

Search-match tool for atomic pair distribution functions and Crystallography Open Database

P. Juhás¹ and L. Pouchard¹

¹*Computational Science Initiative, Brookhaven National Laboratory, Upton, New York 11973, USA*

Search-match programs are routinely used for powder diffraction data to identify phases, verify sample purity and/or find the best initial structure for a detailed structure refinement. We present a prototype search-match tool that brings similar capabilities for atomic pair distribution function (PDF) analysis. We have utilized the Crystallography Open Database (COD)¹ and the DiffPy-CMI software² to compute theoretical x-ray PDFs for the entire COD collection and to extract their characteristic features such as the nearest neighbor distance, positions of the strongest peak, and their probability moments at selected intervals. These characteristics were entered into Elasticsearch³ system together with selected fields from the COD entries allowing to do search-match queries for experimental PDF curves. The large set of simulated PDFs enabled further analytics such as grouping of near-equivalent COD entries, and assessment of scaling-independent features in the PDFs to facilitate similarity searches even for phases that are not yet in COD. Elasticsearch is a powerful and extensible search engine which allows to search and collate multiple document collections using either programmatic or natural-language queries. We will demonstrate such cross analytics with experimental records from the XPD powder diffraction beamline at the NSLS-II synchrotron and also discuss real-time PDF analysis during data acquisition at XPD.

¹ S. Gražulis, D. Chateigner, R. T. Downs, A. F. T. Yokochi, M. Quirós, L. Lutterotti, E. Manakova, J. Butkus, P. Moeck, and A. Le Bail, *J. Appl. Crystallogr.* **42** (2009), 10.1107/S0021889809016690.

² P. Juhás, C. L. Farrow, X. Yang, K. R. Knox, and S. J. L. Billinge, *Acta Crystallogr. A* **71**, 562 (2015).

³ “Elasticsearch - Open Source Search & Analytics,” <https://elastic.co> (2018).