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

A Satellite Workshop to the 29th European Crystallographic Meeting

Practical session: 1. Catalogue of metadata standards

DDDWG T

DDDWG Work Plan

Metadata

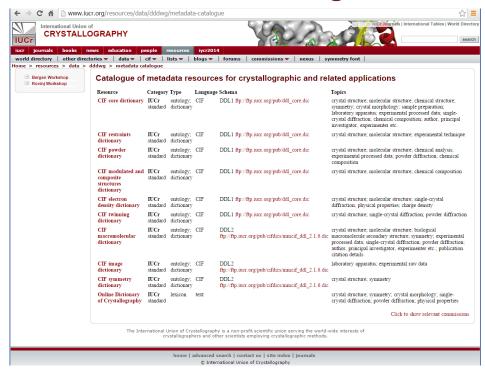
DDDWG Triennial Report 2011-2014

- In the Madrid Congress the Commissions were charged at the DDD inaugural meeting to define the metadata that should accompany their raw data ...
- While each IUCr Commission needs to specify 'technical' metadata *i.e.* those specific to their experimental raw data there is also a need to review 'generic' metadata *e.g.* who 'owns' a data set, details of research grants, embargo periods etc. A higher-level classification of the domain of study may be needed. *E.g.* a synchrotron facility might need to define different data storage policies for, say, X-ray diffraction images *versus* X-ray tomography images. Such policies could be automatically implemented if data sets had characteristics identifying what sort of scientific study they represent.

DDDWG Work Plan

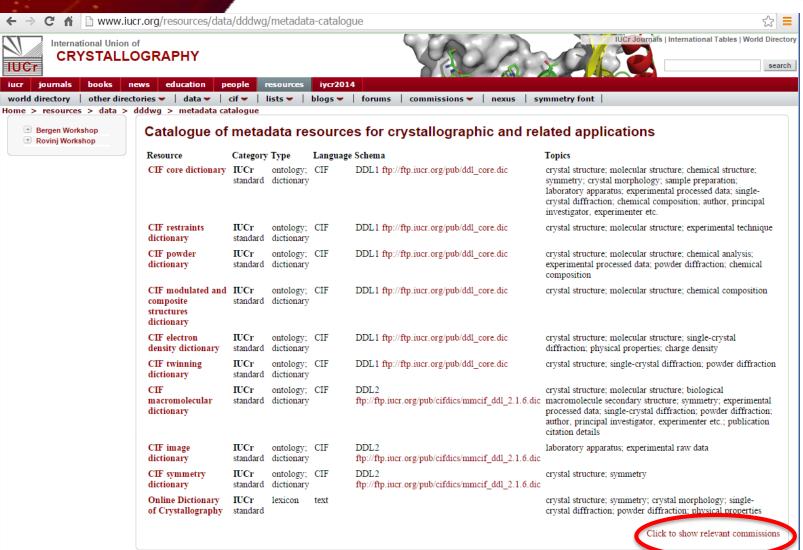
Commissions were charged at the DDD inaugural meeting to define the metadata

Metadata Catalogue



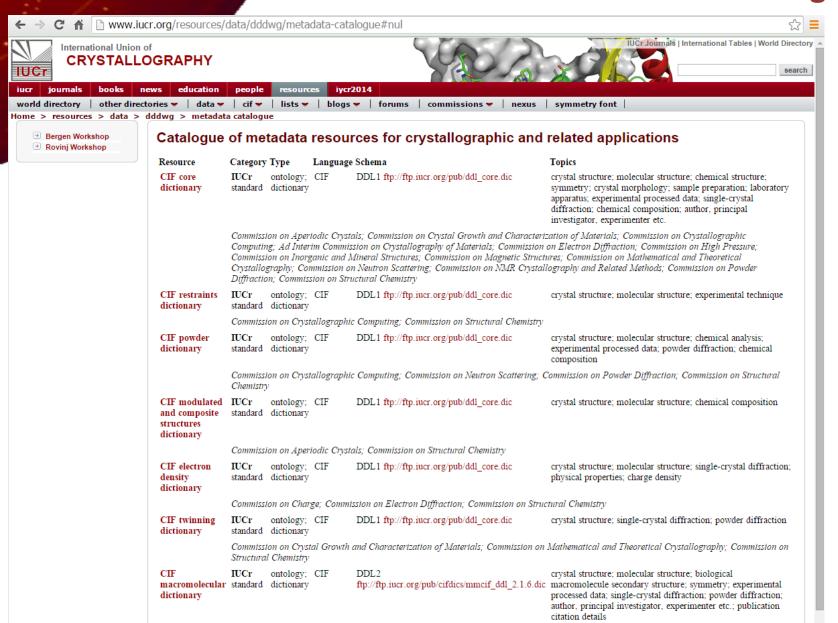
http://www.iucr.org/resources/data/dddwg/metadata-catalogue

Metadata catalogue



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

Metadata catalogue





Catalogue of metadata resources for crystallographic and related applications

Resource	Category	Туре	Language	Schema	Topics
CIF core dictionary	IUCr standard	ontology; dictionary		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.
	Ad Interim Con Inorganic and .	nmission on Mineral Str	Crystallograph uctures; Commi	ny of Materials; Commission on Electron D ssion on Magnetic Structures; Commission	ntion of Materials; Commission on Crystallographic Computing; iffraction; Commission on High Pressure; Commission on on Mathematical and Theoretical Crystallography; Commission Commission on Powder Diffraction; Commission on Structural

Metadata catalogue – more detail

Link to resource itself?

Controlled vocabulary

Controlled vocabulary

Controlled vocabulary

Controlled vocabulary

Controlled vocabulary

Controlled vocabulary for IUCr

Commissions. Other stakeholders?

Documentation?

Name of property	Example	Comments
Resource name	CIF core dictionary	

ftp://ftp.iucr.org/pub/cif core.dic

ftp://ftp.iucr.org/pub/ddl_core.dic

Commission on Aperiodic Crystals;

crystal structure; molecular

structure; chemical structure;

standard

ontology; dictionary

IUCr

CIF

DDL1

symmetry; ...

Commission on ...

Resource URI

Category

Authority

Language

Schema URI

Commissions

Schema

Topics

Type

Metadata catalogue - detail

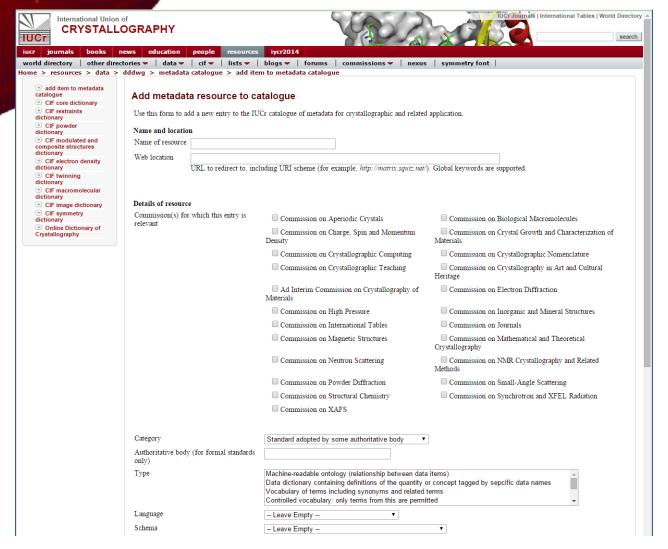
Catalogue of metadata resources for crystallographic and related applications

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.
	Ad Interim Con Inorganic and	nmission on Mineral Stri	Crystallography uctures; Commiss	of Materials; Commission on Electron Diffraction sion on Magnetic Structures; Commission on Math	Materials; Commission on Crystallographic Computing; n; Commission on High Pressure; Commission on mematical and Theoretical Crystallography; Commission mion on Powder Diffraction; Commission on Structural
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
	Commission or Structural Che		Macromolecules	; Commission on Crystallographic Computing; Co	mmission on Powder Diffraction; Commission on
Online Dictionary of Crystallography	IUCr standard	lexicon	text		crystal structure; symmetry; crystal morphology; single- crystal diffraction; powder diffraction; physical properties
	Crystal Growth	h and Chara	cterization of Ma	nterials; Commission on Crystallographic Computi	on Charge, Spin and Momentum Density; Commission on ing; Commission on Crystallographic Nomenclature; ral Heritage; Ad Interim Commission on Crystallography

of Materials; Commission on Electron Diffraction; Commission on High Pressure; Commission on Inorganic and Mineral Structures; Commission on International Tables; Commission on Journals; Commission on Magnetic Structures; Commission on Mathematical and Theoretical Crystallography; Commission on Neutron Scattering; Commission on NMR Crystallography and Related Methods; Commission on Powder Diffraction; Commission on

Small-Angle Scattering; Commission on Structural Chemistry; Commission on Synchrotron and XFEL Radiation; Commission on XAFS

Metadata catalogue - populating





Metadata catalogue - identification

- Should the web location refer to the machine-readable resource itself, or to a documentation / home page?
- Should there be separate entries for separate versions?

Add metadata resource to catalogue

Use this form to add a new entry to the IUCr catalogue of metadata for crystallographic and related application.

Name and location

Name of resource

CIF core dictionary

Web location

ftp://ftp.iucr.org/pub/cif_core.dic

URL to redirect to, including URI scheme (for example, http://matrix.squiz.net/).



Metadata catalogue - stakeholders

- Very IUCr specific
- Remove from public form?
- Add other categories of stakeholder?

Commission on Aperiodic Crystals

Commission on Charge, Spin and Momentum Density

☐ Commission on Crystallographic Computing

☐ Commission on Crystallographic Teaching

■ Ad Interim Commission on Crystallography of Materials

☐ Commission on High Pressure

■ Commission on International Tables

■ Commission on Magnetic Structures

■ Commission on Neutron Scattering

■ Commission on Powder Diffraction

☐ Commission on Structural Chemistry

■ Commission on XAFS

☐ Commission on Biological Macromolecules

☐ Commission on Crystal Growth and Characterization of Materials

Commission on Crystallographic Nomenclature

☐ Commission on Crystallography in Art and Cultural Heritage

☐ Commission on Electron Diffraction

☐ Commission on Inorganic and Mineral Structures

□ Commission on Journals

Commission on Mathematical and

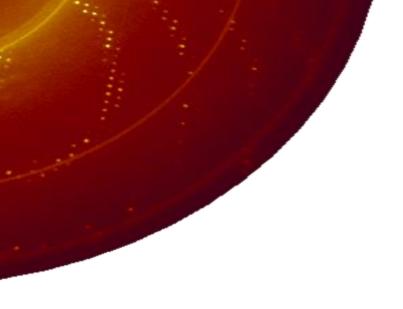
Theoretical Crystallography

Commission on NMR Crystallography and Related Methods

terated intellious

■ Commission on Small-Angle Scattering

☐ Commission on Synchrotron and XFEL Radiation



Metadata catalogue - category

- Only one allowed
- Additional categories?
- "Authority" can be relatively informal
- Should we require documentation for the authority?

Category

Standard adopted by some authoritative body
De facto standard arising from widespread use
Draft or proposed standard
Working system in local use (e.g. within a facility)
Other

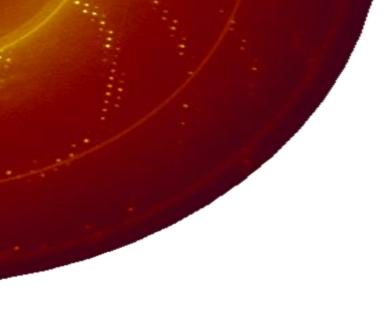
Authoritative body (for formal standards only)



- Multiple "type" allowed
- Only one language / schema

Type	Machine-readable ontology (relationship between data items)					
-JF-	Data dictionary containing definitions of the quantity or concept tagged by sepcific data names Vocabulary of terms including synonyms and related terms Controlled vocabulary: only terms from this are permitted List of keywords (differs from controlled vocabulary in that additional terms may be inserted by user) Other					
	Lexicographic dictionary of terminology					
Language						
	Crystallographic Information Framework					
	W3C Web Ontology Language					
	XML schema					
	Unstructured text					
Schema	▼					
	Distinger, Definition Language version 1					
	Dictionary Definition Language version 1					
	Dictionary Definition Language version 2					
	Dublin Core					
	METS (Metadata Encoding and Transmission Standard)					

URI of schema



Metadata catalogue – topics

- Multiple choice
- Community critique / suggestions especially welcomed

Topics covered by resource molecular biological macromolecule chemical structure crystal structure secondary structure structure symmetry crystal sample preparation chemical analysis morphology chemical reaction laboratory safety hazard, toxicity etc. experimental raw data apparatus experimental single-crystal powder diffraction physical properties processed data diffraction chemical charge density experimental technique author, principal composition investigator, experimenter etc. funding body, publication grant number etc. citation details