26TH CONGRESS AND GENERAL ASSEMBLY OF THE INTERNATIONAL UNION OF CRYSTALLOGRAPHY

https://iucr2023.org

Andy Götz (ESRF) **for LEAPS-WG3 13/9/2023**

PaNOSC and ExPaNDS projects have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements 823852 and 857641, respectively.

Talk outline

1. What is the IUCr?

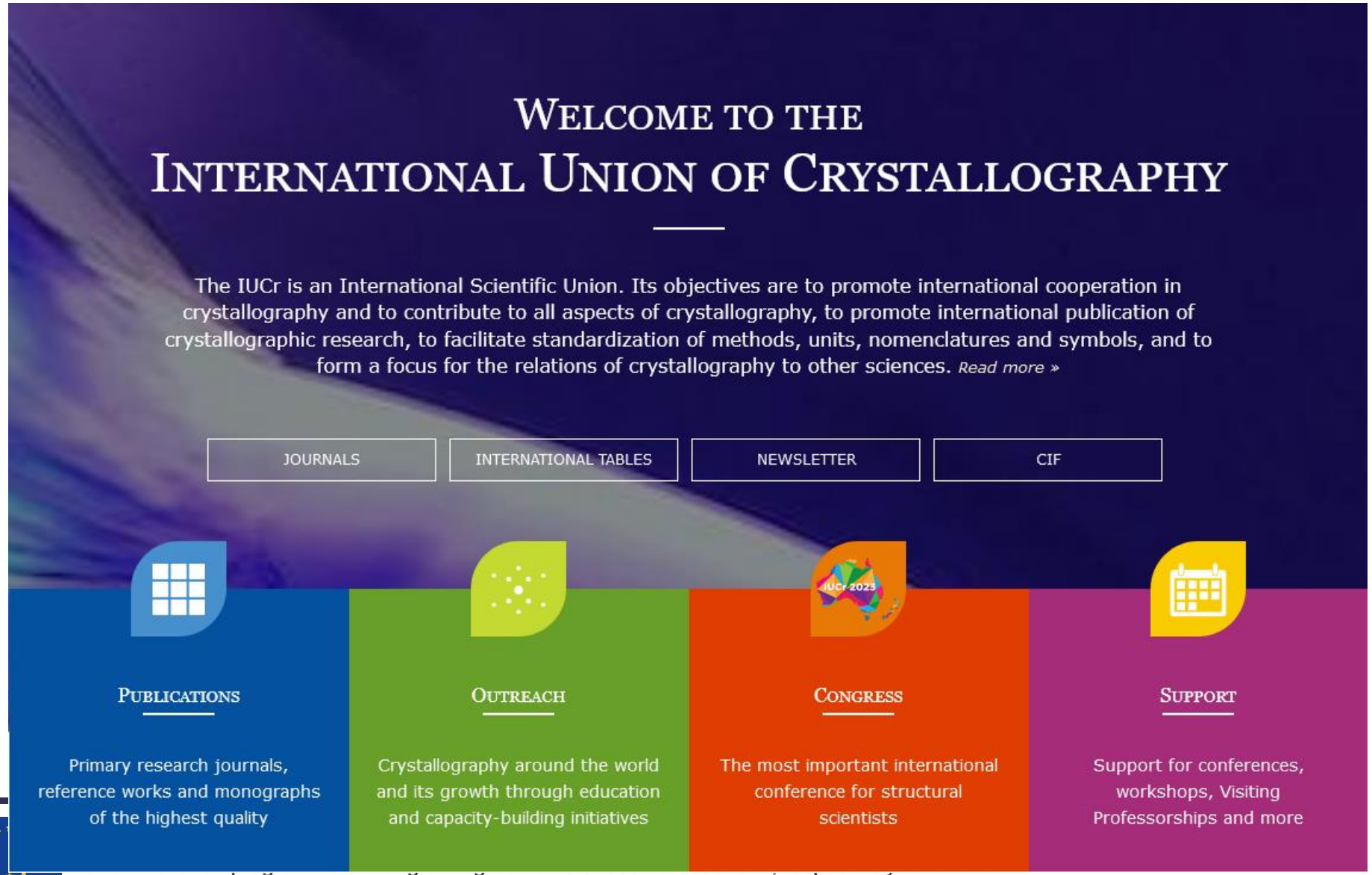
2. Sessions on data management and processing

3. Some impressions from the conference

4. Next IUCr 2026

PaNOSC and ExPaNDS projects have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements 823852 and 857641, respectively.

The IUCr is the learned society representing crystallographers around the world



https://iucr.org

Committee on Data (CommDat)

Membership

- J.R. Helliwell (Chair, UK; john.helliwell@manchester.ac.uk)
- B. McMahon (Secretary, UK)
- H.J. Bernstein (USA)
- A. Brink (South Africa)
- I. Bruno (UK)
- S. Coles (UK)
- K. Dziubek (Poland)
- A. Goetz (France)
- S. Kabekkodu (USA)
- L.M.J. Kroon-Batenburg (The Netherlands)
- G. Kurisu (Japan)
- W. Minor (USA)
- S. Storm (Germany)
- L. Van Meervelt (Belgium)
- J. Hester (Australia) (COMCIFS liaison)

Consultants

- S. Androulakis (Australia)
- M.P. Blakeley (France)
- G. Bricogne (UK)
- S. Grazulis (Lithuania)
- B. Matthews (UK)
- A. Sarjeant (USA)
- D. Szebenyi (USA)
- E.F. Weckert (Germany)
- J. Trewhella (Australia)

https://www.iucr.org/iucr/governance/advisory-committees/committee-on-data

PaNOSC and ExPaNDS projects have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements 823852 and 857641, respectively.

Workshop on *Raw diffraction data reuse: the good, the bad and the challenging*

Organized by

Loes Kroon-Batenburg (Netherlands), Selina Storm (Germany), John Helliwell (UK) and Brian McMahon (UK) for the IUCr Committee on Data

https://www.iucr.org/resources/data/commdat/melbourne-workshop

CIF Dictionary Workshop

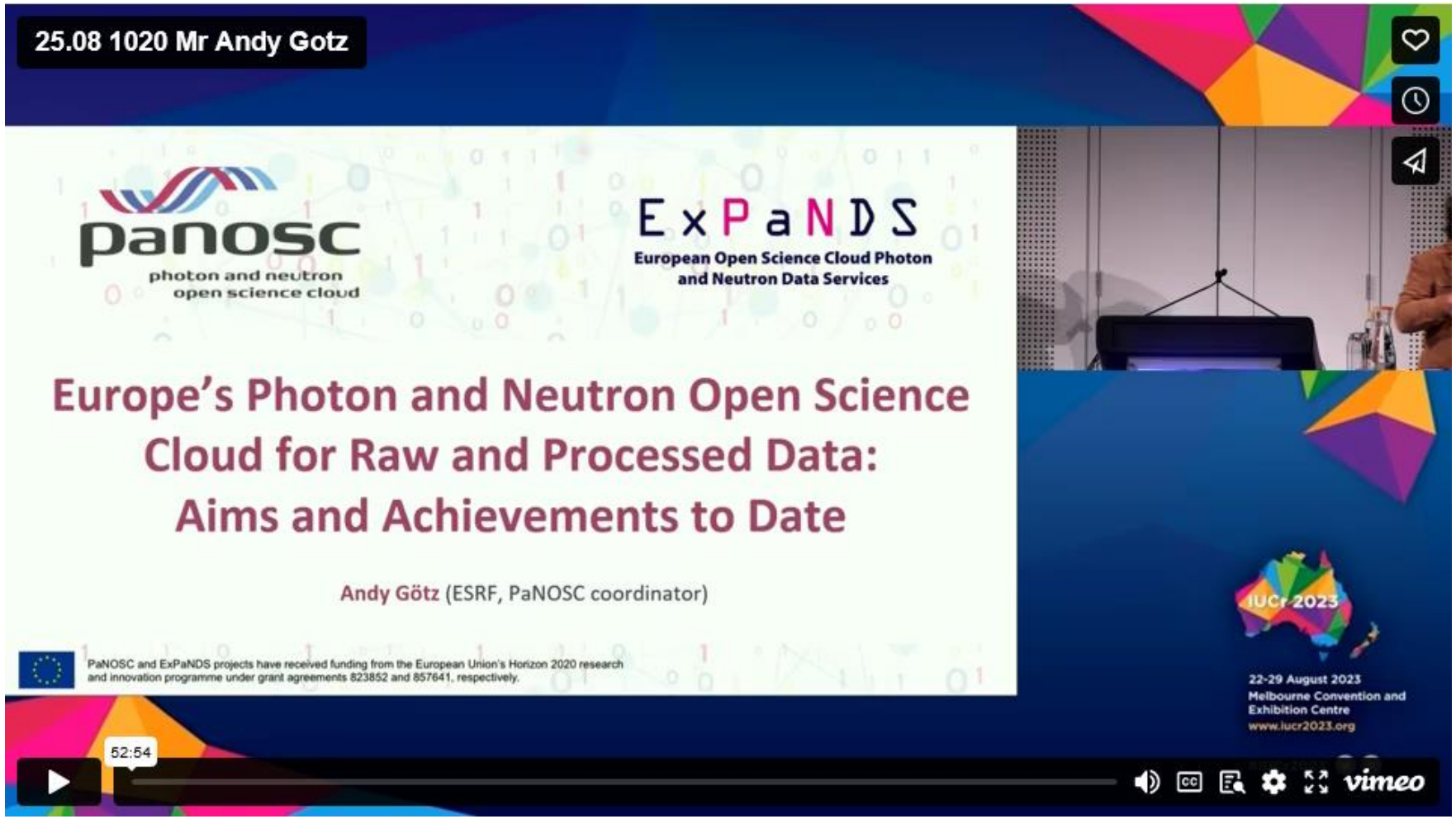
Session 1

TODAY - Friday
at 1:10 pm – 3:30 pm

Room: 216	Room: 220	Room: 207
Chairs: Loes Kroon-Batenburg & James Hester	**Chairs: Nicholas Schwarz & Andrew Gotz**	**Chairs: Brinda Vallat & John Helliwell**
Session: A118 Raw Diffraction Data Reuse: Warts and All	**Session: A109** Data-Driven Science: Current Status and Outlook	**Session: A014** Databases and Data Management

PaNOSC and ExPaNDS projects have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements 823852 and 857641, respectively.

Keynote address on PaNOSC + ExPaNDS



https://vimeo.com/862997399

PaNOSC and ExPaNDS projects have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements 823852 and 857641, respectively.

Committee on Data workshop

Session 1: Facility and raw data archive providers Part I

Chair: John R. Helliwell

08:20-08:30 (*00:20-00:30*)	**Chair**	Opening remarks	
08:30-08:50 (*00:30-00:50*)	**Andreas Moll** (Australia)	Scientific computing and data management at the Australian Synchrotron	**Abstract** | **Presentation** (2.4 MB)
08:50-09:10 (*00:50-01:10*)	**Anton Barty** (Germany)	Managing and curating data flows at PETRA IV	**Abstract** | **Presentation** (5.8 MB)
09:10-09:30 (*01:10-01:30*)	* **Bridget Murphy** (Germany)	DAPHNE4NFDI: DAta from PHoton and Neutron Experiments for NFDI	**Abstract** | **Presentation** (3.2 MB)
09:30-09:50 (*01:30-01:50*)	**Andy Götz** (France)	Making the most of data from the ESRF	**Abstract** | **Presentation** (3.4 MB)

https://www.iucr.org/resources/data/commdat/melbourne-workshop/programme#ab

PaNOSC and ExPaNDS projects have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements 823852 and 857641, respectively.

Committee on Data workshop

Session 2: Facility and raw data archive providers Part II

Chair: * Selina Storm

10:10-10:40 (*02:10-02:40*)	**Alun Ashton** (Switzerland)	Scientific computing, data sharing and reuse at PSI	**Abstract** | **Presentation** (4.0 MB)
10:40-11:00 (*02:40-03:00*)	**Genji Kurisu** (Japan)	X-tal Raw Data Archive (XRDa): A crystallographic raw diffraction image archive in Asia	**Abstract** | **Presentation** (3.4 MB)
11:00-11:20 (*03:00-03:20*)	**Fabio Dall'Antonia** (Germany)	Handling of big data at the European XFEL	**Abstract** | **Presentation** (6.3 MB)
11:20-11:40 (*03:20-03:40*)	**Wladek Minor** (USA)	A subject specific repository for MX (proteindiffraction.org)	**Abstract** | **Presentation** (4.3 MB)
11:40-12:00 (*03:40-04:00*)	**Alexandra Tolstikova** (Germany)	Processing data in serial crystallography on-the-fly: what kind of raw data do we want to store?	**Abstract** | **Presentation** (2.1 MB)

https://www.iucr.org/resources/data/commdat/melbourne-workshop/programme#ab

PaNOSC and ExPaNDS projects have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements 823852 and 857641, respectively.

Committee on Data workshop

Session 3: Raw data reusers

3.1: Macromolecular crystallography

Chair: John R. Helliwell

12:40-13:00 (*04:40-05:00*)	**Gerard Bricogne** (UK)	The raw, the cooked and the medium-rare: unmerged diffraction data as a rich source of opportunities for data re-use and improvements in both methods and results	**Abstract** | **Presentation** (3.2 MB)
13:00-13:20 (*05:00-05:20*)	**David Aragao** (UK)	Experiences with MX data reuse at Diamond	**Abstract** | **Presentation** (4.1 MB)
13:20-13:40 (*05:20-05:40*)	**Eugene Krissinel** (UK)	Raw data reuse: what it means for CCP4	**Abstract** | **Presentation** (3.8 MB)
13:40-14:00 (*05:40-06:00*)	* **Melanie Vollmar** (UK)	Reusing raw data for machine learning in MX	**Abstract** | **Presentation** (2.9 MB)

https://www.iucr.org/resources/data/commdat/melbourne-workshop/programme#ab

PaNOSC and ExPaNDS projects have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements 823852 and 857641, respectively.

Committee on Data workshop

3.2: Chemical crystallography

Chair: Loes Kroon-Batenburg

14:20-14:40 (*06:20-06:40*)	**Jim Britten** (Canada)	Use of raw data for diffraction space visualization: What are we missing in an integrated HKL file?	**Abstract** | **Presentation** (2.0 MB) | **Powerpoint (with embedded video)** (225 MB)
14:40-15:00 (*06:40-07:00*)	**Simon Coles** (UK)	The increasing diversity of small molecule data: can one size fit all?	**Abstract** | **Presentation** (2.2 MB)

3.3: Powder diffraction

Chair: Loes Kroon-Batenburg

15:00-15:20 (*07:00-07:20*)	**Nicola Casati** (Switzerland) and * **Elena Boldyreva** (Russia)	Powder diffraction raw data	**Abstract** | **Presentation** (2.9 MB) | **Video presentation** (39.7 MB)
15:20-15:40 (*07:20-07:40*)	**Miguel A. G. Aranda** (Spain)	Powder diffraction data sharing and reuse: advantages and possible practical obstacles	**Abstract** | **Presentation** (0.5 MB)
15:40-15:55 (*07:40-07:55*)	**Loes Kroon-Batenburg** (Netherlands) / * **Selina Storm** (Germany)	Summing up: the role of *IUCrData*'s new *Raw Data Letters* in serving all the above	***Raw Data Letters* (LKB)** (0.5 MB) | **Workshop summary (SS)** (0.1 MB)

Session on Data-Driven Science

Chairs: Nicholas Schwarz & Andrew Gotz	**Dr Max Nanao** *Invited Speaker 1*
Session: A109 Data-Driven Science: Current Status and Outlook	**#1385 - Dr Masaki Yamamoto** *Contributed talk 3*
#94 - Dr Filip Leonarski *Contributed talk 1*	**#1651 - Dr Stefan Brandstetter** *Contributed talk 4*
#534 - Dr Yasumasa Joti *Contributed talk 2*	**Prof Ian Foster** *Invited Speaker 2*

PaNOSC and ExPaNDS projects have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements 823852 and 857641, respectively.

IUCr Journals has launched *IUCrData's Raw Data Letters*

Scientists are encouraged to publish raw data!

panosc
photon and neutron
open science cloud

ExPaNDS
European Open Science Cloud Photon and Neutron Data Services

IUCrData
ISSN 2414-3146

Crystal structure of the second extracellular domain of human tetraspanin D9: twinning and diffuse scattering

Viviana Neviani, Martin Lutz, Wout Oosterheert, Piet Gros and Loes Kroon-Batenburg*

Department of Chemistry, Structural Biochemistry, Bijvoet Centre for Biomolecular Research, Faculty of Science, Utrecht University, Utrecht, The Netherlands. *Correspondence e-mail: l.m.j.kroon-batenburg@uu.nl

Received 20 April 2021
Accepted 1 May 2021

Keywords: twinning; diffuse scattering; tetraspanin $CD9_{EC2}$.

Remarkable features are reported in the diffraction pattern produced by a crystal of tetraspanin $CD9_{EC2}$, the structure of which was described previously [Oosterheert *et al.* (2020). *Life Sci. Alliance*, **3**, e202000883]. $CD9_{EC2}$ crystallized in space group *P*1 and was twinned. Concurrent with the twinning, diffuse streaks were seen in the direction perpendicular to the twinning interface. Preliminary conclusions are made on packing disorder and potenti... implications for the observed molecular structure. It is envisa... diffraction images could be very useful for method... remove the diffuse scattering to extract ... to model the effect of packin...

Raw diffraction data
HDF5 data file, DOI: https://doi.org/10.5281/zenodo.1234567
Metadata ImgCIF file, DOI: https://doi.org//10.1107/S2414314622000384/me6134.cif

[CheckCif for Raw Data]

checkImgCIF report

```
ImgCIF checker version 2022-07-16

Checking block 5886687 in ...

Running check...
======

... PASS

...: All frames present: PASS
   All frames present and correct for SCAN1

Testing: Detector surface axes used properly: PASS

Testing: Pixel size and origin described correctly: ...

Testing: Check calculated beam centre: PASS

Testing: Check principal axis is aligned with X: PAS...
Testing presence of archive:

Testing: All archives are accessible: PASS

Running checks with downloaded images
=========================================

Testing image 4: Image type and dimensions: PASS
```

Raw data table generated from the CIF

Raw data	
DOI	https://doi.org/10.5281/zenodo.5886687
Data archive	Zenodo
Data format	
... source	
	-287.22
Pixel size (mm)	0.075 × 0.075
No. of pixels	4148 × 4362
No. of scans	1
Exposure time per frame (s)	
Scan axis	ω, X
Start angle, increment per frame (°)	0.0, 0.1
Scan range (°)	360.0
No. of frames	3600

... Uni...
...espe...

Contact Fabio Dall'Antonio (XFEL) for more info and help!

Impressions

1. Photon and Neutron sources are the main sources of data for cutting edge crystallography
2. Faster and better detectors are coming e.g. Pilatus4, Citius, CryoEM tomography
3. PDB is the gold standard for data, there is an opportunity for working closer together – great keynote by RCSB director “You have got to love the PDB”!
4. Positive feedback on PaNOSC+ExPaNDS activities from a number of players e.g. Hercules, IUCr editors, DLS, Spring8, scientists

PaNOSC and ExPaNDS projects have received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreements 823852 and 857641, respectively.

Feedback

1. The PaNOSC Data Commons of FAIR data has sparked real interest in the crystallography community
2. Raw data from our facilities are currently not FAIR, we have offers to help validate raw data cited in the PDB
3. LEAPS facilities are asked to provide data in formats which can be published in the Raw Data Letters
4. LEAPS facilities are encouraged to propose data managers who can participate in the IUCr Committee on Data

Q: are all LEAPS+LENS facilities still committed to FAIR data and the next phase of PaNOSC in the future?

PaNOSC and ExPaNDS projects have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements 823852 and 857641, respectively.

Next IUCr General Assemblies

IUCr 2026 in Calgary

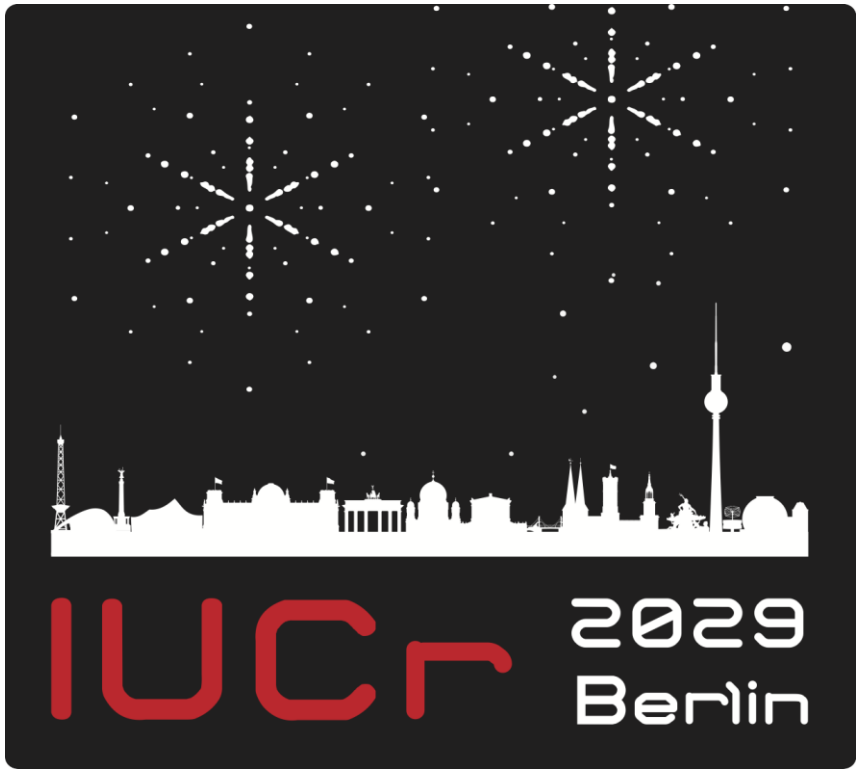
EMBRACING STRUCTURAL SCIENCE

IUCr 2026

CALGARY, CANADA

AUGUST 11-18

IUCr 2029 in Berlin



PaNOSC and ExPaNDS projects have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements 823852 and 857641, respectively.