

FACT and FAIR with big data allows objectivity in science: the view of crystallography

John R Helliwell, Department of Chemistry, University of Manchester, M13 9PL, UK
john.helliwell@manchester.ac.uk

A publication is an important narrative of the work done and interpretations made by researchers securing a scientific discovery. As The Royal Society neatly states though “Nullius in verba” (“Take nobody’s word for it”), whereby the role of the underpinning data is paramount. Thereby the objectivity that preserving those data with the article gives is because readers can check each decision of the authors. But, how to achieve full data archiving? This is the raw data archiving challenge, in size and need for correct metadata. Processed diffraction data and final derived molecular coordinates archiving in crystallography has achieved an exemplary state of the art relative to most fields. One can credit IUCr with developing exemplary peer review procedures, of narrative, underpinning structure factors and coordinates data and validation report, through its checkcif development and submission system introduced for Acta Cryst. C and subsequently developed for its other chemistry journals. The crystallographic databases likewise have achieved amazing success and sustainability these last 50 years or so. The wider science data scene is celebrating the FAIR data accord, namely that data be *Findable, Accessible, Interoperable and Reusable*. Some social scientists also emphasise more than FAIR being needed, it should be “FACT”, which is an acronym meaning *Fair, Accurate, Confidential, and Transparent*, this being the issue of ensuring reproducibility not just reusability. [Confidentiality of personal data not being relevant to our data obviously.] Acta Cryst. B, C, E and IUCrData are the closest I know to being both FACT and FAIR where, I repeat for due emphasis: the narrative, the automatic ‘general’ validation checks and the underpinning data are checked thoroughly by subject specialists (i.e. the specialist referees). IUCr Journals are also the best that I know of for encouraging and then expediting the citation of the doi for a raw diffraction data set in a publication; examples can be found in IUCrJ, Acta Cryst D and Acta Cryst F. The wish for a checkcif for raw diffraction data has been championed by the IUCr DDDWG (Diffraction Data Deposition Working Group) and its successor, the IUCr Committee on Data.