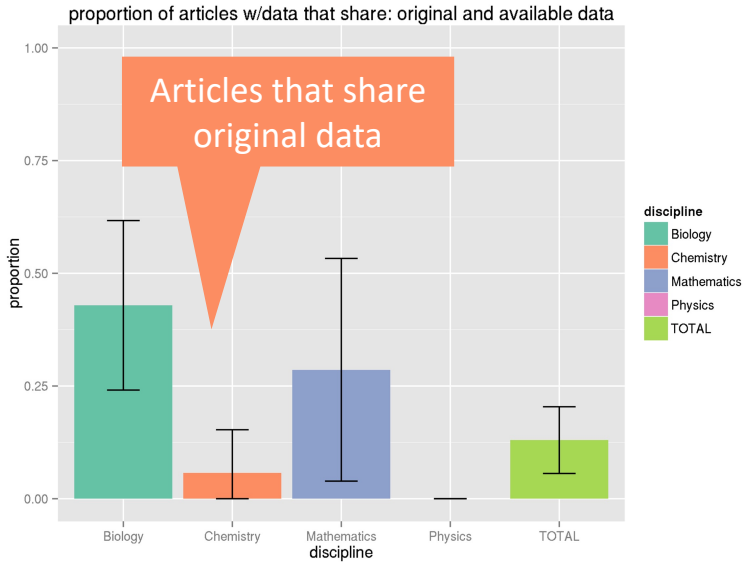
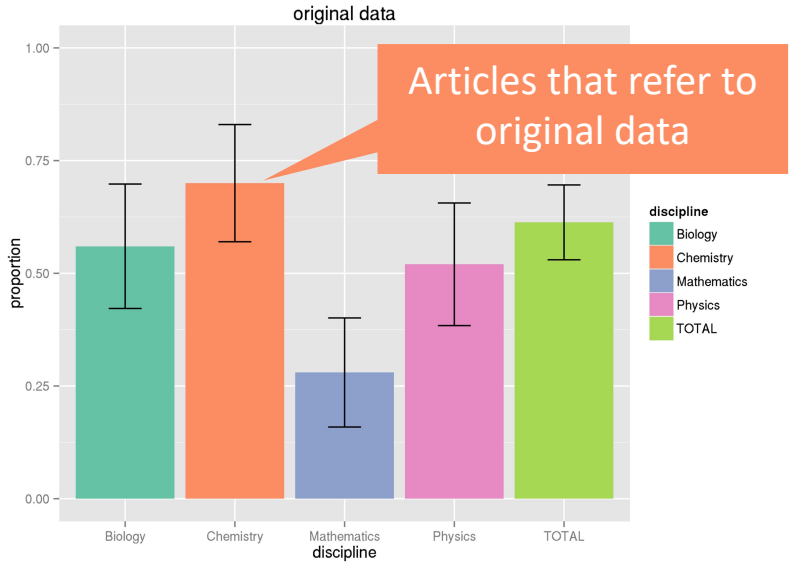
# Journal Data Policies: Getting started with a Framework and Examples from Society Journals

Data Sharing Seminar Series

2021-07-02

Leah McEwen, moderator

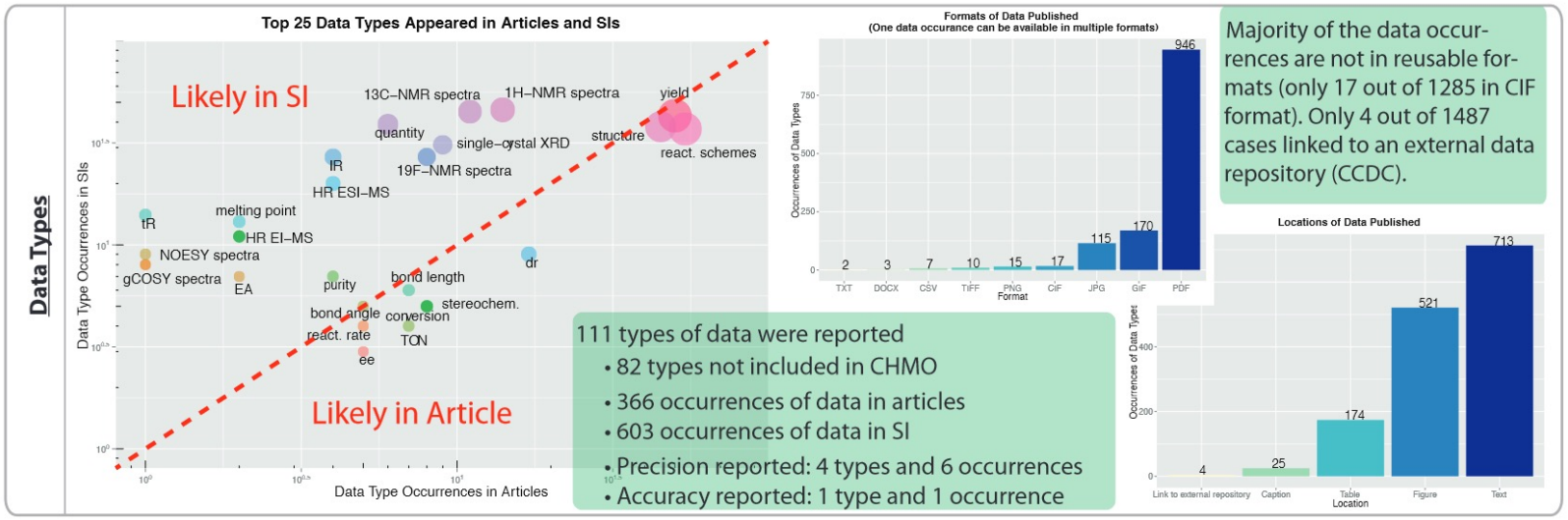



**Womack, Ryan P. (2015).** Research Data in Core Journals in Biology Chemistry, Mathematics, and Physics. *PLOS One*. https://doi.org/10.1371/journal.pone.0143460

# Chemists commonly share data…

**… but primarily in Supporting Information as a PDF or as a static image in the manuscript…**




***Thielen, Joanna, & Li, Ye. (2015).*** *Profiling common types of research data and methods published by organic synthesis chemists at the University of Michigan. Paper presented at the SLA 2015 Annual Conference & Info Expo, Boston, MA.* http://hdl.handle.net/2027.42/111832

CHMO = Chemical Methods Ontology

# Different Chemistry Data Types

Different Chemistry Journals

| Crystallography Data | | Spectra | | | | Chromatographs | Purity and Physical Property Data | | | | Other Data (examples) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Single Crystal XRD | Powder XRD | NMR | IR | UV-vis | Mass | GC, HPLC | Elemental Analysis HRMS | Melting/Boiling Point range | Optical rotation | Kinetic/Equilibrium | Protein/Nucleic Acid XRD or NMR sequence data | Magnetic | Computational |
| RV; R | R; M | O; M | O; M | O; M | O; M | E | R; M | R; M | O; M | R; M | R; R | R; M | R; M |
| RV; R | R; M | O; M | O; M | O; M | O; M | E | R; M | R; M | O; M | R; M | R; R | R; M | R; M |
| RV; R | E | O; M | O; M | O; M | O; M | E | R; M | R; M | O; M | E | R; R | E | R; M |
| RV; R | E | O; M | O; M | O; M | O; M | E | | | | O; M | R; R | E | R; M |
| R; R | E | R; M | R; M | R; M | R; M | E | | | | E | R; R | E | E |
| R; R | E | R; M | R; M | R; M | | | | | | E | R; R | E | O; R |
| R; R | E | R; M | R; M | R; M | R; M | E | | | | E | R; R | E | O; R |
| R; R | R; R | E | E | E | E | E | E | E | E | E | R; R | E | E |
| RV; R | O; M | R; M | R; M | R; M | R; M | O; M | R; M | O; M | O; M | E | R; R | R; M | R; M |
| RV; R | O; M | R; M | R; M | R; M | R; M | O; M | | | | E | R; R | E | E |
| R; R | R; R | R; M | R; M | R; M | R; M | E | O; M | O; M | E | E | E | E | E |
| R; R | E | R; M | R; M | R; M | R; M | O; M | R; M | R; M | R; M | E | E | E | E |
| R; R | R; R | R; M | R; M | R; M | R; M | E | O; M | O; M | E | E | E | E | E |
| RV; R | E | R; M | O; M | O; M | O; M | O; M | O; M | O; M | E | E | R; R | E | O; M |
| O; R | E | E | E | E | E | E | E | E | E | E | O; R | E | O; M |
| RV; R | E | R; M | R; M | R; M | R; M | E | R; M | E | E | E | RV; R | E | E |
| R; R | E | E | E | E | E | E | E | E | E | E | E | E | E |

**Crystallography / Structural Biology**

**Spectroscopy**

- *Vincent Scalfani, RDA CRDIG Open Meeting, ACS Spring Meeting, San Francisco, March 2017*
  https://doi.org/10.6084/m9.figshare.8870144.v1

# What about Chemistry journal data policies?

**Journal Policy**

**Required in Repository**

**Required in Manuscript**

**Optional in Manuscript**

**Not required**

# FAIR Chemical Data
## Publishing Guidelines Workshop
## Chemical Structures and Spectra

29-30 March 2019, Orlando FL

**Workshop Chairs**: Leah McEwen and Vincent Scalfani
**Advisors:** Angie Hunter, Ian Bruno, Guy Jones, and Dave Martinsen
NSF OAC – Award No. 1838958 and 1838960

**WORKSHOP GOALS:**

1. **Workflow**: Develop a digital data publishing model across stakeholders
2. **Guidelines**: Formulate consistent guidelines for publishing FAIR chemical data for common data types
3. **Value Proposition**: Review re-use cases for chemical characterization data
4. **Coalition**: Initiate process for ongoing coordination and stakeholder engagement

Publishers • Databases • Repositories • Software Developers
Researchers • Librarians • Standards Organisations • Data Initiatives

*Scalfani VF, McEwen L. NSF OAC 2019 Workshop: FAIR Publishing Guidelines for Spectral Data and Chemical Structures. Published 2019. https://osf.io/psq7k/*

# TOWARDS FAIRER SPECTRA DATA

## KEY OUTCOMES

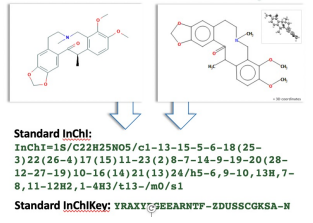
- **FAIR Spectra Publishing Pilot**
  - Pilot encouraging submission of FAIR spectra data to ACS Journals https://doi.org/10.1021/acs.orglett.0c00383
- **IUPAC Project to define a standard metadata specification for Spectra**
  - Format-agnostic https://iupac.org/project/2019-031-1-024
  - Describes the contents of a spectroscopic data collection
  - Describes structures and analyses relating to the data
- **Draft guidelines for chemists to publish chemical structure and spectra data files**

Original draft: https://osf.io/fqd82/
Modified with supporting materials: https://osf.io/vcsnp/

# IUPAC standards that support FAIR data activities

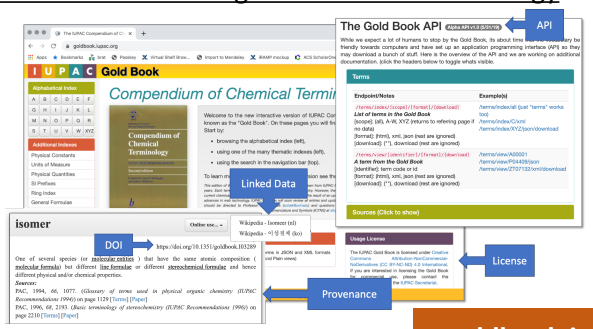
**InChI**: IUPAC International Chemical Identifier

- Small molecule organics
- Reactions
- Mixtures *(provisional)*
- QR Code spec *(provisional)*
- In progress:
  - Organometallics
  - Large molecules
  - Stereochemistry
  - Tautomers




www.inchi-trust.org

**Gold Book**: IUPAC Digital Chemical Terminology




goldbook.iupac.org

**IUPAC Data Standards – FAIR updates**

- Thermophysical & chemical properties
  - ThermoML XML format (update in progress)
  - 2 million data points
- Spectroscopic data
  - JCAMP-DX ASCII legacy format
  - Current project development: format-agnostic, FAIR minimal metadata standard

trc.nist.gov/ThermoML.html

**In progress areas**

- Machine processable critically evaluated data compilations:
  - Isotopic Abundances and Atomic Weights
  - Solubility
- Digital representations:
  - SMILES+ (standardizing interpretation)
  - Units of measure
  - Physical constants

**Other activities**

- Cheminformatics Standards in Practice - *special issue*
- Data reporting guidelines
- International symposia and workshops

## Key enablers of FAIR

- Persistent Identifiers
- Rich Metadata
- Repositories
- Standard Open Protocols
- Knowledge Representation
- FAIR Vocabularies
- Linked Data
- Usage Licences
- Provenance
- Community Standards

Committee on Publications and Cheminformatics Data Standards: iupac.org/committee/024

# Preparing data files and manuscript in parallel