

# Demonstration of the production and use of IUPAC FAIRSpec Data Collections and IUPAC FAIRSpec Finding Aids from "FAIRSpec-ready" data collections

ACS National Meeting, Aug. 20, 2025

**Robert M. Hanson**, Henry S. Rzepa, Stefan Kuhn,  
John R. Cort, Swagat Malla

**[IUPAC Project 2019-031-1-024](#)**

# Our mandate

From the *NSF Guide to Proposal Preparation*:

*Proposals must include a document of no more than two pages uploaded under “Data Management Plan” .... This supplementary document should **describe how the proposal will conform to NSF policy on the dissemination and sharing of research results** ... and may include ... the standards to be used for data and metadata format and content.*

*Investigators are expected to share with other researchers ... the **primary data**, samples, physical collections, **and other supporting materials** created or gathered in the course of work under NSF awards.*

# Our mandate

From the *CHE Data Management and Sharing Plan Guidance*:

**Access to data and data sharing practices and policies.** ...Describe your plans, if any, for providing such general access to data, including **websites maintained by your research group**, and **direct contributions to public databases or software repositories** (e.g., **NMRShiftDB**,.... Consider using **digital object identifiers (DOI)** ... for **suitably-archived, publishable data sets**.

**Archiving of Data.** ...Will there be an **easily accessible index** that documents where all archived data are stored and how they can be accessed?

# Unaddressed issues: findability and reusability

The key missing term here is **collection**




# Unaddressed issues: findability and reusability

How do we organize it? How do we make it available?



# Our approach:

 INTERNATIONAL UNION OF  
PURE AND APPLIED CHEMISTRY



Contact | Login | Join | IYPT 2019  Shop

WHO WE AREWHAT WE DOEVENTSPROJECTSNEWS

 **PROJECTS**

**WHAT IS AN IUPAC PROJECT**

**FAQS ON THE PROJECT SUBMISSION AND APPROVAL PROCESS**

**PROJECT SUBMISSION FORM AND GUIDELINES**

**ADVICE FOR PROJECT REVIEWERS**

**PROJECT REVIEW PROCEDURE**

## Development of a Standard for FAIR Data Management of Spectroscopic Data


<b>Project No.:</b>	2019-031-1-024
<b>Start Date:</b>	18 Mar 2020
<b>End Date:</b>	
<b>Cite:</b>	<a href="https://iupac.org/project/2019-031-1-024">https://iupac.org/project/2019-031-1-024</a>
<b>Division:</b>	Committee on Publications and Cheminformatics Data Standards

**Chair**  
Robert M. Hanson

**Members**  
Mark Archibald  
Ian Bruno  
Stuart Chalk  
Antony Davies  
Damien Jeannerat  
Robert J. Lancashire  
Jeffrey Lang  
Henry Rzepa


# Our approach:

 INTERNATIONAL UNION OF  
PURE AND APPLIED CHEMISTRY



Contact | Login | Join | IYPT 2019  Shop

WHO WE AREWHAT WE DOEVENTSPROJECTSNEWS

 **PROJECTS**

**WHAT IS AN IUPAC PROJECT**

**FAQS ON THE PROJECT SUBMISSION AND APPROVAL PROCESS**

**PROJECT SUBMISSION FORM AND GUIDELINES**

**ADVICE FOR PROJECT REVIEWERS**

**PROJECT REVIEW PROCEDURE**

## Development of a Standard for FAIR Data Management of Spectroscopic Data

<b>Project No.:</b>	2019-031-1-024
<b>Start Date:</b>	18 Mar 2020
<b>End Date:</b>	
<b>Cite:</b>	<a href="https://iupac.org/project/2019-031-1-024">https://iupac.org/project/2019-031-1-024</a>
<b>Division:</b>	Committee on Publications and Cheminformatics Data Standards

**Chair**  
Robert M. Hanson

**Members**  
Mark Archibald  
Ian Bruno  
Stuart Chalk  
Antony Davies  
Damien Jeannerat  
Robert J. Lancashire  
Jeffrey Lang  
Henry Rzepa

## Guiding Principles for the FAIR Management of Spectroscopic Data

### IUPAC Specification for the FAIR Management of Spectroscopic Data in Chemistry (IUPAC FAIRSpec) - Guiding Principles

*Robert M. Hanson, Damien Jeannerat, Mark Archibald, Ian Bruno, Stuart J. Chalk, Antony N. Davies, Robert J. Lancashire, Jeffrey Lang and Henry S. Rzepa*

Pure and Applied Chemistry, 2022

<https://doi.org/10.1515/pac-2021-2009>

#### 1. FAIR Management of data should be an ongoing concern.

- A. FAIR management of data must be an explicit part of research culture.
- B. FAIR management of data should be of intrinsic value.
- C. Good data management requires distributed curation.
- D. Experimental work is by nature iterative.

#### 2. Context is important.

- A. Digital objects are generally part of a collection.
- B. Chemical properties are related to chemical structure.
- C. Data relationships are diverse and develop over time.
- D. FAIR management of data should allow for validation.

#### 3. FAIR management of data requires curation

- A. Data reuse relies upon practical findability.
- B. Data has to be organized to be accessible.
- C. Data interoperability requires well-designed metadata.
- D. Value is in the eye of the reuser.

#### 4. Metadata must be standardized and registered.

- A. Register key metadata.
- B. Assign a variety of persistent identifiers.
- C. Enable metadata crosswalks.
- D. Allow for value-added benefits.

#### 5. FAIR data management standards should be *modular, extensible, and flexible*

- A. Modularity allows specialization.
- B. Allow for future needs.
- C. Respect format and implementation diversity.
- D. All data formats should be valued.



just out...

## Organic Chemistry

# FAIRSpec-Ready Spectroscopic Data Collections – Advice for Researchers, Authors, and Data Managers (IUPAC Technical Report)

19 August 2025, Version 2

### Version History

**Aug 19, 2025 Version 2**

[Jan 02, 2025 Version 1](#)

### Version Notes

Revisions requested by referees.

<https://doi.org/10.26434/chemrxiv-2025-vqhnd-v2>

# A must read!

The key is to provide well-crafted structure representations.

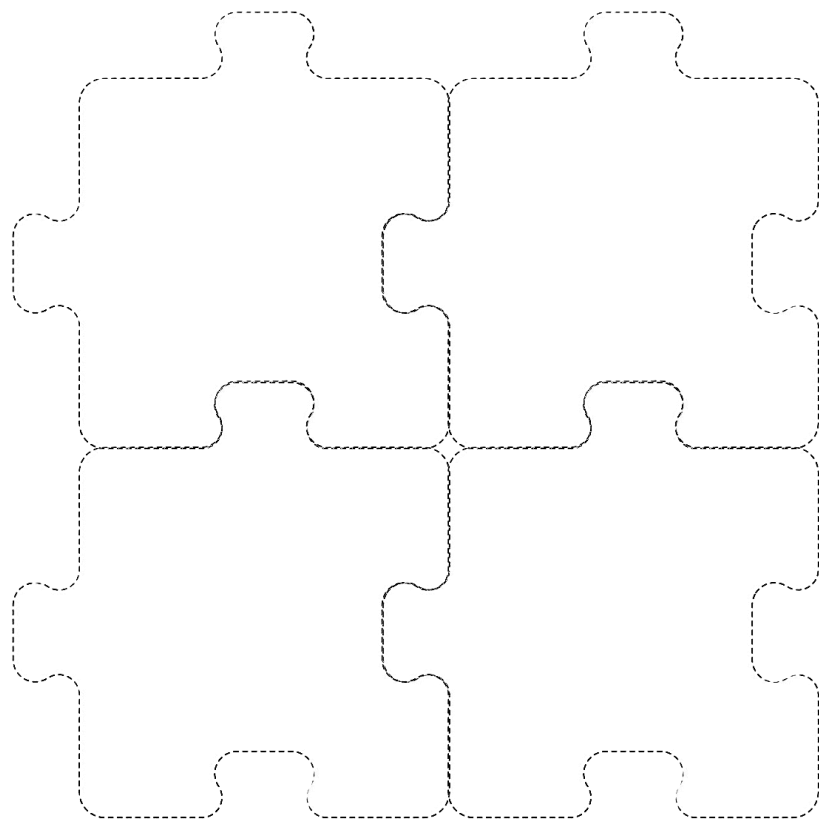
Table 1: Common digital chemical structure representations — pros and cons	
Representation type	Considerations
MDL-MOL Version 2000	Benefits: high interoperability, can contain 2D or 3D coordinates  Limitations: limited ability to convey bonding
MDL-MOL Version 3000	Benefits: all the features of Version 2000, with expanded capabilities to describe special bonding types and multi-atom connection bonding (as in ferrocene)  Limitations: less widely implemented to date in toolkits and informatics platforms.
CDXML	Benefits: well-specified structural format generated and read by popular drawing programs, allows for the expression of "nicknames" such as Ph and TBS, highly versatile in expressing nuances of bonding, may provide warnings of inappropriate bonding or charge states  Limitations: generally, less interoperable than MOL (at least at the time of this writing); ambiguity can arise from using bonding and atom elements for nonmolecular depictions, such as titles, labels, and drawn lines; specification is no longer formally maintained, but last published working specification version is available (ref. 12)
CDX	Benefits: well-specified binary equivalent of CDXML, easily converted to CDXML with open-source tools  Limitations: binary format lacks the human readability aspects of MOL and CDXML; see note for CDXML in relation to specification

# A must read!

...and appropriate dataset representations.

Table 2: Digital dataset representation examples	
Representation type	Considerations
Original instrument data (for example, the NMR FID and associated parameter files)	<p>Benefits: highest integrity with no information loss; high potential for reuse; allows for alternative and/or automated (re)analysis; potential for generation of all other representations; allows possibility of fraud detection; allows the most reliable automated metadata extraction.</p> <p>Limitations: vendor-specific format may require licensing access to a vendor-specific reader; may not express important aspects of the data processing used in the analysis; must contain enough of the key parameters for further processing; may no longer be documented; shortest expected lifetime.</p>
Original data exported to an alternative standardized format, such as JCAMP-DX(FID) <sup>38</sup> NMR-Star <sup>39</sup> or nmrML <sup>40</sup>	<p>Benefits: highest possibility of reuse and interoperability; vendor-agnostic open format; allows for automated production of additional representations; recognized by regulators as potentially the longest expected lifetime.</p> <p>Limitations: Depending on implementations potential for loss of data resolution or precision; potential for loss of some metadata fields.</p>
Instrument-processed data such as transformed NMR data in the form of a spectrum or spectra	<p>Benefits: most generally and immediately informative to the practicing scientist or educator; can be examined in detail and repurposed.</p> <p>Limitations: does not allow for early-processing adjustments, such as in NMR spectroscopy phasing, line broadening or use of non Fourier-transform methods; may require proprietary software to read; information loss compared to original data; less scope for fraud detection.</p>
Third party-processed data	<p>Benefits: concise; convenient if this is the standard process in each laboratory; may include meaningful annotation added by the originator such as integration or peak identification"; May be the only method of getting FAIR metadata to be associated with the data.</p> <p>Limitations: format may require proprietary software to read; possibly limited ability to extract key metadata from proprietary formats; may disallow alternative analysis; may be more subject to fraud or other sorts of spectral editing (removal of solvent peaks, for example)</p>

# The data management puzzle

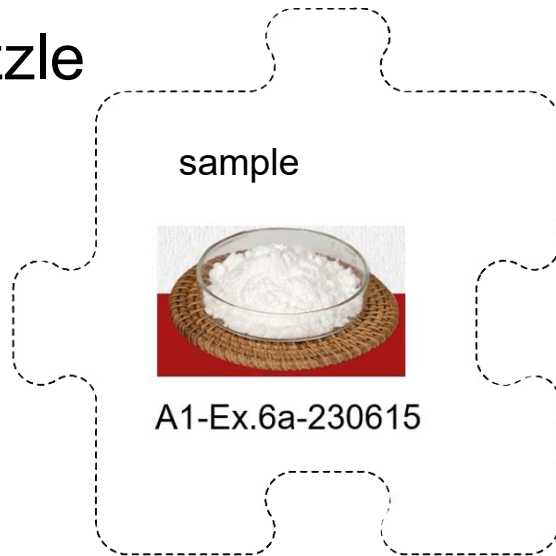


<b>7</b>	1 5 8 9	5 6 8	1 2 6 9	1 3 5 6 9	2 3 5 6 9	<b>4</b>	1 2 5	1 2 3
1 4 5 6 9	<b>2</b>	4 5 6	1 4 6 9	<b>7</b>	3 4 5 6 9	1 5 6 8	1 3	
1 4 5 6	1 4 5	<b>3</b>	1 2 4 6	1 5 6	<b>8</b>	1 2 5 6 7	1 2 5 7	<b>9</b>
2 4 8 9	4 7 8 9	2 4 7 8	<b>5</b>	1 6 8 9	4 6 9	<b>3</b>	1 2 4 7	1 2 4 7 8
3 4 5 8	<b>6</b>	4 5 7 8	1 4 8	<b>2</b>	3 4	1 5 7 8	<b>9</b>	1 4 7 8
2 3 4 5 8 9	4 5 8 9	<b>1</b>	4 8 9	3 8 9	<b>7</b>	2 5 8	2 4 5	<b>6</b>
1 2 5 6 8	1 5 7 8	2 5 6 7 8	<b>3</b>	5 6 8	2 5 6	<b>9</b>	1 2 4 7	1 2 4 7 8
1 2 5 8	<b>3</b>	2 5 7 8	2 7 8 9	<b>4</b>	2 5 9	1 2 7 8	<b>6</b>	1 2 7 8
2 4 6 8	4 7 8	<b>9</b>	2 6 7 8	6 8	<b>1</b>	2 7 8	<b>3</b>	<b>5</b>

# The data management puzzle

We start with a sample.

It has various characteristics, most importantly an ***identifier***



# The data management puzzle

We run tests on it, creating instrument datasets and reports.

Now we have a ***data collection***.

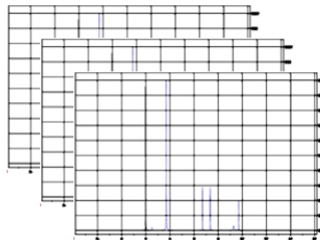
The sample identifier is the *metadata* that establishes the connection between the sample and its spectra.

sample



A1-Ex.6a-230615

spectra



# The data management puzzle

We think on this.

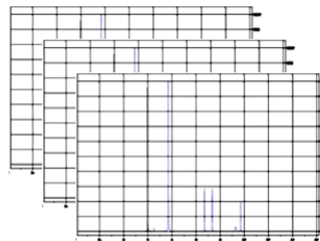
(Or pass it to AI??)

sample



A1-Ex.6a-230615

spectra



analysis



# The data management puzzle

Our goal is to make a connection between sample and structure.

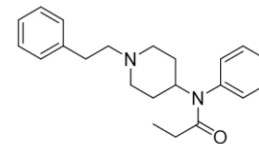
Note how indirect this really is!

sample

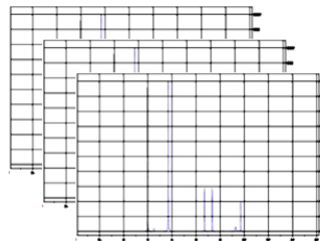


A1-Ex.6a-230615

structure



spectra



analysis

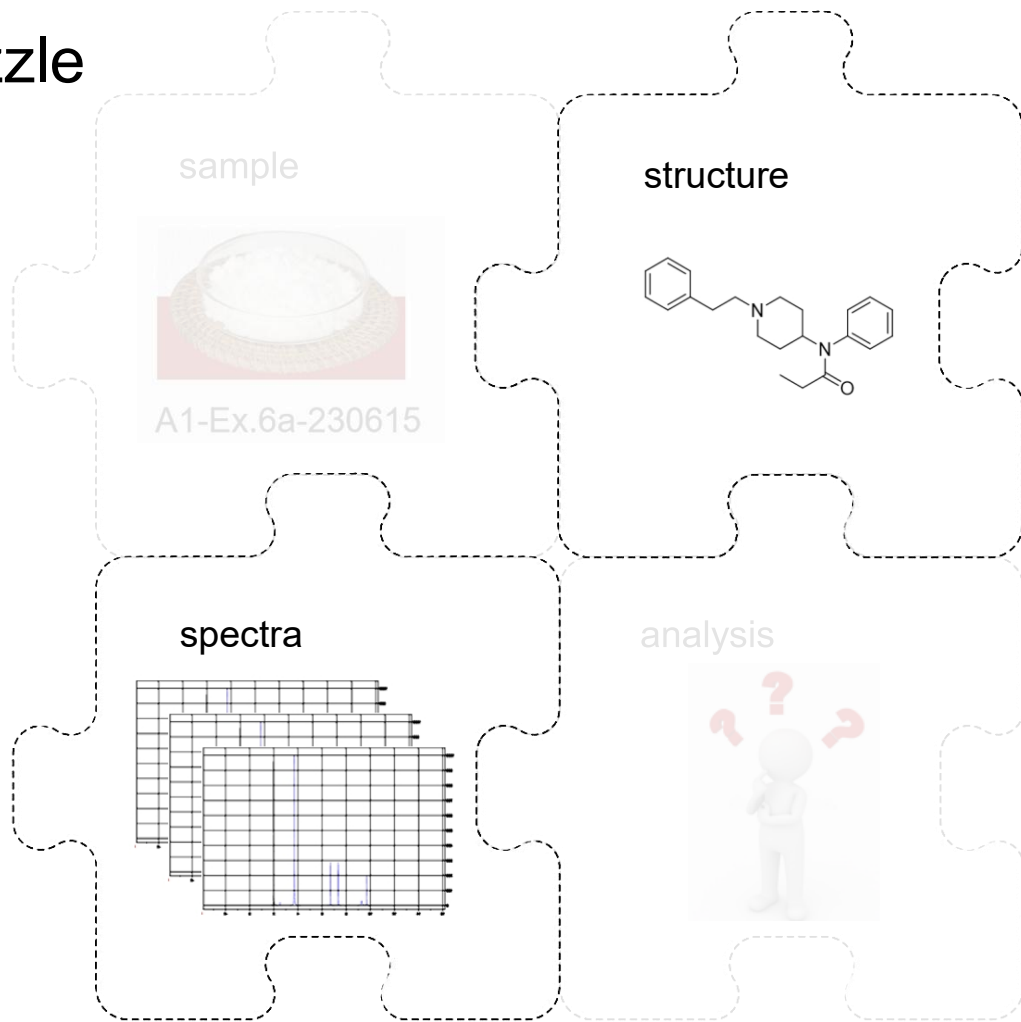




# The data management puzzle

But, in the end, we often just deliver two of the four pieces – the structure and the spectra.

We call this a ***compound association***



# The data management puzzle

The goal of IUPAC Project 2019-031-1-024 is to develop standards for the maintenance and presentation of these relationships.

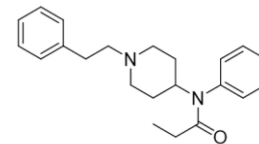
The overall context is much larger than just “SI for a paper”.

sample

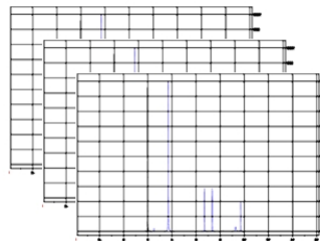


A1-Ex.6a-230615

structure



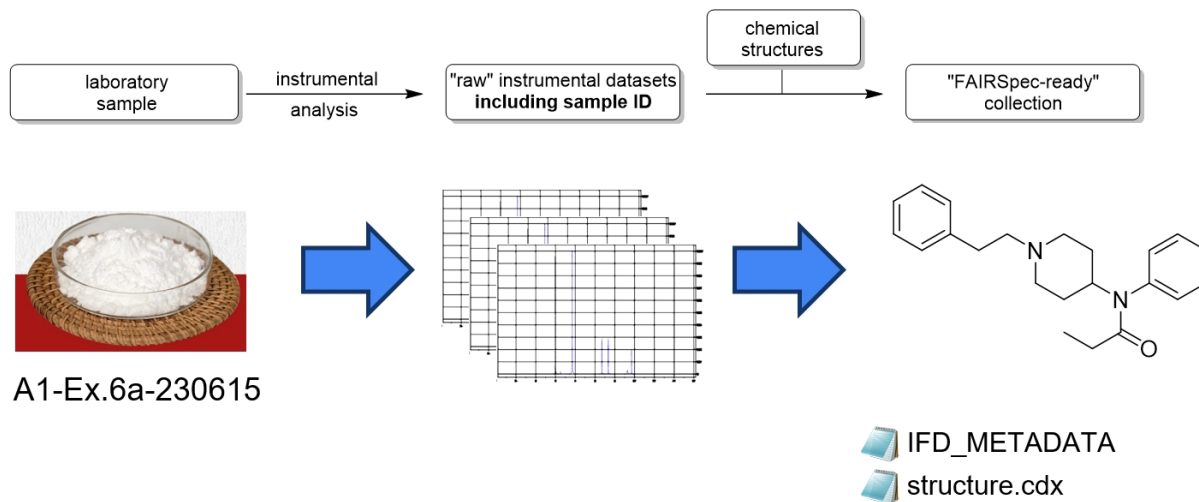
spectra



analysis



# The data management puzzle

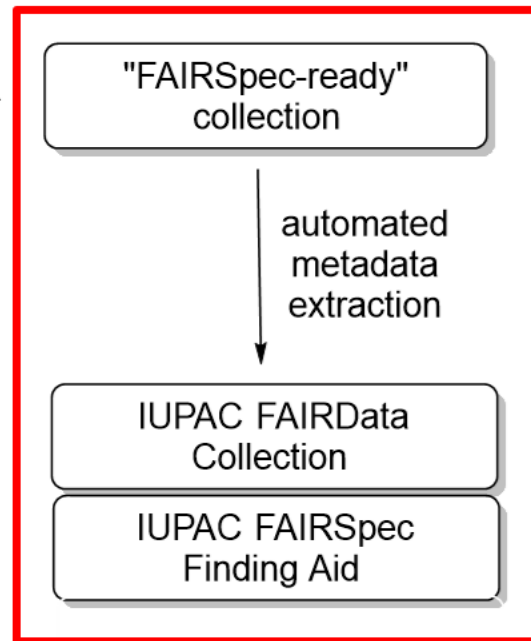


The goal of IUPAC Project 2019-031-1-024 is to develop standards for the maintenance and presentation of these relationships *throughout the entire research workflow*.

# The IUPAC FAIRSpec Solution

Our solution is to enable authors and data managers to create what we are calling a **“FAIRSpec-ready” collection**. Then let automation take that the rest of the way to an **IUPAC FAIRSpec Collection** with its associated **IUPAC FAIRSpec Finding Aid**.


The good news is that **we have tested this with over a dozen datasets**, *and it works*.



# Key concept – *Finding Aid*

a term borrowed from the field of archival science:

## ◆ AI Overview

A finding aid is a descriptive tool used in archives and special collections to help researchers navigate and understand the contents of a collection. It provides a detailed description of the collection, including its creator, history, scope, and organization, often down to the box and folder level. Think of it as a roadmap that helps researchers identify relevant materials within a collection. 

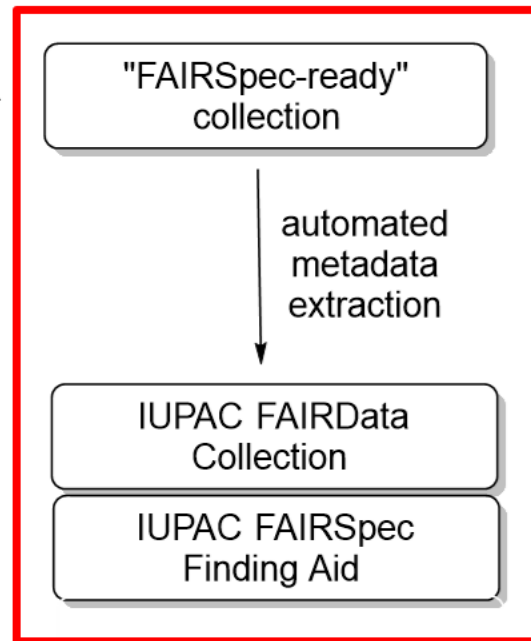
# Key concept – *Finding Aid*

An IUPAC FAIRSpec Finding Aid can be represented as a JSON document.

```
▼ IUPAC.FAIRSpec.findingAid:
  version: "IFD 0.1.0-beta+2025.07.24;FAIRSpec 0.1.0-beta+2025.07.23"
  created: "2025-08-20T05:37Z"
  createdBy: "https://github.com/IUPAC/IUPAC-FAIRSpec/blob/main/src/main/java/com/i"
  ► contents: { resourceCount: 2, collections: (5)[...] }
  ► resources: { 1: {...}, FAIRSpecDataCollection: {...} }
  ▼ collectionSet:
    propertyPrefix: "IFD.property.collectionset"
    resourceID: "FAIRSpecDataCollection"
    ► ifdProperties: { source_data_license_name: "cc-by-nc-4.0", source_data_license_uri: '
  ▼ itemsByID:
    ► samples: { itemsByID: {...} }
    ► structures: { itemsByID: {...} }
    ► spectra: { itemType: "org.iupac.fairdata.contrib.fairspec.dataobject.nmr.FAIRSp
      "org.iupac.fairdata.contrib.fairspec.dataobject.FAIRSpecDataObject;org
      itemsByID: {...} }
    ► sample-spectra associations: { itemsByID: {...} }
    ► compounds: { itemsByID: {...} }
```

# The IUPAC FAIRSpec solution

The focus of this presentation is to present several implementations of this process in different contexts.

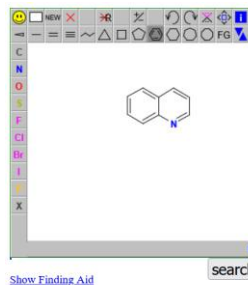


# Demonstrations

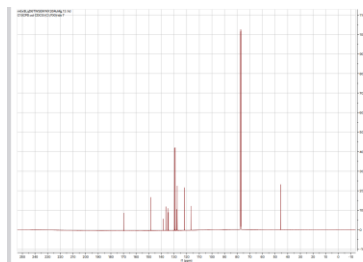
What follows are three demonstrations, in three very different contexts:

1. An undergraduate NMR laboratory using automation.
2. A set of publication spectroscopic supporting information package.
3. A repository holding shared data related to a publication.

All use the same HTML landing page; only the IUPAC FAIRSpec Finding Aids are different.



2 Compounds  
Compound 6. 2-phenyl-N-(quinolin-8-yl)acetamide  
Compound 8. tert-butyl 4-(quinolin-8-ylcarbamoyl)piperidine-1-carboxylate amide.



10.14469\_hpc\_14635

[summary](#) [search](#)

Collections: [Compounds\(51\)](#) [Structures\(51\)](#) [Spectra\(122\)](#)

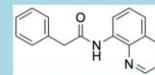
## Compound 6. 2-phenyl-N-(quinolin-8-yl)acetamide

10.14469/HPC/14712

6. 2-phenyl-N-(quinolin-8-yl)acetamide

Synthesized according to the general procedure. Purified by flash column chromatography (EtOAc : Cyclohexane 10% à 30%). Yield = 1.16g off-white solid, 4.42 mmol, 88%. Analytical data is in accordance with literature reports. <sup>1</sup>H NMR (700 MHz, CDCl<sub>3</sub>) δ (ppm) = 9.92 (s, 1H), 8.76 (d, J = 7.5 Hz, 1H), 8.69 (dd, J = 4.2, 1.5 Hz, 1H), 8.11 (dd, J = 8.2, 1.4 Hz, 1H), 7.51 (t, J = 7.9 Hz, 1H), 7.47 (dd, J = 8.1, 1 Hz, 1H), 7.45 (d, J = 7.5 Hz, 2H), 7.40 (m, 3H), 7.33 (t, J = 7.3 Hz, 1H), 3.90 (s, 2H). <sup>13</sup>C NMR (176 MHz, CDCl<sub>3</sub>) δ (ppm) = 169.7, 148.2, 138.5, 136.5, 134.8, 134.59, 129.7, 129.1, 128.0, 127.5, 127.5, 121.75, 121.7, 116.6, 45.5. <sup>1</sup>H NMR [C<sub>1</sub>H<sub>7</sub>N<sub>2</sub>O<sup>+</sup>H]<sup>+</sup>: Expected 263.1179, Observed 263.1178. Melting point: 71 - 73 °C

Predicted  
Spectra  
<sup>1</sup>H <sup>13</sup>C  
COSY, HMB



[inchi](#)  
[fixed\\_inchi](#)  
[smiles](#)  
[inchikey](#)  
[cdxml](#) [6.cd.cdxml](#) (6.6 KB)  
[mol](#) [6.cd.mol](#) (2.3 KB)

[9]

10.14469/HPC/14754  
6. 2-phenyl-N-(quinolin-8-yl)acetamide. <sup>13</sup>C  
Raw spectrometer NMR data  
[6-C.jds](#) (207.5 KB)  
[6-C.mnova](#) (423.6 KB)  
[6-C.pdf](#) (74.8 KB)  
[6-C.zip](#) (1 MB)

### IFD Properties...

<code>nmr.expt_absolute_temperature</code>	298.1491
<code>nmr.expt_description</code>	1D
<code>nmr.expt_dimension</code>	1D
<code>nmr.expt_id</code>	C13CPD.ucl
<code>nmr.expt_nucl1</code>	<sup>13</sup> C
<code>nmr.expt_nucl2</code>	<sup>1</sup> H
<code>nmr.expt_offset_freq1</code>	176.124279947708
<code>nmr.expt_offset_freq2</code>	700.3528014
<code>nmr.expt_pulse_program</code>	zgpg30
<code>nmr.expt_solvent</code>	CDCl <sub>3</sub>
<code>nmr.expt_solvent_inChI</code>	InChI=1S/CHCl3/c2-1(3)/h1H/t1D
<code>nmr.expt_solvent_inChIKey</code>	HEDRZPFGACZZDS-MICDWDJSA-N
<code>nmr.expt_solvent_common_name</code>	chloroform-d



# Demonstrations

Note that everything from here on in this presentation is something you can do yourself by going to our GitHub project:

<https://github.com/IUPAC/IUPAC-FAIRSpec>

or its associated GitHub pages website

<https://iupac.github.io/IUPAC-FAIRSpec>

## IUPAC FAIRSpec GitHub Project pages

Welcome to the GitHub project web pages for [IUPAC Project 2019-031-1-024, Development of a Standard for FAIR Data Management for Spectroscopic Data](#). At this site we highlight our development of IUPAC FAIRSpec Finding Aids.

See [publications](#) for published and submitted publications.

## CDX/CDXML specification

But, first, 352 pages retrieved from multiple points on the wayBack machine preserving the CDX/CDXML specification. (Note that some images were not retrievable.) This specification is important, as it is the basis for one of the most widely used and accessible formats in the chemistry community for communicating structural information. The specification is detailed and extensive, and has been implemented in Jmol, allowing an extractor to use it as a basis for creating value-added structure representations such as molecular formulas, InChI strings, SMILES, and MOL-format descriptions, which can be used both for display and validation. The pages are also available as a [ZIP file](#).



last updated 2025-06-24 (BH)

## IUPAC-FAIRSpec

Welcome to the GitHub project for [IUPAC Project 2019-031-1-024, Development of a Standard for FAIR Data Management for Spectroscopic Data](#).

Check out the [GitHub web pages site](#) for this project to see what we are talking about.

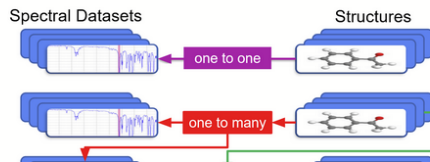
This is an active project. Our (rather minimal, admittedly) progress reports can be found in [documents/reports](#)

Our current working specification can be found at this site under [documentation/specifications](#).

This GitHub project provides public copies of all presentations and publications of the IUPAC Project, as well as a reference Java implementation of the Standard as a Java library as well as a reference Java implementation of an "IUPAC FAIRSpec data and metadata extractor". This GitHub project is currently under intensely active development. It is reasonably well tested (as of 11/2024) and, though public, is only meant for demonstration purposes. **implement these preliminary standards with this in mind** as they periodically change as new issues emerge.

The principal goal of the project is to define standardized metadata associated with complex collections of spectroscopic data in the area of chemistry -- NMR, IR, Raman, MS, etc. The specification is modular and has been worked out primarily in the area of NMR spectroscopy at this time.

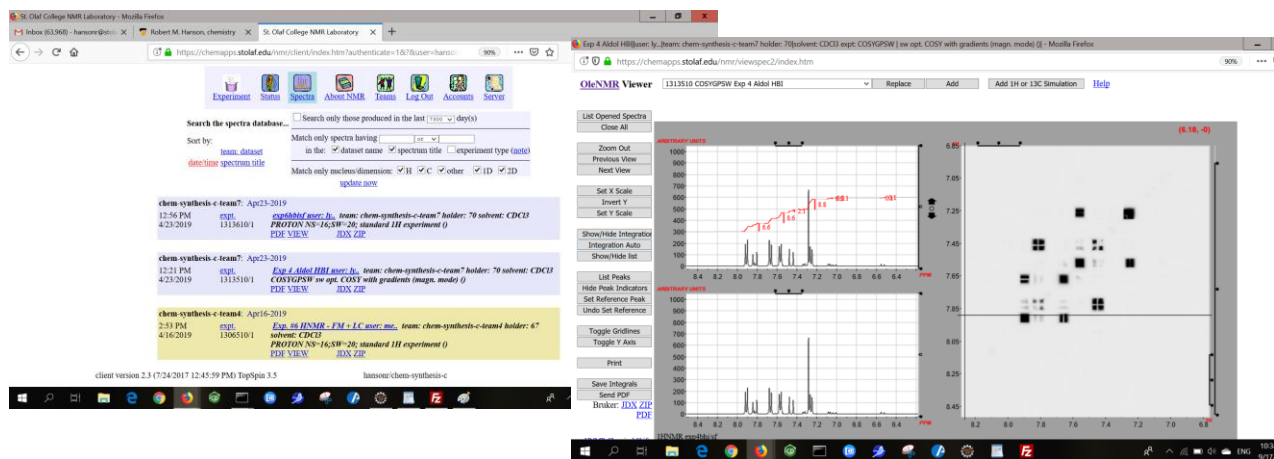
### One to One and One to Many FAIR Relationships



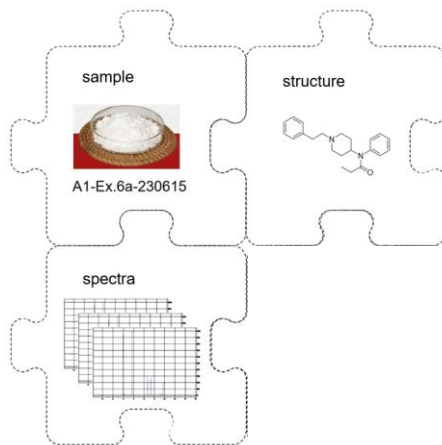
# Demonstration 1

St. Olaf College Organic Chemistry I laboratory course.

Fully automated web-based interface to IconNMR



## Demo 2024.1 [examples2/v6-stolaf](#)



<https://iupac.github.io/IUPAC-FAIRSpec>

# Demonstration 2

Fourteen datasets deposited as supporting information for publications.

Each finding aid is fully searchable for text, properties, and substructures.

## Demo 2025.1 [examples2/v6-acs](#)

**Property Search**

☐ compounds ☐ structures ☒ spectra

☐ 1\_nmr.expt\_dimension

- ☐ 1D (48)
- ☐ 2D (6)

☐ 1\_nmr.expt\_id

- ☐ unspecified (27)
- ☐ C13CPD (2)
- ☐ NUS\_HMBC.kaist
- ☐ NUS\_HSQC.kaist (2)
- ☐ PROTON (2)

☐ 1\_nmr.expt\_nucl2

- ☐ unspecified (48)
- ☐ 13C (6)

☐ 1\_nmr.expt\_offset\_freq1

- ☐ 100.62031505 (12)
- ☐ 100.622829328806 (12)
- ☐ 400.12240072 (2)
- ☐ 400.12247074 (19)
- ☐ 400.13280091 (8)

☐ 1\_nmr.expt\_offset\_freq2

- ☐ unspecified (36)
- ☐ 100.617799793 (4)
- ☐ 100.62031505 (2)
- ☐ 400.12160048 (7)
- ☐ 400.13160052 (5)

☐ 1\_nmr.expt\_pulse\_program

- ☐ hmbcgndqf (2)
- ☐ hsqcdetgp (4)
- ☐ zg30 (24)
- ☐ zgpg30 (24)

☐ 1\_nmr.expt\_solvent

- ☐ Acetone (10)
- ☐ C6D6 (4)
- ☐ CDCl3 (40)

☐ 1\_nmr.expt\_solvent\_InChI

- ☐ InChI=1S/C3H6O/c1-3(2)4/h1-2H3/1D3,2D3 (10)
- ☐ InChI=1S/C6H6/c1-2-4-6-5-3-1/h1-6H/1D,2D,3D,4D,5D,6D (4)
- ☐ InChI=1S/CHCl3/c2-1(3)4/h1H/1D (40)

☐ 1\_nmr.expt\_solvent\_InChIKey

- ☐ CSCPPACGZOCGX-WFGJKAKNSA-N (10)
- ☐ HEDRZPFGACZZDS-MICDWDOSA-N (40)
- ☐ UHOVQNZJYSORNB-MZWXYZOWSA-N (4)

☐ 1\_nmr.expt\_solvent\_common\_name

- ☐ acetone-d6 (10)
- ☐ benzene-d6 (4)
- ☐ zg30 (24)
- ☐ chloroform-d (40)

structure

c1ccc(cc1)C2=CC(=O)N(C2)Cc3ccccc3

spectra

This page and its associated [IUPAC FAIRSpec Finding Aid](#) were automatically generated by IFDEExtractor.java ([GitHub site](#)). It is a demonstration page for [IUPAC Project 2019-031-1-024](#). Development of a Standard for FAIR Data Management of Spectroscopic Data.

Select a Finding Aid ▾

[acs.joc.0c00770](#) [view IFD.findingaid.json](#) [IFD.collection.zip \(25.2 MB\)](#)

**Collections:** [Compounds\(11\)](#) [Structures\(11\)](#) [Spectra\(54\)](#)  
**Title** [Total Synthesis of \(+\)-Pentalone A and \(+\)-Iso-A82773C](#)  
**Authors** [Geon Kim; Taewan Kim; Sunkyu Han](#)  
**DOI** [https://doi.org/10.1021/acs.joc.0c00770 \(metadata\)](#)  
**Data Origin** [https://indownloader.figshare.com/files/22567817/24.5 MB](#)  
**FAIRSpecDataCollection** [\\_acs.joc.0c00770\\_/IFD.collection.zip \(25.2 MB\)](#)

[acs.orglett.0c00571](#) [view IFD.findingaid.json](#) [acs.orglett.0c00571-IFD.collection.zip \(199.6 MB\)](#)

**Collections:** [Compounds\(20\)](#) [Structures\(30\)](#) [Spectra\(114\)](#)  
**Title** [Synthesis of Novel Heterocycles by Amide Activation and Unpinning Cyclization](#)  
**Authors** [Haoqi Zhang; Margaux Riomet; Alexander Roller; Nuno Maulide](#)  
**DOI** [https://doi.org/10.1021/acs.orglett.0c00571 \(metadata\)](#)  
**Data Origin** [https://indownloader.figshare.com/files/21975525 \(189.9 MB\)](#)  
**FAIRSpecDataCollection** [https://chemapps.stolaf.edu/iupac/examples2/v6-acs/acs.orglett.0c00571-IFD.collection.zip \(199.6 MB\)](#)

[acs.orglett.0c00624](#) [view IFD.findingaid.json](#) [IFD.collection.zip \(15.9 MB\)](#)

**Collections:** [Compounds\(41\)](#) [Structures\(42\)](#) [Spectra\(80\)](#)  
**Title** [Intermolecular Vicinal Diaminative Assembly of Tetrahydroquinolines via Metal-free Oxidative \[4 + 2\] Cycloaddition Strategy](#)  
**Authors** [Dangui Wang; Huabin Yu; Shaohan Sun; Fangrui Zhong](#)  
**DOI** [https://doi.org/10.1021/acs.orglett.0c00624 \(metadata\)](#)  
**Data Origin** [https://indownloader.figshare.com/files/21947274 \(15.2 MB\)](#)  
**FAIRSpecDataCollection** [\\_acs.orglett.0c00624\\_/IFD.collection.zip \(15.9 MB\)](#)

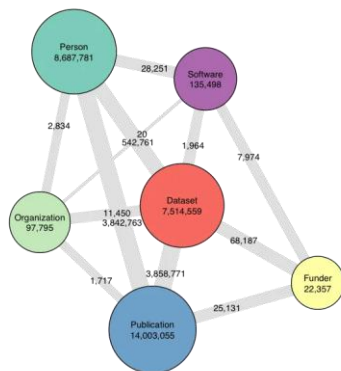
[acs.orglett.0c00755](#) [view IFD.findingaid.json](#) [IFD.collection.zip \(105.7 MB\)](#)

**Collections:** [Compounds\(26\)](#) [Structures\(28\)](#) [Spectra\(78\)](#)  
**Title** [Synthesis of a P-Glycoprotein Inhibitor and Its High-Energy \(2\)-Isomer by Carbenoid Eliminative Cross-Coupling](#)  
**Authors** [Subhash D. Tanspur; Mohamed F. El-Mansry; Paul R. Blakemore](#)  
**DOI** [https://doi.org/10.1021/acs.orglett.0c00755 \(metadata\)](#)  
**Data Origin** [https://indownloader.figshare.com/files/22150197 \(105.6 MB\)](#)  
**FAIRSpecDataCollection** [https://chemapps.stolaf.edu/iupac/examples2/v6-acs/acs.orglett.0c00755-IFD.collection.zip \(105.7 MB\)](#)

# Demonstration 3

A finding aid created by “crawling” a chain of metadata documents at DataCite.

The metadata was richly annotated manually with links to related documents.



IMPERIAL

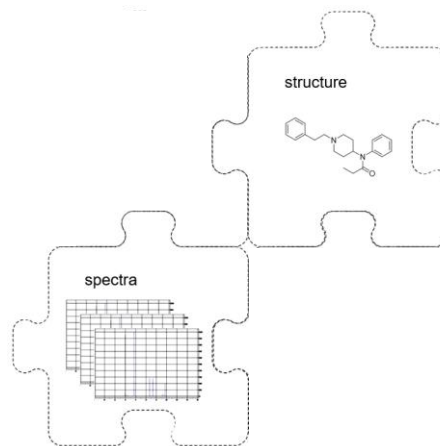
## Research Data Repository

Browse Add Collection Deposit Data Help Search Engine: Google Search Query:

## Syntheses and Characterization of Main Group, Transition Metal, Lanthanide and Actinide Complexes of Bidentate Acylpyrazolone Ligands

DOI: [10.14469/hpc/10386](https://doi.org/10.14469/hpc/10386) Metadata

Demo 2025.6 [examples2/v6-crawler2](#)



10.14469\_hpc\_10386

[summary](#) [search](#)

Collections: [Compounds\(57\)](#) [Structures\(57\)](#) [Spectra\(228\)](#)

Title *Syntheses and Characterization of Main Group, Transition Metal, Lanthanide and Actinide Complexes of Bidentate Acylpyrazolone Ligands*

Authors [Thomas Mies](#), [Andrew J. P. White](#), [Henry S. Rzepa](#), [Luciano Barluzzi](#), [Layfield](#), [Anthony G. M. Barrett](#)

DOI <https://doi.org/10.1021/acs.inorgchem.3c01506> (metadata)

Data Title *Syntheses and Characterization of Main Group, Transition Metal, Lanthanide and Actinide Complexes of Bidentate Acylpyrazolone Ligands*

Data DOI <https://doi.org/10.14469/HPC/10386> (metadata)

Description The synthesis of acylpyrazolone salts and their complexes of main group elements, lanthanides, and actinides are described and characterized inter alia by means of X-ray crystallography, NMR, and IR spectroscopies. The complexes consist of two bidentate ligands bound to the metal atom, resulting in a structurally diverse set of complexes (distorted) octahedral, pentagonal-bipyramidal, or antiprismatic arrangements. Some complexes are also characterized in the solid state including heteropolymers/lanthanide complexes.

<https://iupac.github.io/IUPAC-FAIRSpec>

# Thank you!

Note that everything from here on in this presentation is something you can do yourself by going to our GitHub project:

<https://github.com/IUPAC/IUPAC-FAIRSpec>

or its associated GitHub pages website

<https://iupac.github.io/IUPAC-FAIRSpec>

## IUPAC FAIRSpec GitHub Project pages

Welcome to the GitHub project web pages for [IUPAC Project 2019-031-1-024, Development of a Standard for FAIR Data Management for Spectroscopic Data](#). At this site we highlight our development of IUPAC FAIRSpec Finding Aids.

See [publications](#) for published and submitted publications.

## CDX/CDXML specification

But, first, 352 pages retrieved from multiple points on the wayBack machine preserving the CDX/CDXML specification. (Note that some images were not retrievable.) This specification is important, as it is the basis for one of the most widely used and accessible formats in the chemistry community for communicating structural information. The specification is detailed and extensive, and has been implemented in Jmol, allowing an extractor to use it as a basis for creating value-added structure representations such as molecular formulas, InChI strings, SMILES, and MOL-format descriptions, which can be used both for display and validation. The pages are also available as a [ZIP file](#).



last updated 2025-06-24 (BH)

## IUPAC-FAIRSpec

Welcome to the GitHub project for [IUPAC Project 2019-031-1-024, Development of a Standard for FAIR Data Management for Spectroscopic Data](#).

Check out the [GitHub web pages site](#) for this project to see what we are talking about.

This is an active project. Our (rather minimal, admittedly) progress reports can be found in [documents/reports](#)

Our current working specification can be found at this site under [documentation/specifications](#).

This GitHub project provides public copies of all presentations and publications of the IUPAC Project, as well as a reference Java implementation of the Standard as a Java library as well as a reference Java implementation of an "IUPAC FAIRSpec data and metadata extractor". This GitHub project is currently under intensely active development. It is reasonably well tested (as of 11/2024) and, though public, is only meant for demonstration purposes. **implement these preliminary standards with this in mind** as they periodically change as new issues emerge.

The principal goal of the project is to define standardized metadata associated with complex collections of spectroscopic data in the area of chemistry -- NMR, IR, Raman, MS, etc. The specification is modular and has been worked out primarily in the area of NMR spectroscopy at this time.

### One to One and One to Many FAIR Relationships

