maintainers (Antanas Vaitkus) to spend parts of his PhD time at Actelion in Switzerland to work on the cif2sdf conversion and especially to improve the calculation of bonds from atom coordinates, which has some issues especially with organo-metallic structures. He did a marvelous job improving the bond calculation logic by looking into lots of original papers, added validation code to produc warnings and errors. After returning to Vilnius he established the conversion as a regularly occurring process that not only creates an SD-file, but also creates a dwar file in their SVN repository.

We download the dwar, remove fishy structures, apply some minor changes, add the organic/metal-organic/inorganic classification, apply a template, and put it on this website for download.

Thomas