CSD One Million: Reasons why the crystallographic community is exemplary

I. J. Bruno, N. T. Johnson, M. P. Lightfoot, S. C. Ward

The Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB2 1EZ, United Kingdom

bruno@ccdc.cam.ac.uk

The recent surpassing of 1 million structures in the Cambridge Structural Database [1] offered a moment for celebration and an opportunity to reflect. Achieving this significant milestone is a testament to the exemplary initiatives and engagement emanating from the crystallographic community over many decades. The development of semantic representation formats [2], the cultivation of joined-up publishing workflows, and the broad adoption of standards all pre-empted the principles, guidelines and practices that have come to dominate the discourse around research data preservation and reuse today [3]. We cannot however rest on our laurels. The curation activities of organisations such as the Cambridge Crystallographic Data Centre remain of critical importance and must continue to evolve. We must ensure that our data resources remain relevant and can be readily utilised by the data-driven approaches being applied to the complex scientific problems of today.

This presentation will offer reflections on the successes of the crystallographic community that have been critical in ensuring the outputs of the past can conform to the expectations and demands of the future. It will highlight how these have enabled a wealth of structural chemistry knowledge to be applied across industry and academia to innovate and educate [4]. Additionally, it will look at the challenges and opportunities presented by an evolving research publication landscape, new experimental and computational methods, and the desire for greater reproducibility and richer reuse of structural chemistry data.

[1] Groom C. R., Bruno I. J., Lightfoot M. P. & Ward S. C. (2016). *Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater.* **72(2)**, 171.

[2] Hall S. R. & McMahon B. (2016). *Data Sci. J.* **15(3)**, 1.

[3] Wilkinson M. D., Dumontier M., Aalbersberg IjJ., *et al.* (2016). *Sci Data*. **3(1)**, 1.

[4] Taylor R., Wood P. A. (2019) *Chem Rev.* **119(16)**, 9427.

Keywords: data; community; standards; FAIR; CSD