

# Importance of powder diffraction raw data archival in a curated database for materials science applications

Soorya Kabekkodu\* and Thomas Blanton

International Centre for Diffraction Data, 12 Campus Blvd., Newtown Square, Pennsylvania 19073, USA. \*Correspondence e-mail: kabekkodu@icdd.com

Received 15 April 2024

Accepted 4 July 2024

Edited by L. Dawe, Wilfrid Laurier University, Waterloo, Ontario, Canada

This article is part of a collection of articles from the IUCr 2023 Congress in Melbourne, Australia, and commemorates the 75th anniversary of the IUCr.

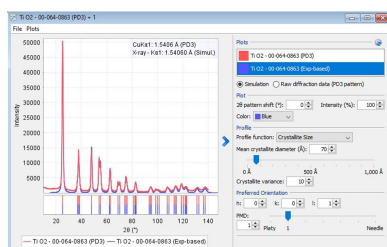
**Keywords:** powder diffraction; raw data; phase identification; Powder Diffraction File.

In recent years, there is a significant interest from the crystallographic and materials science communities to have access to raw diffraction data. The effort in archiving raw data for access by the user community is spearheaded by the International Union of Crystallography (IUCr) Committee on Data. In materials science, where powder diffraction is extensively used, the challenge in archiving raw data is different to that from single crystal data, owing to the very nature of the contributions involved. Powder diffraction (X-ray or neutron) data consist of contributions from the material under study as well as instrument specific parameters. Having raw powder diffraction data can be essential in cases of analysing materials with poor crystallinity, disorder, micro structure (size/strain) *etc.* Here, the initiative and progress made by the International Centre for Diffraction Data (ICDD<sup>R</sup>) in archiving powder X-ray diffraction raw data in the Powder Diffraction File<sup>TM</sup> (PDF<sup>R</sup>) database is outlined. The upcoming 2025 release of the PDF-5+ database will have more than 20 800 raw powder diffraction patterns that are available for reference.

## 1. Introduction

The powder X-ray diffraction (PXRD) method, first demonstrated in 1916 by Debye and Scherrer, has remained a vital technique in materials characterization for over a century and is an indispensable analytical tool. One advantage of the powder X-ray diffraction (PXRD) method is that it often serves as a fingerprint of the solid-state materials under study. The powder diffraction method is a vital characterization tool for phases where growing a single crystal is difficult (or not attainable in practice) or the phase of interest is a part of a mixture.

In earlier decades the PXRD technique's applications were limited to qualitative and semi-quantitative phase analysis and macroscopic stress measurement (Dinnebier & Billinge, 2008). The principal limitation of PXRD was the very nature of powder diffraction where three-dimensional reciprocal space is usually projected on a one-dimensional angular ( $2\theta$ ) axis resulting in severe peak overlaps. The advances made in the instrumentation and methodologies in the later part of the 20th century have opened up a wide variety of applications, from complex materials identification, quantification, and structure elucidation to micro structure and texture analysis (Gilmore *et al.*, 2019). In recent years, the advent of automated diffractometers (equipped with special environment sample cells) and two-dimensional (2D) detectors have made *in-situ* monitoring of complex chemical reactions and non-ambient parametric studies possible (Gilmore *et al.*, 2019).



In the powder diffraction method, ‘powder’ refers to a large number of tiny crystallites randomly oriented. Sometimes PXRD is referred to as the polycrystalline diffraction method. A theoretical estimate is that at least 50 000 crystallites in the X-ray illuminated specimen volume are necessary to obtain a random powder diffraction pattern (Smith, 2001; Whitfield *et al.*, 2019). However, in practice, many samples are not powders, and materials under study can range from solid blocks of a metal to polymer gels. Unlike the single-crystal method, powder diffraction provides a wealth of information on the bulk material as shown in Fig. 1.

## 2. The Powder Diffraction File

The Powder Diffraction File (PDF), managed and maintained by the International Centre for Diffraction Data (ICDD, www.icdd.com), is a powerful database for materials characterization that has been used extensively by the scientific community for more than eight decades. The ICDD is a non-profit scientific organization dedicated to collecting, editing, publishing and distributing powder diffraction data to identify materials. Starting with approximately 1000 PXRD data entries on printed cards in 1941, the database has grown to contain over one million unique material data sets. The history, growth and development of the Powder Diffraction File has been summarized in various publications (Messick, 2012; Kaduk, 2019; Kabekkodu, Dosen & Blanton, 2024 and references therein). The Powder Diffraction File in Relational Database (RDB) format contains extensive chemical, physical, bibliographic and crystallographic data including atomic coordinates enabling characterization and computational analysis. In this paper we describe the ICDD’s implementation of archiving raw data and its availability for use by analysts in materials characterization.

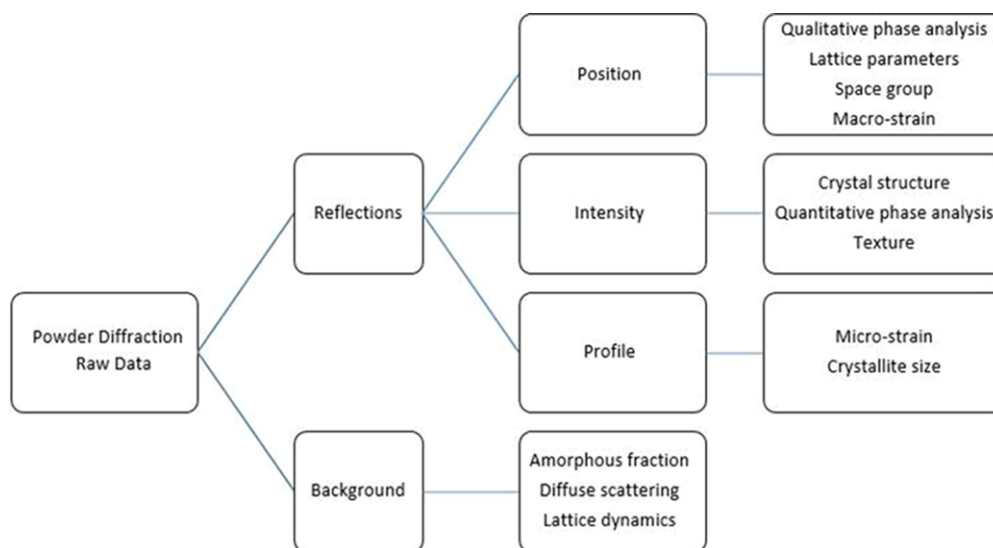
d	$\frac{I}{I_0}$	$\frac{I}{I_0}$	d in Å $\lambda = .708$	$\frac{I}{I_0}$	d in Å $\lambda = .708$	$\frac{I}{I_0}$
20.0	9.9	2.47	20.0	1.00		
	1.00	0.19	9.9	0.19		
		0.10	3.40	0.05		
			3.22	0.05		
			3.15	0.05		
<i>BiO<sub>2</sub> · C<sub>10</sub>H<sub>8</sub></i>			2.75	0.03		
<i>Bismuth beta-naphthal</i>			2.67	0.10		
			2.06	0.03		
			1.73	0.05		
			1.59	0.10		
			1.61	0.05		
			1.57	0.05		
			1.52	0.05		
			1.30	0.05		
			1.240	0.05		
Z =						
a =	b =	c =				
A =	C =		1.190	0.05		
D =						
n =	ω =	ε =				

**Figure 2**  
The first PDF card showing handwritten experimentally observed *d* spacings and relative intensities.

### 2.1. Archiving raw powder diffraction data

The interest in access to raw data and its publication for the diffraction user community is beneficial to the fields of crystallography and materials science, and the IUCr has been promoting this initiative, as has the ICDD for many years. The review papers aptly titled ‘Science in the Data’ (Helliwell *et al.*, 2017, and reference therein) and ‘Raw diffraction data and reproducibility’ (Kroon-Batenburg *et al.*, 2024) elegantly encapsulate the need for raw data in crystallography. These papers detail archiving diffraction data in chemical crystallography which is predominantly single crystal. The challenges in archiving powder diffraction raw data will be discussed in later sections.

Set 1 of the PDF published in 1941, on a 3 in × 5 in paper card (Fig. 2) was a listing of experimentally observed interplanar spacings (*d* spacings) and relative intensities (*I/I*<sub>0</sub>) characteristic of the compound. The ICDD first incorporated amorphous and poorly crystalline patterns in Set 38



**Figure 1**  
Information content of a powder diffraction pattern (adapted from Dinnebier & Billinge, 2008).

(published in 1979), represented as a drawing of an amorphous silicate. The conventional listing of *d*-spacing and relative intensities was no longer sufficient to detail the diffraction pattern aspects and for identifying these types of poorly crystalline phases. The powder diffraction data (whole pattern) was published as an image (Fig. 3) back then due to the limitation of an analog book format.

This was the genesis of the ICDD's effort to archive raw powder diffraction data. In 1984 supported by the ICDD, a round robin study conducted on systematic errors found

access to powder X-ray diffraction raw data was helpful in establishing the early guidelines for the deposition of raw data. (Schreiner & Fawcett, 1984). An outcome of this round robin was that the ICDD established a distinctive program called Grant-in-Aid (GiA) to acquire high-quality powder diffraction data of targeted materials. In the late 1980s, the ICDD emphasized GiA participants to deposit their raw powder diffraction data. The deposition of powder diffraction raw data is mandatory in the GiA program today. All new grantees are expected to submit powder diffraction raw data collected

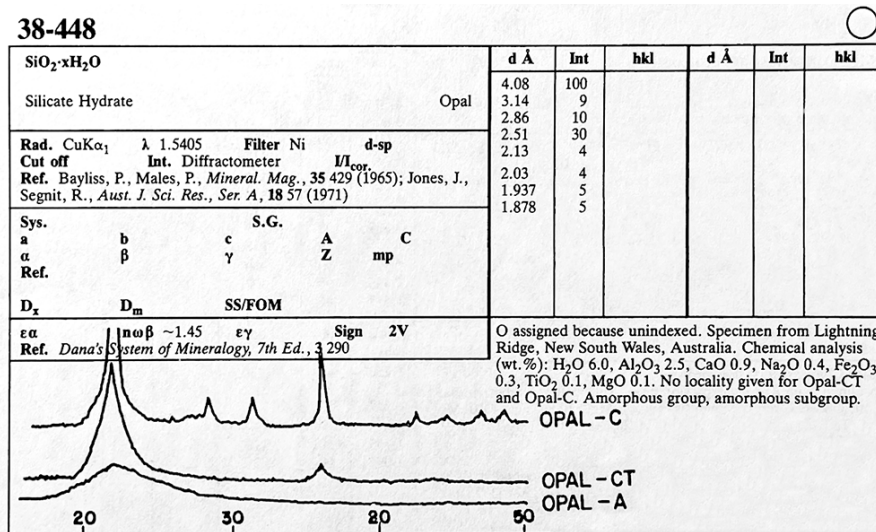


Figure 3 PDF card of the mineral opal with a PXR image to facilitate phase identification.

PDF #	QM	Chemical Formula	Compound Name	Mineral Name	Database Comments	Raw Diffraction Data
00-066-0177	● O	SiO <sub>2</sub> · xH <sub>2</sub> O	Silicon Oxide Hydrate	Opal	Color: Pink. Sample Preparation: The sample was manually ground into an agate mortar and sieved to obtain powder with a grain diameter <33 μm. The powders were then pressed and mounted. Sample Source or Locality: Specimen from Peru. Structures: Ordered structure. C-type opal, 100% cristobalite. Article Title: XRPD patterns of opals: a brief review and new results from recent studies. Temperature of Data Collection: 273 K. Warning: Unindexed pattern.	
00-066-0178	● O	SiO <sub>2</sub> · xH <sub>2</sub> O	Silicon Oxide Hydrate	Opal	Sample Preparation: The sample was manually ground into an agate mortar and sieved to obtain powder with a grain diameter <33 μm. The powders were then pressed and mounted. Raw Data Comment: Reflection at 3.34 Å from trace quartz impurity was omitted. Sample Source or Locality: Specimen from Australia. Article Title: XRPD patterns of opals: a brief review and new results from recent studies. Temperature of Data Collection: 273 K. Warning: Unindexed pattern.	
00-066-0179	● O	SiO <sub>2</sub> · xH <sub>2</sub> O	Silicon Oxide Hydrate	Opal	Sample Preparation: The sample was manually ground into an agate mortar and sieved to obtain powder with a grain diameter <33 μm. The powders were then pressed and mounted. Sample Source or Locality: Specimen from Mexico. Structures: CT-type fire opal. 29% cristobalite and 71% tridymite. Article Title: XRPD patterns of opals: a brief review and new results from recent studies. Temperature of Data Collection: 273 K. Warning: Unindexed pattern.	
00-066-0180	● O	SiO <sub>2</sub> · xH <sub>2</sub> O	Silicon Oxide Hydrate	Opal	Color: Green. Sample Preparation: The sample was manually ground into an agate mortar and sieved to obtain powder with a grain diameter <33 μm. The powders were then pressed and mounted. Sample Source or Locality: Specimen from Peru. Structures: CT-type opal. 45% cristobalite and 55% tridymite. Article Title: XRPD patterns of opals: a brief review and new results from recent studies. Temperature of Data Collection: 273 K. Warning: Unindexed pattern.	
00-066-0181	● O	SiO <sub>2</sub> · xH <sub>2</sub> O	Silicon Oxide Hydrate	Opal	Sample Preparation: The sample was manually ground into an agate mortar and sieved to obtain powder with a grain diameter <33 μm. The powders were then pressed and mounted. Sample Source or Locality: Specimen from Australia. Structures: CT-type fire opal. 15% cristobalite and 85% tridymite. Article Title: XRPD patterns of opals: a brief review and new results from recent studies. Temperature of Data Collection: 273 K. Warning: Unindexed pattern.	

Figure 4 Powder diffraction raw data of different opal minerals in PDF-5+ 2024.

on a NIST (National Institute for Standards and Technology, [www.nist.gov](http://www.nist.gov)) Standard Reference Material (SRM) in order to evaluate the diffraction data quality prior to the grant approval. More detailed information about this grant is available on the ICDD website (<https://www.icdd.com/grant-in-aid/>). Other sources of raw data found in the PDF are from published literature, author contributed and private communications. Active involvement of the scientific community is crucial in the development and growth of any scientific database. The ICDD has been encouraging scientists to deposit their powder diffraction data using a web portal (<https://www.icdd.com/data-submission/>). Compared to the earlier example (Fig. 3) of the mineral opal, Fig. 4 shows raw PXRD data of different opals in PDF-5+ Release 2024 illustrating the progress made in archiving and presentation of raw PXRD data.

### 2.2. Data curation

All of the data published in the Powder Diffraction File goes through a multi-tier editorial process. Each entry in the PDF has an editorially assigned quality mark (Kabekkodu, Dosen & Blanton, 2024). An editorial comment will describe the reason if an entry does not meet the top-quality mark. The editorial processes of the ICDD's quality management system are unique in that they are ISO 9001:2015 certified.

The challenges in archiving powder diffraction raw data are manifold due to phase impurities, data collection strategies, diffractometer geometry, sample preparation, particle statistics, systematic errors, and preferred orientation. These topics are covered in depth in *International Tables for Crystallography Volume H: Powder Diffraction* (2019) and in many reference books (Klug & Alexander, 1954; Jenkins & Snyder, 1996; Pecharsky & Zavalij, 2009; Dinnebier & Billinge, 2008). In our experience in reviewing submitted raw data about 30% of the raw data has issues that require contacting the authors. As an example, published raw data of sodium alginate [Fig. 6a in Chhatbar *et al.* (2009)] shows several sharp crystalline peaks, which is not characteristic of a typical polysaccharide biopolymer. Chhatbar *et al.* (2009) do mention washing sodium alginate with 0.25 M H<sub>2</sub>SO<sub>4</sub>, and further review of the data confirmed the highly crystalline phase to be sodium sulfate (Na<sub>2</sub>SO<sub>4</sub>, thenardite, PDF# 01-070-1541) overlapped with residual sodium alginate.

### 2.3. Importance of metadata

The FAIR (findability, accessibility, interoperability, and reusability) data principles (Wilkinson *et al.*, 2016) have been discussed widely in the scientific data archival effort in recent years. The role of metadata in terms of reproducibility of raw diffraction data were detailed in a recent publication by Kroon-Batenburg *et al.* (2024).

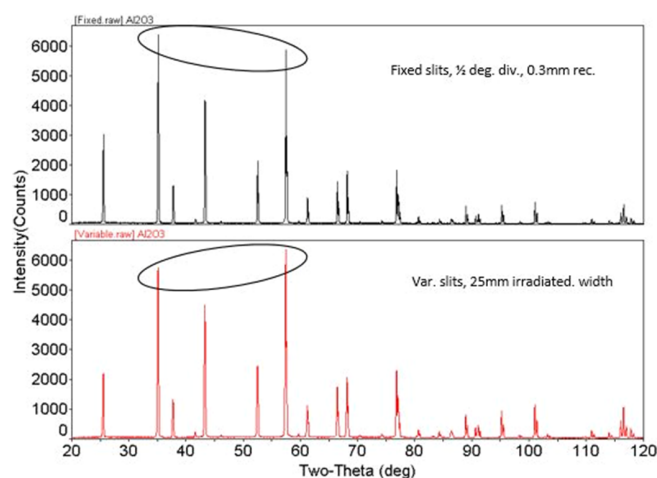
As discussed in the previous section, there are many factors that contribute to the powder diffraction raw data. It is extremely important to capture this information (where available) as metadata during the raw data archival. If one wishes to perform further analysis, for example, Rietveld

refinement (Rietveld, 1967, 1969) then an instrument parameter file is required. A very basic correction for the sample displacement (one of the most commonly seen problems in Bragg–Brentano geometry) requires knowledge of the goniometer radius. Fig. 5 is a simple example demonstrating the influence of fixed versus variable slits on a diffractometer, affecting the normalized relative intensities of the PXRD pattern of NIST SRM 1976 (Al<sub>2</sub>O<sub>3</sub>). Many of the published PXRD raw data deposited as an XY ASCII file and archival of such raw data may not be useful without proper metadata in terms of its reusability and reproducibility. Much of the powder diffraction data is collected using home laboratory diffractometers, and the format varies with across instrument manufacturers.

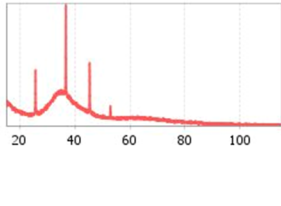
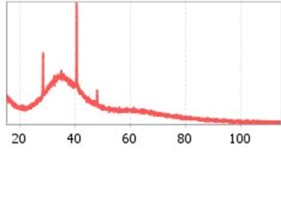
In some cases, the importance of metadata goes beyond the instrument configuration and data collection method adopted. Fig. 6 shows high-temperature polymorphs of silver sulfide, α-Ag<sub>2</sub>S (at 523.15 K) and γ-Ag<sub>2</sub>S (at 923.15 K). The authors (Blanton *et al.*, 2011) deposited metadata attributing the observed diffuse scattering to the highly disordered state of Ag<sup>+</sup> ions at high temperature resembling a liquid-like distribution. Such a description is crucial in using these types of raw data. The ICDD's online data submission tool enables metadata incorporation by prompting the authors to include them.

### 2.4. Raw data reusability

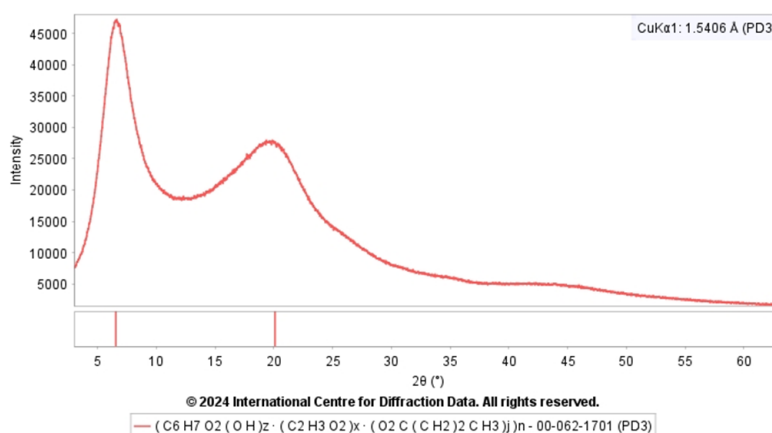
An example of raw data reusability is the determination of the crystal structure oftrandolapril (Reid *et al.*, 2016) using deposited raw data from the ICDD PDF. Most of the reusability of powder diffraction raw data is in the materials characterization where such data is a necessity, especially in characterizing poorly crystalline (clays, for example) or amorphous materials. It is evident from Fig. 7 that having raw data is essential to carry out phase identification in the case of poorly crystalline or amorphous patterns as they cannot be represented satisfactorily as an interplanar spacing (*d*) and relative intensity (*I*) list due to poorly resolved broad peaks.



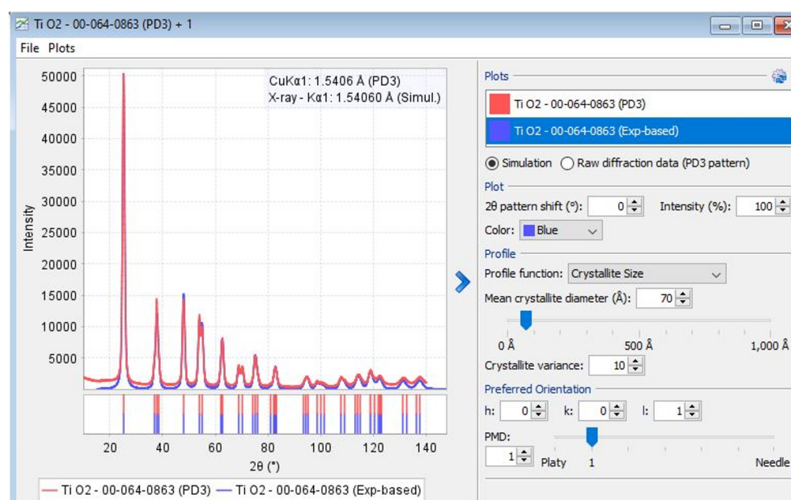
**Figure 5** Comparison of NIST SRM 1976 plate PXRD patterns fixed versus variable slits. Notice the difference in the relative intensities (marked in the figure).

PDF #	QM	Chemical Formula	Alternate Name	Temp (K)	Database Comments	Raw Diffraction Data
00-062-0248	● S	Ag <sub>2</sub> S		523.15	Rietveld refinement. Sample Preparation: Prepared by mixing 100 mL of deionized water with 12.845 g Ag N O <sub>3</sub> and 100 mL of deionized water with 9.067 g Na <sub>2</sub> S · 9 H <sub>2</sub> O into separate glass vessels. The solutions were combined into another glass vessel with 50 mL of deionized water immediately forming a black precipitate. After stirring for 10 minutes, the dispersion was poured into a vacuum filtration apparatus and the precipitate was washed with 1000 mL of deionized water before drying at 323 K for 3 hours. Phase Transition: The sample, when heated from room temperature (β-phase) to 523 K, undergoes a phase transition from monoclinic to body-centered cubic (α-phase) and then another phase transition to face-centered cubic at 923 K (γ-phase). Raw Data Comment: In situ high temperature XRD experiments were performed using a diffractometer equipped with a custom XRD furnace with measurements performed under pure nitrogen gas. Structures: The diffuse scattering seen in the α and γ phases is attributed to Ag <sup>+</sup> ions in a highly disordered state resembling a liquid-like distribution. Temperature of Data Collection: 523.15 K.	
00-062-0249	● S	Ag <sub>2</sub> S	γ-Ag <sub>2</sub> S	923.15	Color: Black. Processing Information: Rietveld refinement. Sample Preparation: Prepared by mixing 100 mL of deionized water with 12.845 g Ag N O <sub>3</sub> and 100 mL of deionized water with 9.067 g Na <sub>2</sub> S · 9 H <sub>2</sub> O into separate glass vessels. The solutions were combined into another glass vessel with 50 mL of deionized water immediately forming a black precipitate. After stirring for 10 minutes, the dispersion was poured onto a vacuum filtration apparatus and the precipitate was washed with 1000 mL of deionized water before drying at 323 K for 3 hours. Phase Transition: The sample, when heated from room temperature (β-phase) to 523 K, undergoes a phase transition from monoclinic to body-centered cubic (α-phase) and then another phase transition to face-centered cubic at 923 K (γ-phase). Raw Data Comment: In situ high temperature XRD experiments were performed using a diffractometer equipped with a custom XRD furnace with measurements performed under pure nitrogen gas. Structures: The diffuse scattering seen in the α and γ phases is attributed to Ag <sup>+</sup> ions in a highly disordered state resembling a liquid-like distribution. Temperature of Data Collection: 923.15 K.	

**Figure 6**  
Metadata describing the diffuse scattering in α-Ag<sub>2</sub>S (at 523.15 K) and γ-Ag<sub>2</sub>S (at 923.15 K).



**Figure 7**  
PXRD raw data of cellulose acetate butyrate (PDF# 00-062-1701).



**Figure 8**  
Estimation of the mean crystallite diameter of TiO<sub>2</sub> (anatase, nano). Simulation (blue) with 70 Å mean crystallite diameter matches satisfactorily with the raw data (red).

Whole raw data matching using a similarity index (Hofmann & Kuleshova, 2005) is one of the best methods to perform a search/match for poorly ordered phases. There are several examples where raw data is essential in characterizing pharmaceutical samples (Fawcett *et al.*, 2019), non-crystalline materials (Fawcett *et al.*, 2020a,b) and polymers (Gates *et al.*, 2014).

Raw powder diffraction patterns of nanocrystalline powders are crucial in describing the microstructural features such as crystallite size. PDF-5+ software allows users to overlay simulated PXRD patterns by varying the crystallite size (Scardi *et al.*, 2006) and is beneficial in estimating the mean crystallite diameter of nanocrystalline powders. This feature is applicable to both user imported as well as archived raw data in the database. As an example, the estimation of the crystallite size of nano anatase (PDF# 00-064-0863) is shown in Fig. 8.

### 3. Future

Raw data that is currently archived by the ICDD is in one-dimension (1D) format primarily due to the nature of data collection using traditional powder diffractometers. In the future, the ICDD will encourage 2D raw data submission which will be beneficial in characterizing materials with texture and preferred orientation. Archiving time-of-flight neutron powder diffraction patterns is currently in development.

### 4. Conclusion

The availability of powder diffraction raw data in a well curated database plays a pivotal role in any successful material characterization or data driven study. The upcoming Release 2025 of the Powder Diffraction File will have more than 20 800 powder diffraction raw data entries. The exponential growth and interest in data-driven research based on machine learning and artificial intelligence make it critical to have a database with reliable data curation. One of the main problems in archiving powder diffraction raw data is the lack of a common format that would encapsulate parameters that have a strong influence on the observed powder diffraction pattern. Scientific journals encouraging authors to submit powder diffraction raw data as powder CIF ([https://www.iucr.org/resources/cif/dictionaries/cif\\_pd](https://www.iucr.org/resources/cif/dictionaries/cif_pd)) along with required metadata will be beneficial for the materials science community.

### Acknowledgements

Editing, curating and producing a quality database and software requires a significant team effort by the ICDD staff, members and editors. We thank the hundreds of researchers and scientists who have contributed to the Powder Diffraction File and the ICDD organization over the past 80+ years.

### References

- Blanton, T., Misture, S., Dontula, N. & Zdzieszynski, S. (2011). *Powder Diffr.* **26**, 114–118.
- Chhatbar, M., Meena, R., Prasad, K. & Siddhanta, A. K. (2009). *Carbohydr. Polym.* **76**, 650–656.
- Dinnebier, R. E. & Billinge, S. J. L. (2008). *Powder Diffraction Theory and Practice*, edited by R. E. Dinnebier & S. J. L. Billinge. Cambridge, UK: RSC Publishing.
- Fawcett, T. G., Gates-Rector, S., Gindhart, A., Rost, M., Kabekkodu, S. N., Blanton, J. R. & Blanton, T. N. (2020a). *Powder Diffr.* **34**, 130–142.
- Fawcett, T. G., Gates-Rector, S., Gindhart, A. M., Rost, M., Kabekkodu, S. N., Blanton, J. R. & Blanton, T. N. (2019). *Powder Diffr.* **34**, 164–183.
- Fawcett, T. G., Gates-Rector, S., Gindhart, A. M., Rost, M., Kabekkodu, S. N., Blanton, J. R. & Blanton, T. N. (2020b). *Powder Diffr.* **35**, 82–88.
- Gates, S. D., Blanton, T. N. & Fawcett, T. G. (2014). *Powder Diffr.* **29**, 102–107.
- International Tables for Crystallography* (2019). Vol. H, *Powder diffraction*, 1st online ed., edited by C. J. Gilmore, J. A. Kaduk, J. A. & H. Schenk. <https://doi.org/10.1107/97809553602060000115>.
- Helliwell, J. R., McMahon, B., Guss, J. M. & Kroon-Batenburg, L. M. J. (2017). *IUCrJ*, **4**, 714–722.
- Hofmann, D. W. M. & Kuleshova, L. (2005). *J. Appl. Cryst.* **38**, 861–866.
- Jenkins, R. & Snyder, R. (1996). *Introduction to X-ray Powder Diffraction*. New York: Wiley-Interscience.
- Kabekkodu, S. N., Dosen, A. & Blanton, T. (2024). *Powder Diffr.* **39**, 47–59.
- Kaduk, J. A. (2019). In *International Tables for Crystallography*, Vol. H, *Powder diffraction*, 1st online ed., edited by C. J. Gilmore, J. A. Kaduk & H. Schenk, ch. 3.7, pp. 304–324, <https://doi.org/10.1107/97809553602060000952>.
- Klug, H. P. & Alexander, L. E. (1954). *X-ray Diffraction Procedures for Polycrystalline and Amorphous Materials*. New York: Wiley-Interscience.
- Kroon-Batenburg, L. M. J., Lightfoot, M. P., Johnson, N. T. & Helliwell, J. R. (2024). *Struct. Dyn.* **11**, 011301.
- Messick, J. (2012). *Powder Diffr.* **27**, 36–44.
- Pecharsky, V. J. & Zavalij, P. Y. (2009). *Fundamentals of Powder Diffraction and Structural Characterization of Materials*, 2nd ed. New York: Springer.
- Reid, J. W., Kaduk, J. A. & Vickers, M. (2016). *Powder Diffr.* **31**, 205–210.
- Rietveld, H. M. (1967). *Acta Cryst.* **22**, 151–152.
- Rietveld, H. M. (1969). *J. Appl. Cryst.* **2**, 65–71.
- Scardi, P., Leoni, M. & Faber, J. (2006). *Powder Diffr.* **21**, 270–277.
- Schreiner, W. N. & Fawcett, T. (1984). *Adv. X-ray Anal.* **28**, 309–314.
- Smith, D. K. (2001). *Powder Diffr.* **16**, 186–191.
- Whitfield, P. A., Huq, A. & Kaduk, J. A. (2019). In *International Tables for Crystallography*, Vol. H, *Powder diffraction*, 1st online ed., edited by C. J. Gilmore, J. A. Kaduk & H. Schenk, ch. 2.10, pp. 200–222, <https://doi.org/10.1107/97809553602060000945>.
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., Gonzalez-Beltran, A., Gray, A. J. G., Groth, P., Goble, C., Grethe, J. S., Heringa, J., 't Hoen, P. A. C., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S. J., Martone, M. E., Mons, A., Packer, A. L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M. A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J. & Mons, B. (2016). *Sci. Data*, **3**, 160018.