Powder Diffraction raw data
Diffraction data

Less and less information easily available:

Single-crystal diffraction
multiple images collected

2D Powder diffraction
single image collected

Powder diffraction after integration
Information loss! (position, overlap etc.)
Powder Diffraction – if so much information is lost, why still used? When is it preferable / unavoidable?

• To represent all the sample,
• To identify the phases in the sample, in particular, to analyse phase composition of multi-component samples (qualitatively and quantitatively),
• to measure the unit cell parameters and volume and their variations with temperature, pressure, as functions of the composition of solid solutions, concentration of defects, strain and stress,
• to estimate the size of crystallites, domains,
• to find the preferred orientation of particles,
• to estimate the content of the amorphous phase(s) in relation to the crystalline phase(s),
• to solve an unknown crystal structure, if a single crystal cannot be obtained,
• to refine a crystal structure that has been preliminary roughly solved based on poor-quality or low-completeness single-crystal diffraction data.
Powder Diffraction: not only peak positions and integral intensities

(a) a perfect single crystal;
(b) a small number (Nc = 10) of domains with uniform pole distribution;
(c) a large number (Nc = 1000) of domains, again with uniform pole distribution;
(d) Nc = 1000 with anisotropic pole distribution.

Smaller sized datasets, less information

Working well for measurements/experiments requiring statistical analysis; e.g. Quantification

**Fitting and refinement are prone to end in false minima**

More often one has to go back and look again at raw data, sometimes after years

Data should then not simply be a sequence of angles and intensities, but include all aspects necessary to interpretation (metadata).

Not only the **angular positions** of the diffraction maxima and their **integral intensities**, but also the **shapes** of the reflections, the **background**, the presence of some **non-indexed reflections**, the **2θ range** measured are important

**Raw data are incomplete without metadata**
Raw diffraction data, processing

Data processed by most used programs are in \(2\theta-I\) or better \(2\theta-I-\sigma(I)\) format

Often though this is not the lowest level of data:

Angular calibration and flatfield
Data processing and metadata

A bare minimum metadata is the wavelength, or wavelength spread.

In many cases, though, data could be of limited value unless the full Instrumental Resolution Function (IRF) is known. This is needed to extract higher level information, beyond the simple fingerprinting. Should normally be kept as metadata.

2D -> 1D, effect of correct angular calibration (angular bias) and area masking.

Wrong detector flatfield may also lead to non-statistical noise.

Information on processing is essential to re-evaluate old data. It is also often a major source of error and after re-processing correctly better information can be extracted.
Things that enter data ...

Fluorescence: pattern with no fluorescence in green, in red the pattern with fluorescence, including a higher background and non-statistical noise.

Sample holder size: Fructose [1 1 0], blue 0.3 mm capillary, green 0.7 mm capillary, red 0.7 mm capillary with 0.3 mm slits (slits have a wrong offset, resulting in a different angular position)

Temperature, sample origin, sample holder, these are simple parameter to record and keep. Many things may influence the final pattern which instead we cannot fully control, composition may lead to fluorescence and may not be known at the beginning, radiation sensitivity may not be known either and could affect the pattern in long measurements (or in synchrotron ones)
Particle Statistics...

The level of spinning, rocking the sample should be recorded, not always a 2D image is available to re-evaluate the quality of the statistics.
No alternative to keeping all images

synchrotron top up, as perturbation (less intensity) every 3 minutes
Keeping sample photographs for non-homogeneous / varying samples

Masking is essential for processing high-pressure data
Innovation

Why re-visit old data? A case study

Rotation of H-bonded fragments (phase transitions in L-serine, OH...O bonds)

E. Boldyreva, H. Sowa et al., 2005, 2006; S. Moggach, S. Parsons et al., 2005, 2006
L-serine III, 8 GPa

Single crystal data / Powder data

Influence of compression rate on phase transitions

L-serine I -> L-serine II: slow nucleation, fast nuclei growth;

L-serine I -> L-serine IV: faster nucleation, very slow nuclei growth

L-serine III-IV, 8 GPa

Powder data from 2005 re-processed in 2015
Instead of conclusions

Data is not information,
Information is not knowledge,
Knowledge is not understanding,
understanding is not wisdom.

C. Stoll