Metadata for raw data from X-ray diffraction and other structural techniques

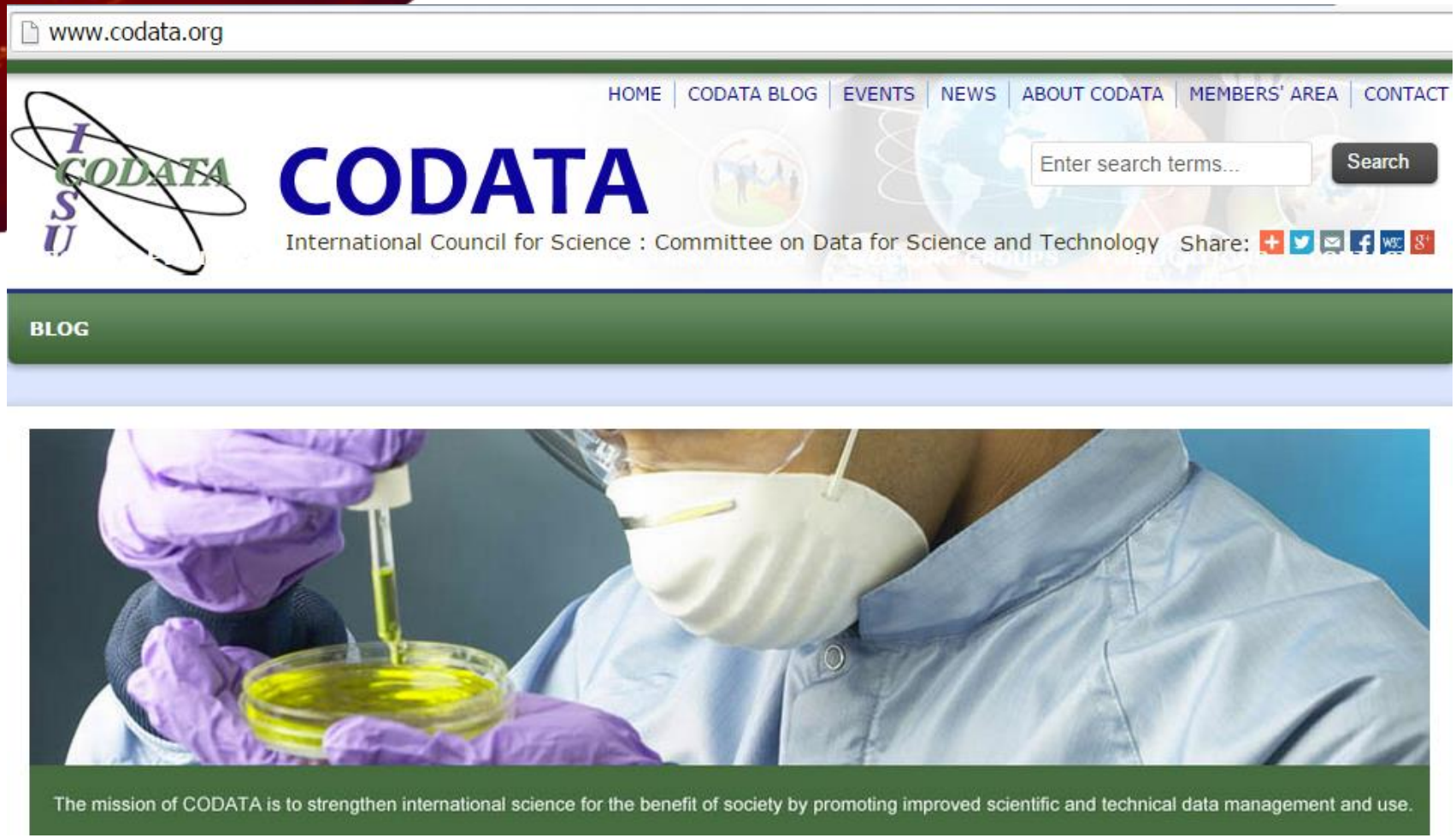
A Satellite Workshop to the 29th European Crystallographic Meeting

CODATA and (meta)data characterisation in the wider world

Brian McMahon

International Union of Crystallography
5 Abbey Square
Chester CH1 2HU
UK
bm@iucr.org

CODATA – ICSU Committee on Data for Science and Technology



http://www.codata.org

CODATA – ICSU Committee on Data for Science and Technology

Some facts

- Founded 1966
- Scientific Committee of ICSU
- Mission: to strengthen international science for the benefit of society by promoting improved scientific and technical data management and use.
- Remit: CODATA is concerned with all types of quantitative data from experimental measurements/observations in the physical, biological, geological and astronomical sciences. Particular emphasis is given to data management problems common to different scientific disciplines and to data used outside the field in which they were generated.
- General objectives: improvement of the quality and accessibility of data, as well as the methods by which data are acquired, managed and analysed; facilitation of international cooperation among those collecting, organizing and using data; promotion of an increased awareness in the scientific and technical community of the importance of these activities.

CODATA – ICSU Committee on Data for Science and Technology

Some facts

- Headquarters: Paris, France
- Members: National Members, International Scientific Unions, Co-Opted Organizations, Supporting Organizations
- Officers: *President* Geoffrey Boulton (UK); *Secretary General* Sara Graves (USA); *Treasurer* John Broome (Canada)
- Executive Director: Simon Hodson (UK)
- Task Groups (*e.g.* Data Citation Standards and Practices; Interoperable Data Publications; Physical Constants of Nature)
- Working Groups: Description of Nanomaterials (CODATA/VAMAS)
- Biennial CODATA Prize: Syd Hall *'for outstanding achievement in the world of scientific and technical data'*
- IUCr Representatives: J. R. Helliwell (2012-); B. McMahon (2000-2012)

CODATA/VAMAS Working Group on Description of Nanomaterials

- Convened 2012 under auspices of CODATA and VAMAS (Versailles Project on Advanced Materials and Standards)
- Chairs: John Rumble (formerly NIST), Steve Freiman, Clayton Teague
- Workshops
  - Paris February 2012 (IUCr: Reinhard Neder)
  - Paris May 2013 (IUCr; Daniel Chateigner)
  - North Carolina
  - Paris April 2014
  - Beijing September 2014
  - Maastricht July 2015
- Uniform Description System for Materials on the Nanoscale v1.0 published February 2015

http://www.codata.org/nanomaterials

CODATA/VAMAS Working Group on Description of Nanomaterials

IUCr participation in the CODATA / VAMAS Working Group on Nanomaterials

Daniel Chateigner
CRISMAT-ENSICAEN
Université de Caen Basse-Normandie
6 Bd. Maréchal Juin
14050 Caen, France
daniel.chateigner@ensicaen.fr

John R. Helliwell
School of Chemistry
University of Manchester
Oxford Road
Manchester M13 9PL, UK
john.helliwell@manchester.ac.uk

Reinhard Neder
Physics Department
University of Erlangen
Staudtstr. 3
D-91058 Erlangen, Germany
reinhard.neder@krist.uni-erlangen.de

Introduction

Nanotechnology is moving towards commercialization, yet numerous scientific questions remain unanswered. One of the most critical challenges is that there is no common nomenclature or description system for nanomaterials that is accepted by a single discipline, let alone by all disciplines. The International Union of Crystallography (IUCr) and its Committee for the Maintenance of the CIF Standard (COMCIFS) are well matched to help with this task set up under the aegis of CODATA/VAMAS.

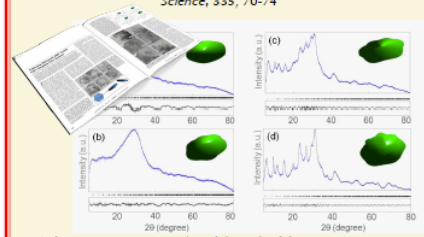
IUCr was invited to participate in a CODATA/VAMAS Working Group in 2012. CODATA is the Committee on Data for Science and Technology of ICSU, the International Council for Science (http://www.codata.org). VAMAS is the Versailles Project on Advanced Materials and Standards (http://www.vamas.org). There have been 3 Workshops so far.

The overall goal is to define the needs of as many scientific disciplines and user communities as possible. This White Paper will be transmitted to ISO Technical Committee 229 on Nanotechnology (http://www.iso.org/iso/iso_technical_committee?commid=381983) as well as other international and national standards development bodies and government agencies. Within this Working Group, crystallographers unambiguously define and characterize the structure of nanomaterials.

Furthermore X-ray reflectivity is suitable when the nano character is only expressed along one direction (stacks, films). Compared to other fields of science, Crystallography appears as a potential leader of quantitative nanomaterial descriptors and definitions. We bring different perspectives to the task including the physical and biological sciences relevant to inorganic, organic and bio nanomaterials.

Capturing Ultrasmall EMT Zeolite from Template-Free Systems

E.-P. Ng, D. Chateigner, T. Bein, V. Valtchev & S. Mintova (2012). *Science*, 335, 70-74



cited in *Science* Magazine 'Breakthrough of the Year 2011'

Crystallographic Techniques

Diffraction (scattering + interference): X-ray, gamma radiation, neutrons, electrons

Reflectivity (specular, off-specular): X-rays and neutrons

Small-angle scattering: X-rays (SAXS), neutrons (SANS)

Tomography (absorption or phase contrasts): X-rays, neutrons, electrons

Imaging (HRTEM)

Spectroscopy: X-ray (XRF, XANES, EXAFS, DAFS)
electrons (EDS)
muons (μSR)
photoluminescence
dynamic light scattering (DLS)

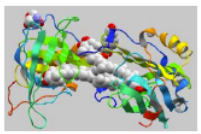
Crystallographic Information Framework (CIF)

CIF dictionaries provide a formal taxonomy of crystallographic terms and ideas. Dictionary entries are constructed in a structured machine-readable manner that facilitates validation and structuring of data: http://www.iucr.org/resources/cif/dictionaries

Dictionaries: Core, Restraints, Powder, Modulated-Composite, Electron Density, Twinning, Macromolecular, Images, Symmetry

Local dictionaries: reflectivity (to come), MPOD (properties), MAUD

Electron, neutron and X-ray scattering and diffraction, at small and wide angles, and imaging techniques offer a physically grounded determination of the coherent size domains (including crystal shape). Our tools can approach the physical state of nanoaggregates.



Acknowledgement

Brian McMahon (IUCr, Chester, UK) is thanked for discussions and assistance with this poster.

- Summary of IUCr contributions presented as a poster to the ACA 2014 Annual Meeting

http://www.ecole.ensicaen.fr/~chateign/danielc/posters/Chateigner_poster_ACA2014_CODATA%20VAMASv3.pdf

CODATA/VAMAS Working Group on Description of Nanomaterials

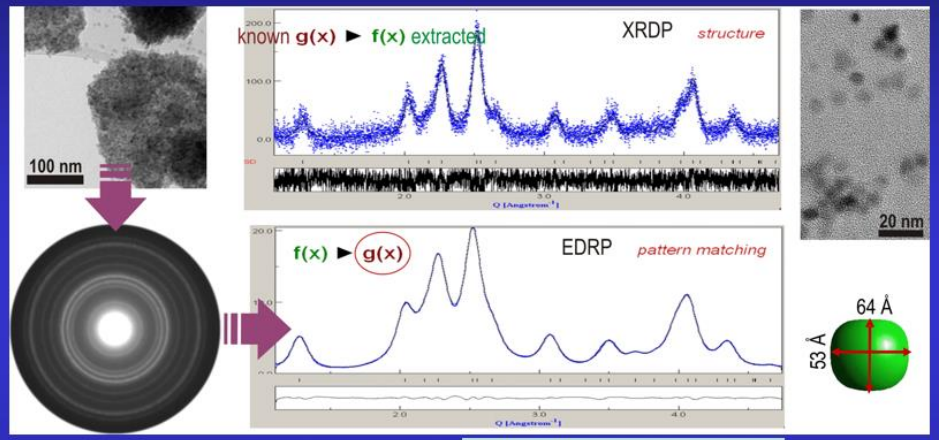
Crystallography Information File (CIF)

CIF dictionaries provide a formal taxonomy of crystallographic terms and ideas. Dictionary entries are constructed in a structured machine-readable manner that facilitates validation and structuring of data: http://www.iucr.org/resources/cif/dictionaries

Dictionaries: Core, Restraints, Powder, Modulated-Composite, Electron density, Twinning, Macromolecular, Images, Symmetry

Local dictionnaries: reflectivity (to come), MPOD (Properties), MAUD

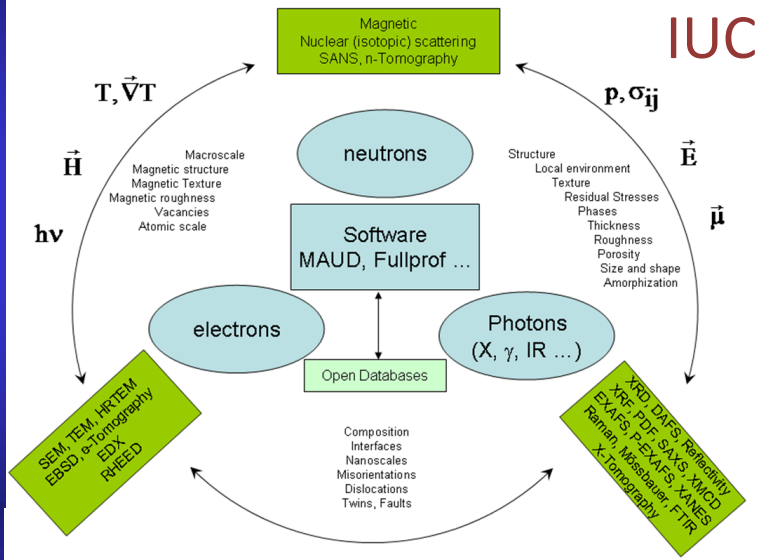
Reflection for 3h (100mg)



TEM in seconds (few µg)

$$\langle R_h \rangle = R_0 + R_1 P_2^0(x) + R_2 P_2^1(x)\cos\varphi + R_3 P_2^1(x)\sin\varphi + R_4 P_2^2(x)\cos 2\varphi + R_5 P_2^2(x)\sin 2\varphi + \ldots$$

IUCr input



Crystallographic Techniques

Diffraction (scattering + interferences): X, γ, n, e^-

Reflectivity (specular, off-specular): x-rays and neutrons

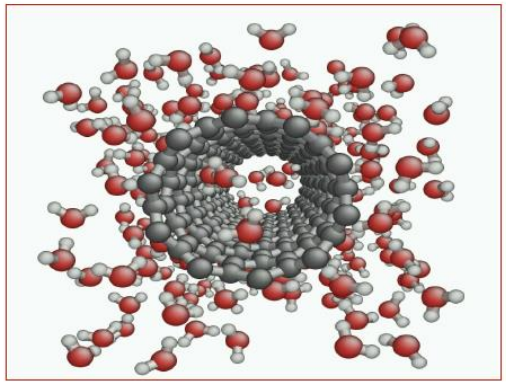
Small-Angles Scattering: x-rays (SAXS), neutrons (SANS)

Tomography (absorption or phase contrasts): x-rays, neutrons, electrons

Spectroscopy: X (XRF, XANES, EXAFS, DAFS)
e^- (EDS)
μ^+ (µSR)

CODATA/VAMAS Working Group on Description of Nanomaterials

Uniform Description System for Materials on the Nanoscale



Prepared by the CODATA-VAMAS Working Group
On the Description of Nanomaterials
www.codata.org/nanomaterials

Version 1.0
1 February 2015

Table 7. Descriptors for the crystal structure of a nano-object

Descriptors for the Crystal Structure of a Nano-Object	
Descriptor	**Definition**
Subcategory: Physical Structure Identification	
Physical structure name	Name of the physical structure within the nano-object that is being described by its crystal structure
Physical structure type	Structure type: layer, shell, surface, etc.
Physical structure location	Location of the physical structure within the nano-object
General nano-object type	Metal, polymer, etc.
Subcategory: Unit Cell Information	
Crystal system	The crystal system of the physical structure (one of seven)
Brevais lattice	The Brevais lattice of the physical structure (one of 14)
Space group	The space group
Miller indices	The appropriate Miller indices
Subcategory: Basic Unit Cell Parameters	
Cell length a	Cell length *a* appropriate for the crystal system value
Cell length b	Cell length *b* appropriate for the crystal system value
Cell length c	Cell length *c* appropriate for the crystal system value
Cell angle alpha	Cell angle *alpha* appropriate for the crystal system value
Cell angle beta	Cell angle *beta* appropriate for the crystal system value
Cell angle gamma	Cell angle *gamma* appropriate for the crystal system value
Cell volume	Measured or calculated cell volume
Cell measurement temperature	Temperature at which crystal structure data were measured

Other organizations with inter-disciplinary interest in scientific metadata

ICSTI – International Council for Scientific and Technical Information

http://www.icsti.org

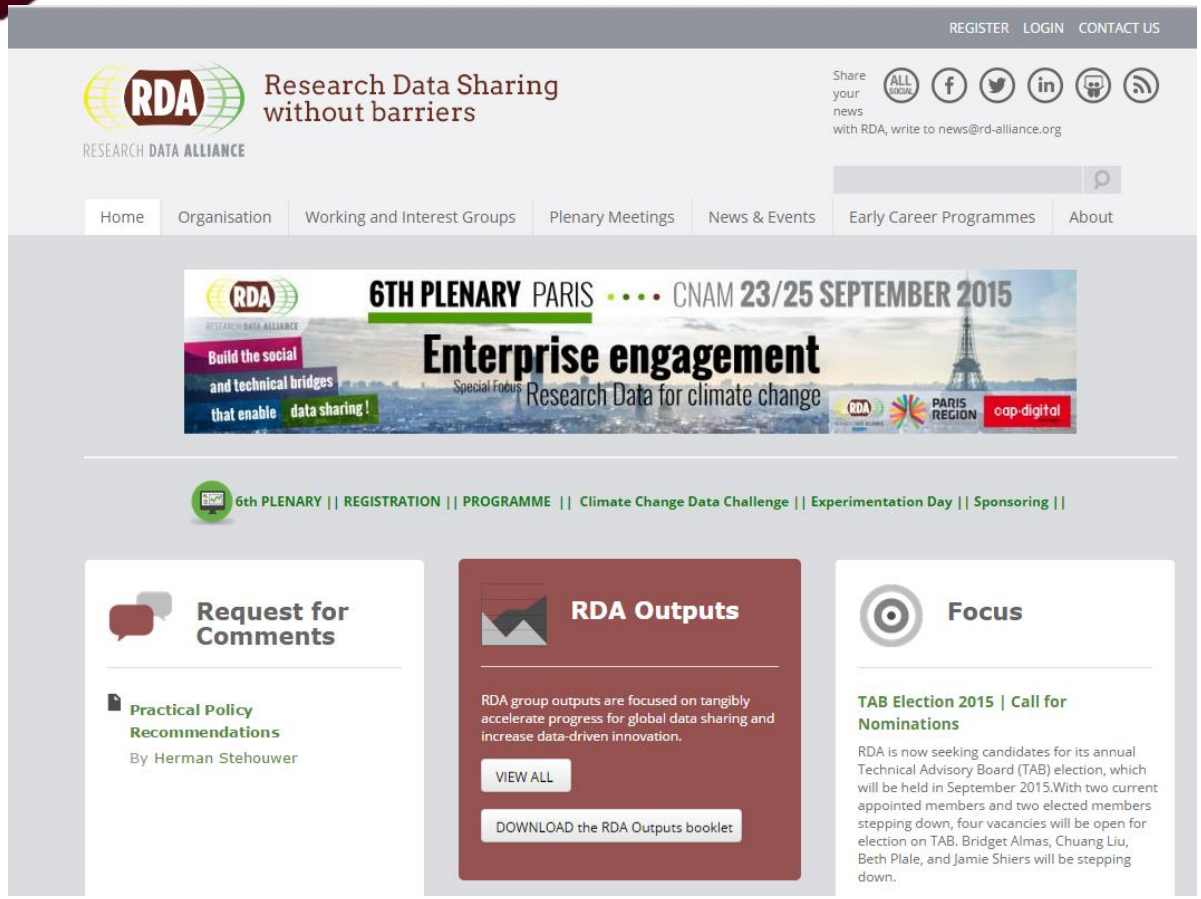
Other organizations with inter-disciplinary interest in scientific metadata

ICSTI – International Council for Scientific and Technical Information

- Affiliate of the International Council for Science (ICSU)
- Mission: facilitate cooperation among stakeholders engaged in the scientific communication process with the aim of improving the effectiveness of scientific research
- Task Groups: CODATA/ICSTI Task Group on Data Citation Standards and Practices
- Workshops/Conferences: Goportis Conference 2013 on Non-Textual Information – *Strategy and Innovation Beyond Text*; Winter Workshop 2012 – *Delivering Data in Science*; Winter Workshop 2010 – *Interactive Publications and the Record of Science*; Conference 2009 – *Managing Data for Science*
- ICSTI Insights: *The Living Publication*

Other organizations with inter-disciplinary interest in scientific metadata

RDA – Research Data Alliance



https://rd-alliance.org

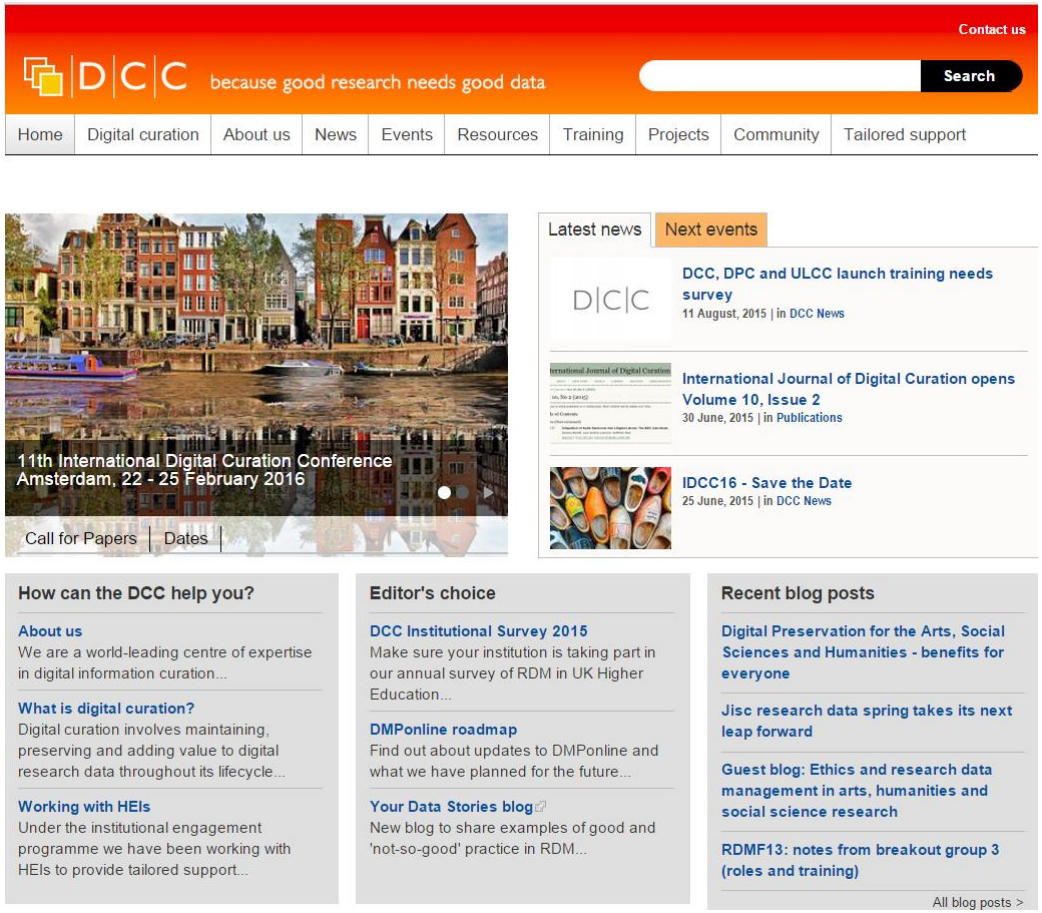
Other organizations with inter-disciplinary interest in scientific metadata

RDA – Research Data Alliance

- Launched 2013
- Core group: European Commission, US National Science Foundation, US National Institute of Standards and Technology, Australian Government Department of Innovation
- Other agencies, companies, associations and institutions
- Individual membership – 3000 from 102 countries
- Plenary Meetings
- Project-oriented methodology

Other organizations with inter-disciplinary interest in scientific metadata

DCC – Digital Curation Centre



http://www.dcc.ac.uk

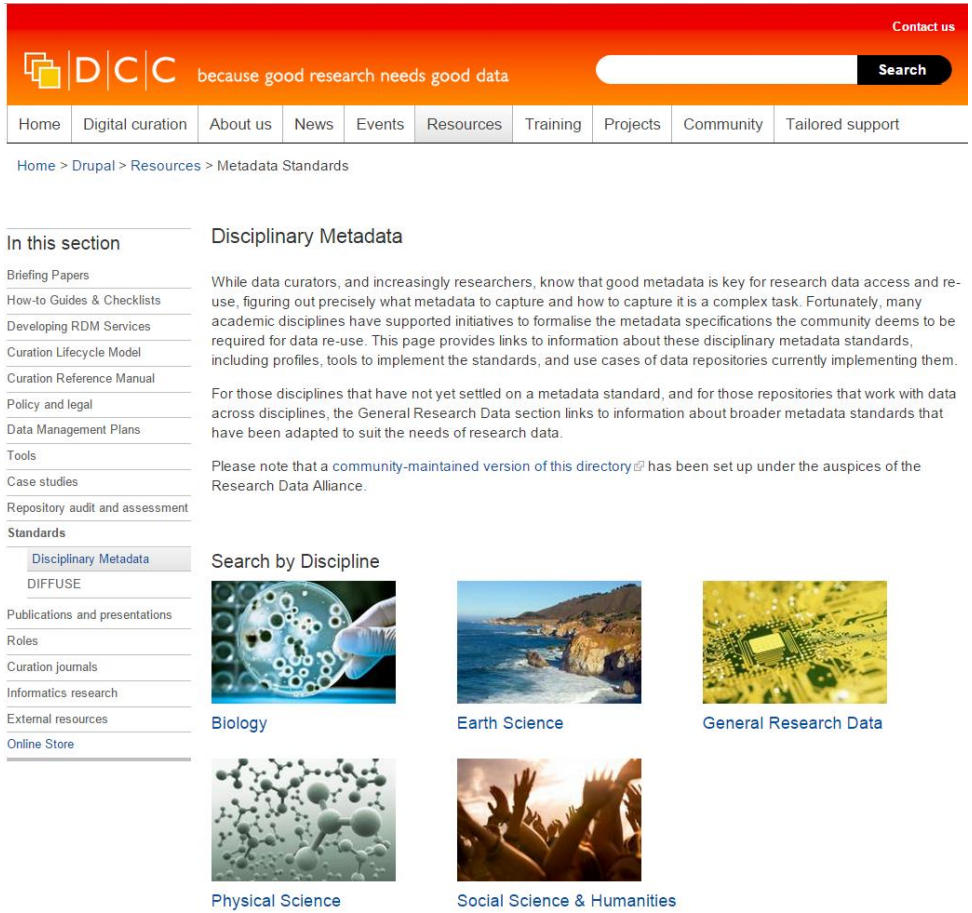
Other organizations with inter-disciplinary interest in scientific metadata

DCC – Digital Curation Centre

- Launched 1 March 2004, following successful response to JISC Circular 6/038 by consortium comprising Universities of Edinburgh and Glasgow (jointly hosting National e-Science Centre), UKOLN at University of Bath, and STFC, which managed Rutherford Appleton and Daresbury Laboratories
- National centre for solving challenges in digital curation that could not be tackled by any single institution or discipline
- Shift from development of curation tools to renewed focus on building capacity, capability and skills for data curation across the UK's higher education research community

Other organizations with inter-disciplinary interest in scientific metadata

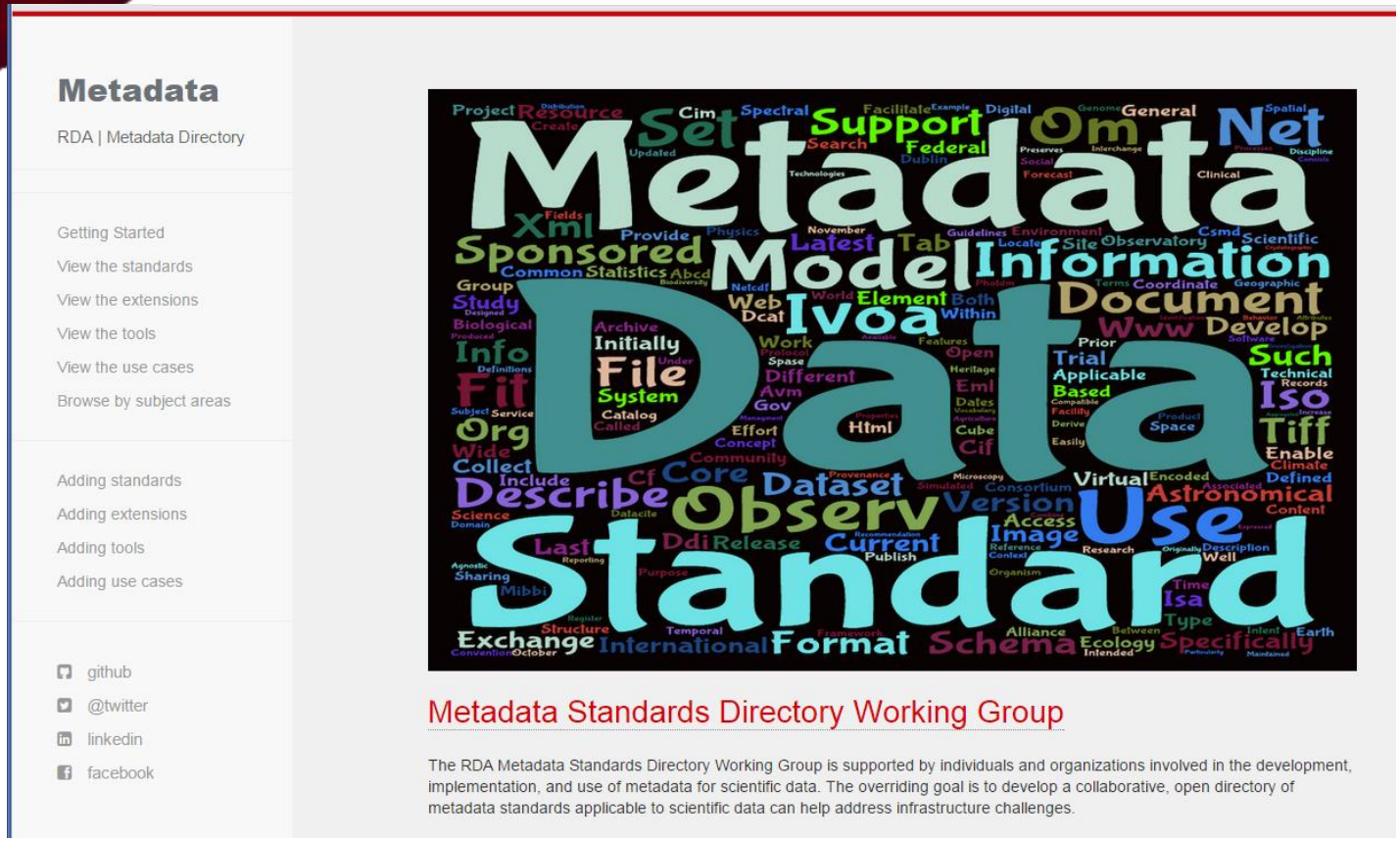
DCC – Disciplinary Metadata



http://www.dcc.ac.uk/resources/metadata-standards

Other organizations with inter-disciplinary interest in scientific metadata

RDA/DCC – Disciplinary Metadata



http://rd-alliance.github.io/metadata-directory/

Other organizations with inter-disciplinary interest in scientific metadata

DCC – Disciplinary Metadata Catalogue

Contact us

DCC because good research needs good data

Search

Home | Digital curation | About us | News | Events | Resources | Training | Projects | Community | Tailored support

Home > Drupal > Resources > Subject Areas > Physical Science

Physical Science

Geography Geoscience Space science Chemistry Crystallography Multi-disciplinary Biochemistry Geology Molecular biology Remote Sensing Astronomy Nuclear and Particle Physics Solar physics Bioinformatics Astrophysics Physics Meteorology Materials Science

Metadata Standards

AVM - Astronomy Visualization Metadata
A standard defining discovery metadata for fully rende

CIF - Crystallographic Information Framework
An extensible standard file format and set of protocols structured data.

CSMD-CCLRC Core Scientific Metadata Model
A study-data oriented model that captures high-level in they produce, primarily tailored for the physical sciences.

FITS - Flexible Image Transport System
Used by the astronomy community to originally describ standards to describe multi-dimensional data including scaling and distortions that may be present.

International Virtual Observatory Alliance Technical Spe
A set of specifications, including metadata standards, t archives into an international virtual observatory.

Observations and Measurements
This standard specifies an XML implementation for the (O&M) conceptual model, including a schema for Sam

PDBx/mmCIF – Protein Data Bank Exchange Dictionary Information Framework
PDBx/mmCIF is the standard archive format used by th metadata and data according to properties defined in t Macromolecular Crystallographic Information Framewo

Curation Reference Manual

Advice, in-depth information and criticism on current techniques and best practice.

CIF - Crystallographic Information Framework

An extensible standard file format and set of protocols for the exchange of crystallographic and related structured data.

CSMD-CCLRC Core Scientific Metadata Model

A study-data oriented model that captures high-level information about scientific studies and the data that they produce, primarily tailored for the physical sciences.

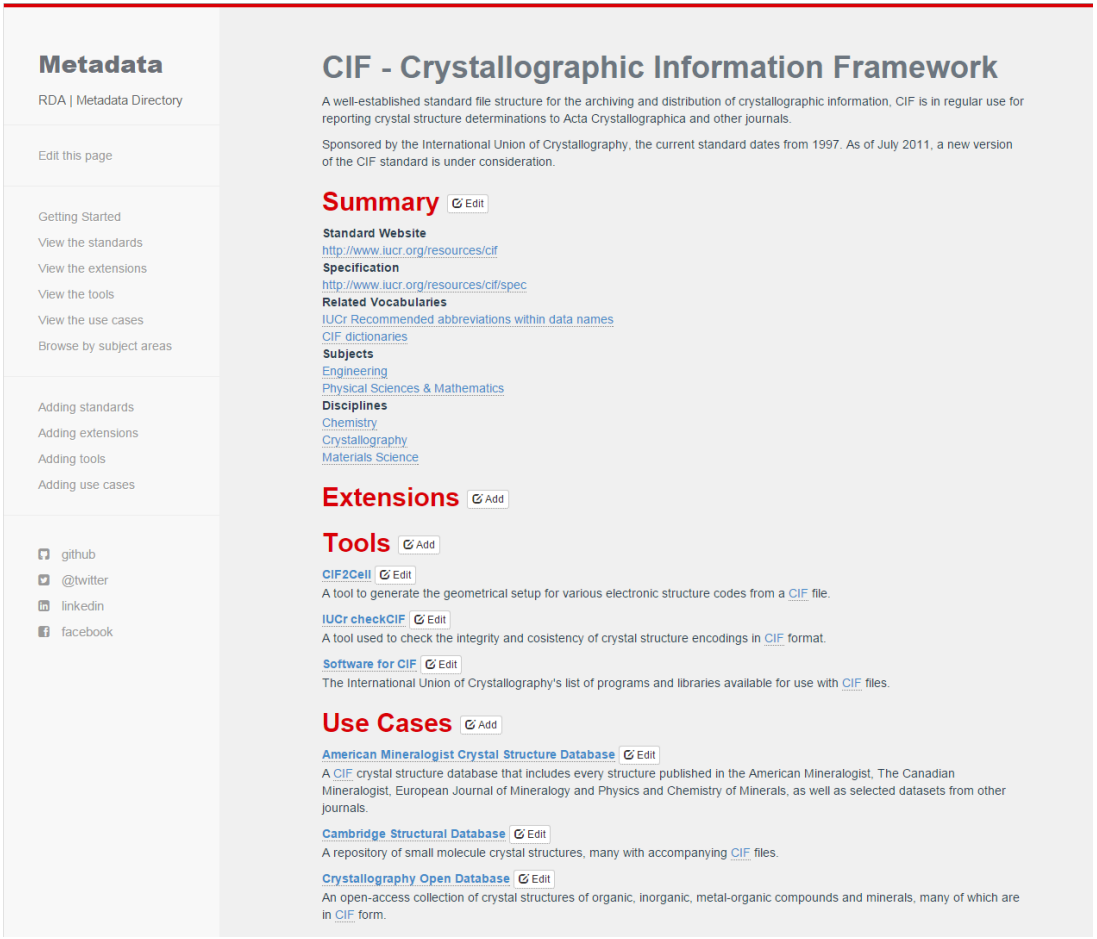
PDBx/mmCIF – Protein Data Bank Exchange Dictionary and the Macromolecular Crystallographic Information Framework

PDBx/mmCIF is the standard archive format used by the Protein Data Bank (PDB). It provides both metadata and data according to properties defined in the PDB Exchange Dictionary and the Macromolecular Crystallographic Information Framework (mmCIF).

http://www.dcc.ac.uk/resources/subject-areas/physical-science

Other organizations with inter-disciplinary interest in scientific metadata

RDA/DCC – Disciplinary Metadata Catalogue



Metadata

RDA | Metadata Directory

Edit this page

Getting Started
View the standards
View the extensions
View the tools
View the use cases
Browse by subject areas

Adding standards
Adding extensions
Adding tools
Adding use cases

github
@twitter
linkedin
facebook

CIF - Crystallographic Information Framework

A well-established standard file structure for the archiving and distribution of crystallographic information, CIF is in regular use for reporting crystal structure determinations to Acta Crystallographica and other journals.

Sponsored by the International Union of Crystallography, the current standard dates from 1997. As of July 2011, a new version of the CIF standard is under consideration.

Summary Edit

Standard Website
http://www.iucr.org/resources/cif
Specification
http://www.iucr.org/resources/cif/spec
Related Vocabularies
IUCr Recommended abbreviations within data names
CIF dictionaries
Subjects
Engineering
Physical Sciences & Mathematics
Disciplines
Chemistry
Crystallography
Materials Science

Extensions Add

Tools Add

CIF2Cell Edit
A tool to generate the geometrical setup for various electronic structure codes from a CIF file.

IUCr checkCIF Edit
A tool used to check the integrity and cosistency of crystal structure encodings in CIF format.

Software for CIF Edit
The International Union of Crystallography's list of programs and libraries available for use with CIF files.

Use Cases Add

American Mineralogist Crystal Structure Database Edit
A CIF crystal structure database that includes every structure published in the American Mineralogist, The Canadian Mineralogist, European Journal of Mineralogy and Physics and Chemistry of Minerals, as well as selected datasets from other journals.

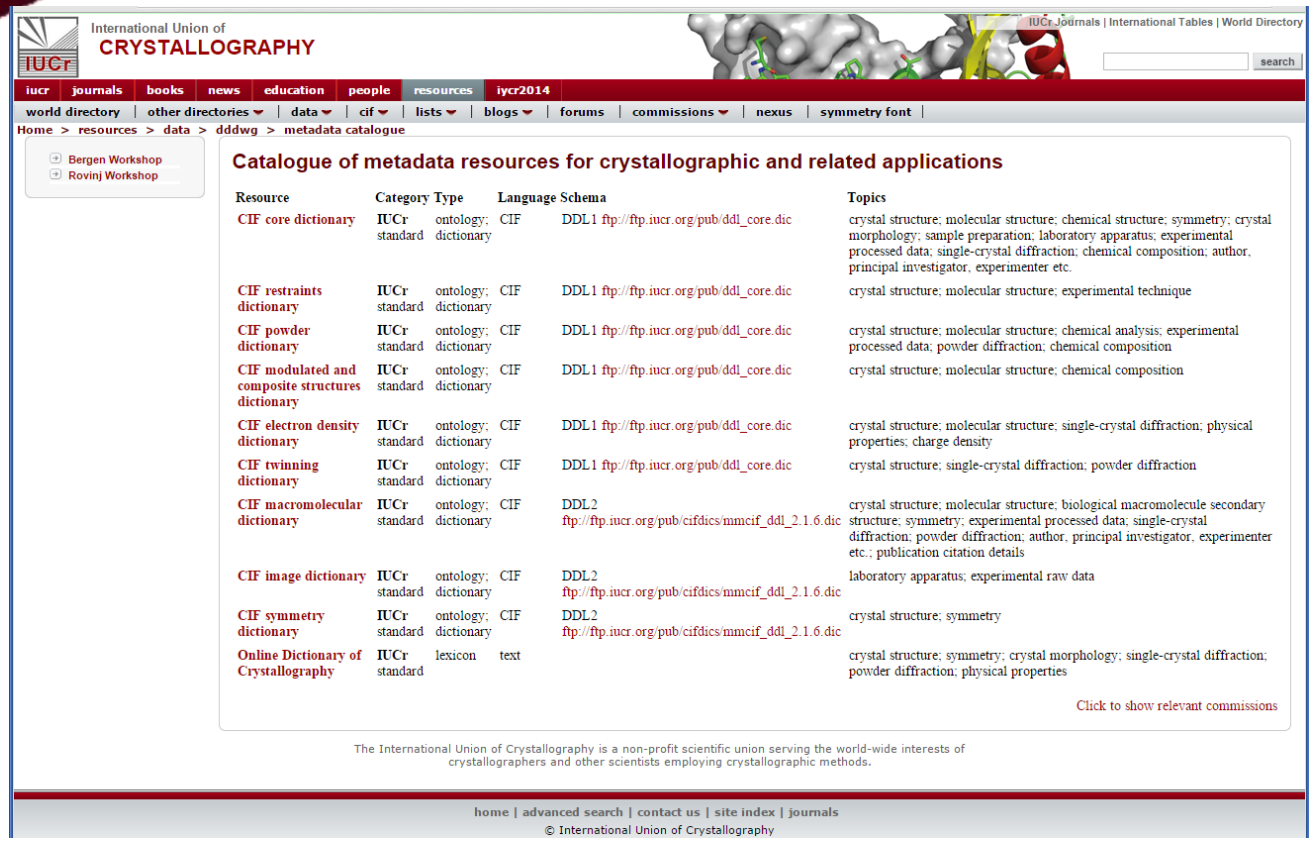
Cambridge Structural Database Edit
A repository of small molecule crystal structures, many with accompanying CIF files.

Crystallography Open Database Edit
An open-access collection of crystal structures of organic, inorganic, metal-organic compounds and minerals, many of which are in CIF form.

http://rd-alliance.github.io/metadata-directory/standards/cif-crystallographic-information-framework.html

IUCr/DDDWG initiative

Catalogue of metadata resources for crystallography



International Union of CRYSTALLOGRAPHY

IUCr Journals | International Tables | World Directory

search

iucr | journals | books | news | education | people | resources | iycr2014

world directory | other directories | data | cif | lists | blogs | forums | commissions | nexus | symmetry font

Home > resources > data > dddwg > metadata catalogue

Bergen Workshop
Rovinj Workshop

Catalogue of metadata resources for crystallographic and related applications

Resource	Category	Type	Language	Schema	Topics
CIF core dictionary	IUCr standard	ontology; dictionary	CIF	DDL1 ftp://ftp.iucr.org/pub/ddl_core.dic	crystal structure; molecular structure; chemical structure; symmetry; crystal morphology; sample preparation; laboratory apparatus; experimental processed data; single-crystal diffraction; chemical composition; author, principal investigator, experimenter etc.
CIF restraints dictionary	IUCr standard	ontology; dictionary	CIF	DDL1 ftp://ftp.iucr.org/pub/ddl_core.dic	crystal structure; molecular structure; experimental technique
CIF powder dictionary	IUCr standard	ontology; dictionary	CIF	DDL1 ftp://ftp.iucr.org/pub/ddl_core.dic	crystal structure; molecular structure; chemical analysis; experimental processed data; powder diffraction; chemical composition
CIF modulated and composite structures dictionary	IUCr standard	ontology; dictionary	CIF	DDL1 ftp://ftp.iucr.org/pub/ddl_core.dic	crystal structure; molecular structure; chemical composition
CIF electron density dictionary	IUCr standard	ontology; dictionary	CIF	DDL1 ftp://ftp.iucr.org/pub/ddl_core.dic	crystal structure; molecular structure; single-crystal diffraction; physical properties; charge density
CIF twinning dictionary	IUCr standard	ontology; dictionary	CIF	DDL1 ftp://ftp.iucr.org/pub/ddl_core.dic	crystal structure; single-crystal diffraction; powder diffraction
CIF macromolecular dictionary	IUCr standard	ontology; dictionary	CIF	DDL2 ftp://ftp.iucr.org/pub/cifdics/mmcif_ddl_2.1.6.dic	crystal structure; molecular structure; biological macromolecule secondary structure; symmetry; experimental processed data; single-crystal diffraction; powder diffraction; author, principal investigator, experimenter etc.; publication citation details
CIF image dictionary	IUCr standard	ontology; dictionary	CIF	DDL2 ftp://ftp.iucr.org/pub/cifdics/mmcif_ddl_2.1.6.dic	laboratory apparatus; experimental raw data
CIF symmetry dictionary	IUCr standard	ontology; dictionary	CIF	DDL2 ftp://ftp.iucr.org/pub/cifdics/mmcif_ddl_2.1.6.dic	crystal structure; symmetry
Online Dictionary of Crystallography	IUCr standard	lexicon	text		crystal structure; symmetry; crystal morphology; single-crystal diffraction; powder diffraction; physical properties

Click to show relevant commissions

The International Union of Crystallography is a non-profit scientific union serving the world-wide interests of crystallographers and other scientists employing crystallographic methods.

home | advanced search | contact us | site index | journals

http://www.iucr.org/resources/data/dddwg/metadata-catalogue

Other generic metadata systems

- DCMI – Dublin Core Metadata Initiative
- NISO – ISO 23081 Records Management Metada Standard
- NISO – Z39.87 Data Dictionary: Technical Metadata for Still Images
- PREMIS – Preservation metadata
- METS – Structural metadata regarding objects within a digital library
- TOTEM – Trustworthy Online Technical Environment Metadata Registry

Concordances

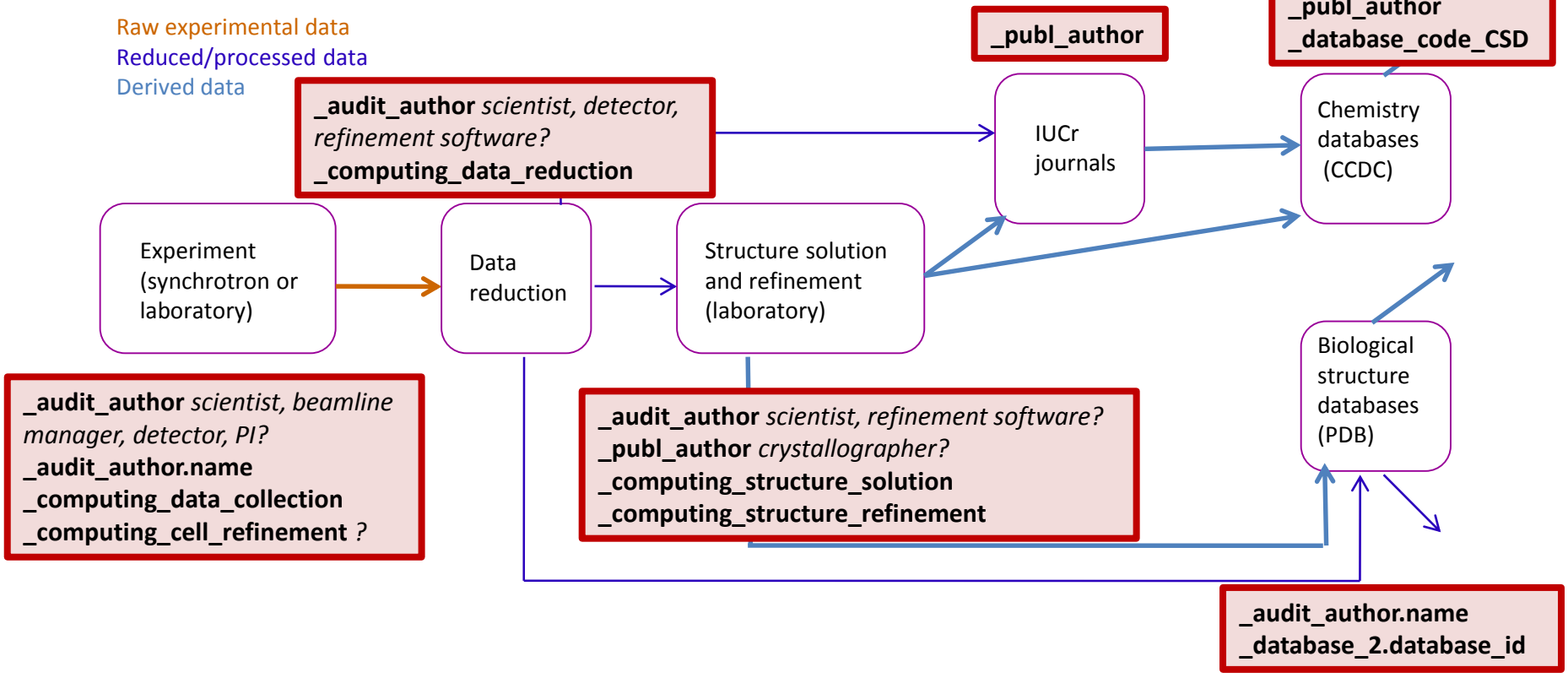
Relationship between generic and disciplinary metadata formulations

	Generic (DC)	Disciplinary (CIF)
Purpose	Cataloguing; classification; identification; discovery; location; provenance	Reproducibility; provenance; validation; discovery; retrieval
Users	Librarians; digital archivists; repository managers; funders	Scientists; software developers; LIMS developers; data managers; analysts
Granularity	Macro	Micro
Scale	Relatively small (DC: 15 terms)	Relatively large (coreCIF: ~720 terms; pdbx/mmCIF: ~4700 terms)
Stability	Static	Extensible (mmCIF had 4100 terms in 2006)
Semantics	Fluid	Precise

Concordances

Contextual significance of a high-level metadata descriptor

DC: Creator
An entity primarily responsible for making the resource



Middleware

Is there a need for an intermediate 'glue' layer of metadata?

	Generic (DC)	*Metametadata*	Disciplinary (CIF)
Purpose	Cataloguing; classification; identification; discovery; location; provenance	Provenance; classification; purpose; schema; discovery	Reproducibility; provenance; validation; discovery; retrieval
Users	Librarians; digital archivists; repository managers; funders	Repository managers; funders; knowledge base builders; end users	Scientists; software developers; LIMS developers; data managers; analysts
Granularity	Macro	Intermediate	Micro
Scale	Relatively small	As required	Relatively large
Stability	Static	Relatively stable, but extensible	Extensible
Semantics	Fluid	Specific	Precise

Metadata for raw data from X-ray diffraction and other structural techniques

A Satellite Workshop to the 29th European Crystallographic Meeting

Sponsors

DECTRIS®

IUCr Journals
CRYSTALLOGRAPHY JOURNALS ONLINE

THE CAMBRIDGE CRYSTALLOGRAPHIC DATA CENTRE

ICODATA ISU

FIZ Karlsruhe
ICSD

WILEY

OxfordCryosystems

BRUKER