

Research Data Management: administration, raw diffraction data, structure factors and coordinates at the UK's National Crystallographic Service (NCS)

S. J. Coles.

UK National Crystallography Service, Chemistry, Faculty of Natural and Environmental Sciences, University of Southampton, Highfield, Southampton, SO17 1BJ, UK.

The need to manage, curate and disseminate data has become paramount in the modern era of academic research. The data explosion that has occurred at the same time has prompted an increased requirement for transparency about its generation and a greater responsibility and accountability for facilities to provide accurate and long term mechanisms for archival and curation.

The NCS has led the way for chemical crystallography for around 15 years in developing approaches to addressing this problem¹. The eCrystals project (<http://ecrystals.chem.soton.ac.uk/>) developed an institutional repository approach to curating and disseminating coordinates, structure factors and a range of other information relating to the 'derived' data from a crystallographic experiment. However, raw diffraction data, although being rigorously archived and in the last 15 years highly curated, is only available on request directly to the NCS. eCrystals has been designed to act as a discipline specific data repository, which has resulted in a pragmatic metadata scheme for the description of its contents and this promotes discovery and reuse of the material it makes available.

There is also a necessary administrative function in running a facility that provides a service and this is intrinsically related to the data itself. Over 30 years of operation the NCS has accumulated a range of databases, spreadsheets and forms to meet ever-changing requirements for administration, tracking and reporting. For the last 15 years a range of NCS projects has been researching and addressing this problem. However, becoming an EPSRC mid-range facility in 2010 prompted a review of requirements and with additional demands for enhanced user interaction and reporting, alongside the Web becoming a more prevalent and mature technology that people readily engage with, "Portal" was conceived. Portal aimed to bring together all the elements described above into a single unified and coherent system.

We have learnt a lot from this work and Portal has largely achieved its goals, however there are significant aspects of the data repository yet to be incorporated and also the need to maintain a modern codebase. We have therefore embarked on an 18-month project address these matters. "Portal – The Next Generation" will be a combination of a laboratory information system and a data repository with specific functions and plug-ins tailored for the operation of a crystallographic facility and its resulting data. The design objectives of this system, development progress and the potential for its availability to the community will be discussed.

1. S.J. Coles et al, *J.Appl.Cryst.*, (2005), **38**, 819–826; M.B. Hursthouse & S.J. Coles, *Crystallogr. Revs.*, (2014), **20:2**, 117-154.