Sustaining our community: trust in science and the role of crystallography

John R Helliwell1 and Chiara Massera2

1 Department of Chemistry, University of Manchester, Manchester M13 9PL, United Kingdom,
2 Dipartimento di Scienze Chimiche, della Vita e della Sostenibilità Ambientale, Università di Parma, Viale delle Scienze 17/A, Parma 43124, Italy

john.helliwell@manchester.ac.uk

Keywords: trust; reliability; reproducibility; replicability; data reuse

Education has a vital role to play in the scientific process so that it can be trusted. There are efforts by disparate science communities to introduce new terms to ensure trust in science. These new terms have merit for discussion in crystallographic teaching commissions and possible adoption by crystallographers too. We have published a recent Teaching and Education article on this topic [1] where we envisage a students’ discussion seminar on ‘trust in science and the role of crystallography’ where the students could explore firstly the domain of crystallography and then, more broadly, history of science examples. Loss of trust reduces the confidence inside a community and, as well, can disparage the way that a community is perceived from outside. Trust in science is built up from different facets. To that end the crystallographic community has for many decades used the word ‘reliability’ as exemplified by its Rfactors as well as other metrics. Other science communities and policy bodies have new terminologies such as FAIR and FACT. These have the following meanings: FAIR data means that the data underpinning a publication are findable, accessible, interoperable, and reusable and is a general term in data science. In fact, crystallography has a long tradition in FAIR data through its databases linked to its publications. FACT, coming from the social sciences, means their data should show fairness, accuracy, confidentiality, and transparency. Whereas FAIR looks at practical issues related to the sharing and distribution of data [2], FACT focuses more on the foundational scientific challenges [3]. These developments connect with efforts to improve reproducibility and replicability in science in general as exemplified by the US National Academies of Sciences, Engineering and Medicine who published in 2019 on the Reproducibility and Replicability of Science [4]. Independently the crystallography community has developed, indeed led the way in introducing, a distinct crystallographic information framework (CIF) of clear ontologies within a CIF file [5]. The International Union of Crystallography has a Committee for the Maintenance of the CIF Standard (https://www.iucr.org/resources/cif/comcifs), established in 1993. Central to this approach is a check of the CIF file; checkCIF reports on the consistency and integrity of crystal structure determinations reported in CIF format. Similarly, any Protein Data Bank deposition involves an extensive advisory PDB validation report (https://www.wwpdb.org/validation/validation-reports) assessing numerous indicators of correctness against the processed diffraction data and expected molecular geometry values. In conclusion crystallography is a discipline where community-agreed processed diffraction data and model validation checks are routinely made, and now being extended in various ways to the primary experimental data (“the raw data as ground truth”) according to community wishes and standards. These approaches provide the best chance for ensuring reliability and thereby trust in what we do. Most importantly these standards preserve our community. We think though that the terminologies from other areas of the sciences could usefully assist the crystallographic community in its policies such as in journal notes for authors, as well as how we engage with the public and students [1].