NFDI4Chem

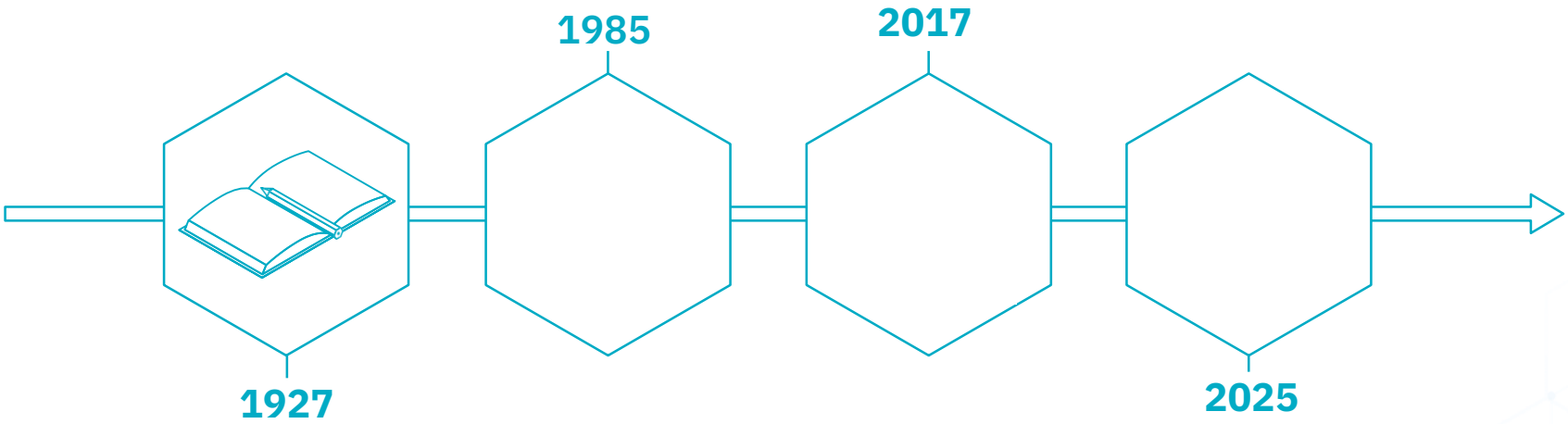
ENHANCE
YOUR
DATA.

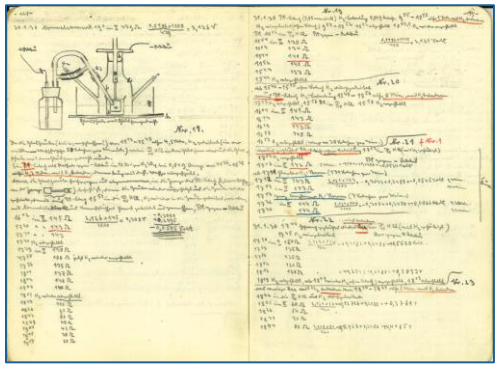
Unlocking Chemical Research Data - The NFDI4Chem Service & Infrastructure Federation

Oliver Koepler, TIB, RSC Webinar „Data Management“, 18 Nov 2025

NFDI4Chem is supported by DFG
under project number 441958208

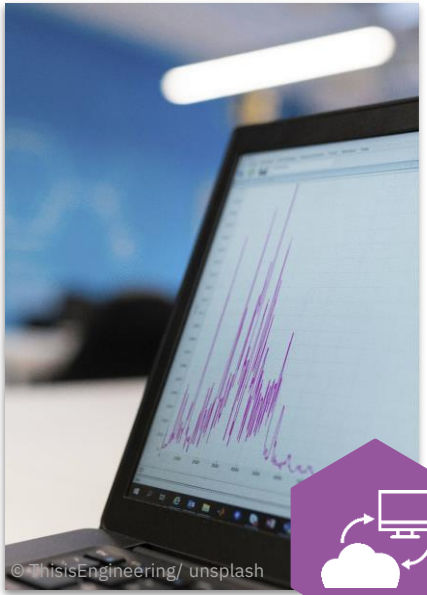
Speed of Evolution vs. Digitisation in Chemistry



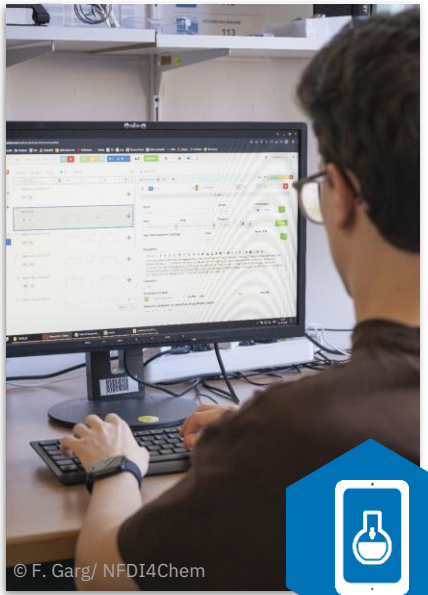


Supporting Users from Data Creation to Re-use

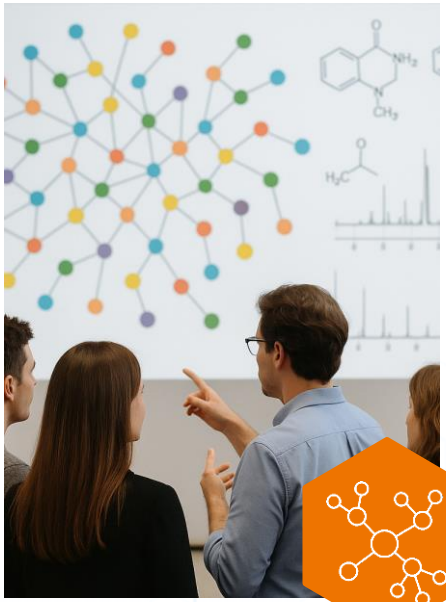
Create Data
Smart Lab



Publish & Access Data
Repositories



Find & Use Data
Metadata, Semantics & Standards



Data Discovery & Reuse
Semantic Data Hub

NFDI4Chem is community-driven

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN
IN PUBLICA COMMODA SEIT 1737

TIB
LEIBNIZ INFORMATION CENTRE FOR SCIENCE AND TECHNOLOGY UNIVERSITY LIBRARY

WWU MÜNSTER

PTB
Physikalisch-Technische Bundesanstalt
National Metrology Institute

technische universität dortmund

RWTH AACHEN UNIVERSITY

GDCh
GESELLSCHAFT DEUTSCHER CHEMIKER

DPhG
Deutsche Pharmazeutische Gesellschaft e.V.

BEILSTEIN INSTITUT

JOHANNES GUTENBERG UNIVERSITÄT MAINZ

JG|U

RPTU
Rheinland-Pfälzische Technische Universität
Kaiserslautern
Landau

KIT
Karlsruhe Institute of Technology

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

Universität Stuttgart

universität ulm

Technische Universität Braunschweig

FREIE UNIVERSITÄT

BAM

UNIVERSITÄT LEIPZIG

UFZ HELMHOLTZ
Centre for Environmental Research

Center for the Transformation of Chemistry

TECHNISCHE UNIVERSITÄT DRESDEN

IPB
Leibniz Institute of Plant Biochemistry

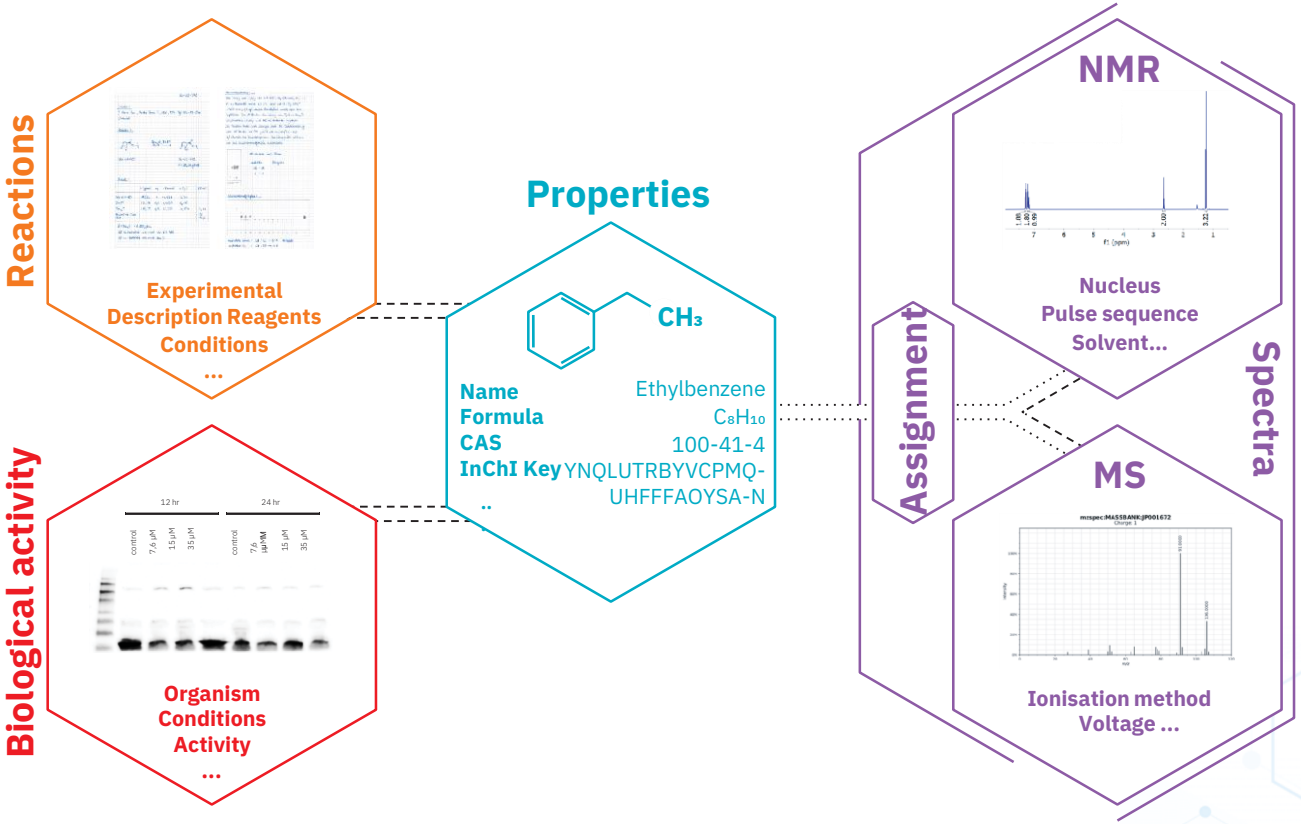
FRIEDRICH-SCHILLER-UNIVERSITÄT JENA

Julius-Maximilians-
UNIVERSITÄT WÜRZBURG

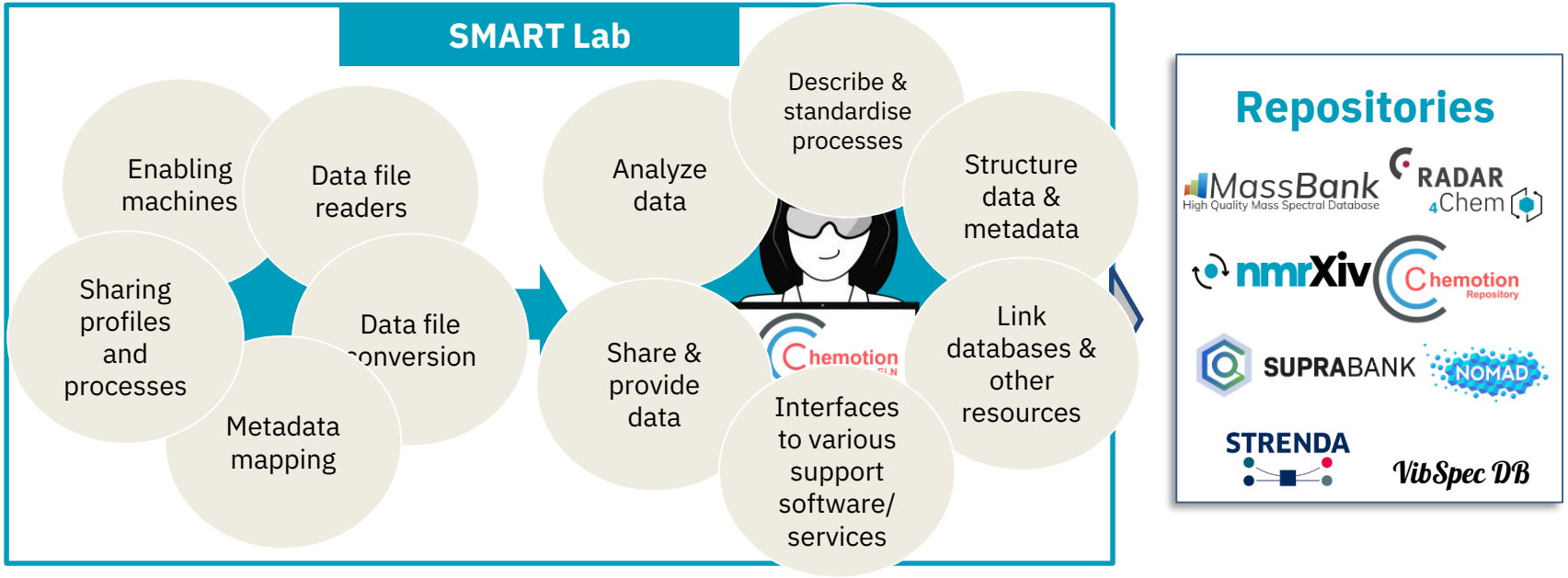
Leibniz ipht
LEIBNIZ-INSTITUT für PHOTONISCHE TECHNOLOGIEN

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

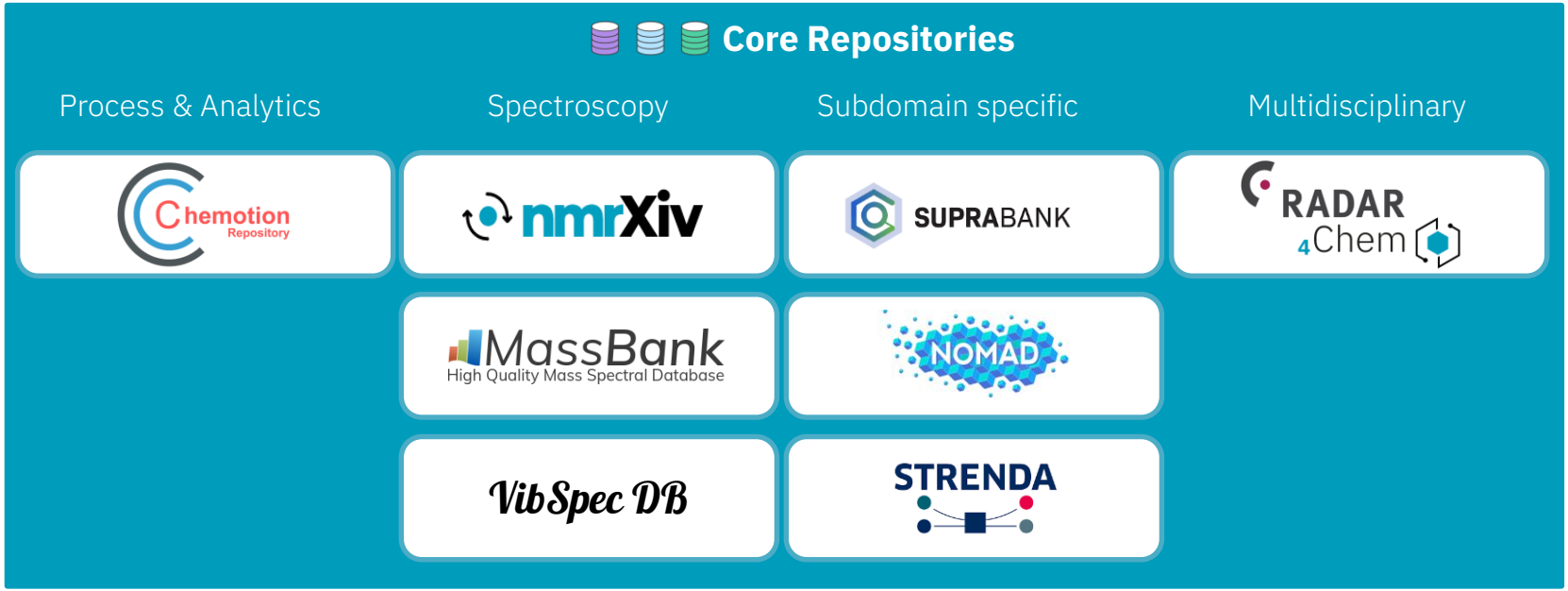
Data we cover - Molecules and related meta(data)



Smart Lab - Enabling Digitisation and Seamless Data Flows



Trusted Federation of Repositories



Harmonisation across resources - MIChI

Chemists, RDM Experts & Computer Scientists develop Minimum Information for Chemical Investigations

Guideline Article


DOI 10.1007/s11306-007-0082-2

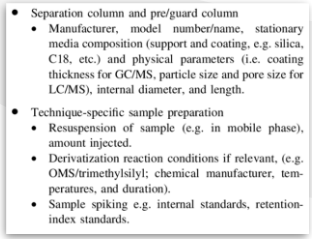
ORIGINAL ARTICLE

Proposed minimum reporting standards

Chemical Analysis Working Group (CAWG) Metabol

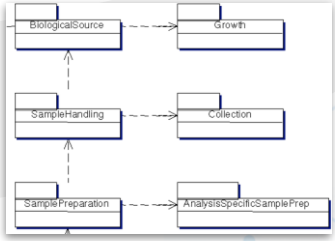
Lloyd W. Sumner · Alexander Amberg · Dave Barrett · Michael
Richard Beger · Clare A. Daykin · Teresa W.-M. Fan · Oliver Fi
Royston Goodacre · Julian L. Griffin · Thomas Hankemeier · Nig
James Harnly · Richard Higashi · Joachim Kopka · Andrew N. L
John C. Lindon · Philip Marriott · Andrew W. Nicholls · Michael
John J. Thaden · Mark R. Viant

Checklist

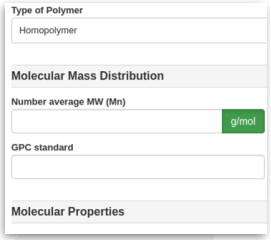

- Separation column and pre/guard column
  - Manufacturer, model number/name, stationary media composition (support and coating, e.g. silica, C18, etc.) and physical parameters (i.e. coating thickness for GC/MS, particle size and pore size for LC/MS), internal diameter, and length.
- Technique-specific sample preparation
  - Resuspension of sample (e.g. in mobile phase), amount injected.
  - Derivatization reaction conditions if relevant, (e.g. OMS/trimethylsilyl; chemical manufacturer, temperatures, and duration).
  - Sample spiking e.g. internal standards, retention-index standards.

MIChI

Data Model



Implementation


Type of Polymer

Homopolymer

Molecular Mass Distribution

Number average MW (Mn)

g/mol

GPC standard

Molecular Properties

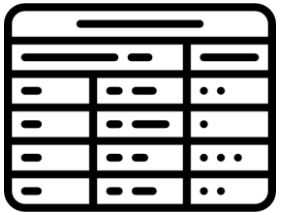
MIChI Development Workflow

Review resources

Create maximum information

Create minimum information draft

Community workshop



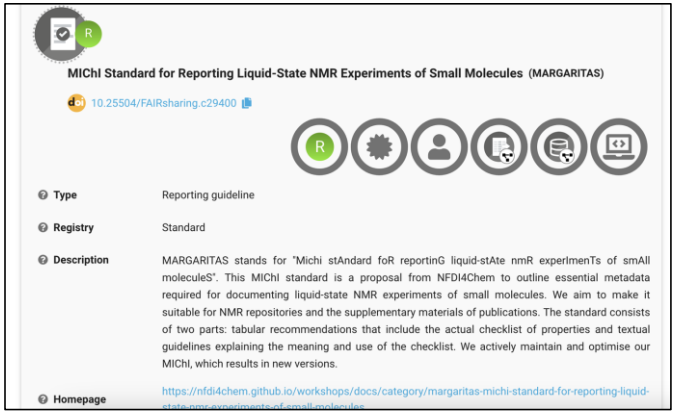
MIChI specifications

FAIRsharing.org
standards, databases, policies

Publication

MIChI in NFDI4Chem

NMR MIChIs: on FAIRSharing
doi.org/10.25504/FAIRsharing.c29400



MIChI Standard for Reporting Liquid-State NMR Experiments of Small Molecules (MARGARITAS)

10.25504/FAIRsharing.c29400

Type	Reporting guideline
Registry	Standard
Description	MARGARITAS stands for "Michi stAndard foR reportinG liquid-stAte nmR experImenTs of smAll moleculeS". This MIChI standard is a proposal from NFDI4Chem to outline essential metadata required for documenting liquid-state NMR experiments of small molecules. We aim to make it suitable for NMR repositories and the supplementary materials of publications. The standard consists of two parts: tabular recommendations that include the actual checklist of properties and textual guidelines explaining the meaning and use of the checklist. We actively maintain and optimise our MIChI, which results in new versions.
Homepage	https://nfdi4chem.github.io/workshops/docs/category/margaritas-michi-standard-for-reporting-liquid-state-nmr-experiments-of-small-molecules

No.	Property	Definition
0	Differential optical absorption spectrophotometry	
0.1	UV-Vis Differential Optical Absorption Sample	
0.1.1	Chemical structure	Chemical structure identifier in the form of an alphanumeric text string
0.1.2	Solvent	
0.1.3	Concentration	
0.1.4	Optical path length	
0.1.5	Sample ID	Identifier for the sample which is unique within the project

UV-Vis MIChI: Community workshop

No.	Property	Definition
1	Chemical Entity	Chemical Entity (CHEBI:24431)
1.1	Identification	
1.1.1	Primary Identifier	A globally unique, persistent value assigned to an entity that distinguishes it from all other entities in a dataset or system. It is a dcterm Identifier
1.1.2	Title/ name/ label	dcterm Title
1.1.3	Alternative identifer	An identifier that refers to the same entity as the primary identifier but comes from a different identification system, source, or context. Not the main reference. It is a dcterm Identifier
1.1.4	Alternative Title name/ label	

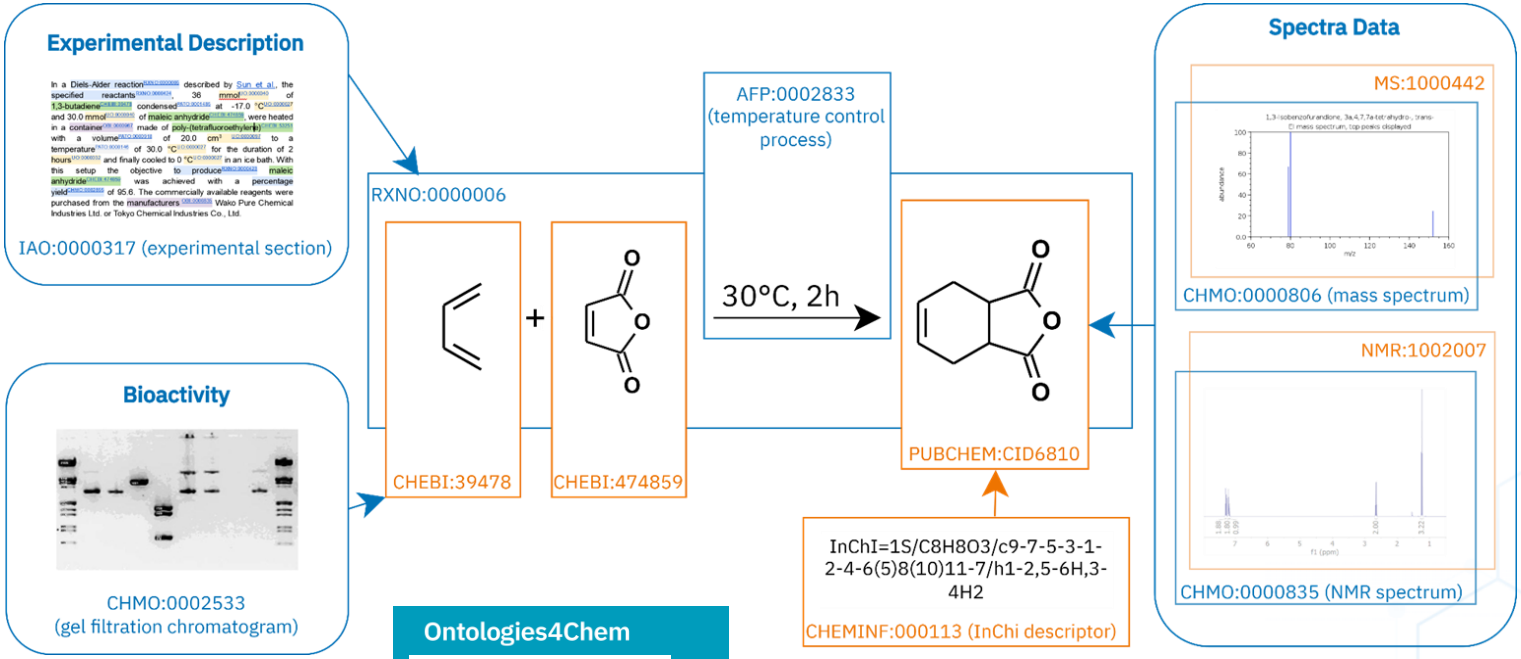
Chemical entity, MS: Minimum information draft

No.	Property	Definition	Export Format
1	Mass Spectrometry Assay		
1.1	MS Sample		
1.1.1	Analysed Compound and/or Sample ID	A compound analysed in an mass spectrometry (MS) sample using data acquired from an mass spectrometry (MS) analysis	- mol - ChEBI ID or IRI that is a leaf of the ontology term molecular entity (CHEBI:23367) or - Pubchem ID
			ChEBI ID or IRI that is a leaf of the ontology cular entity (CHEBI:23367)
			with unit: unit (UO:1000099)

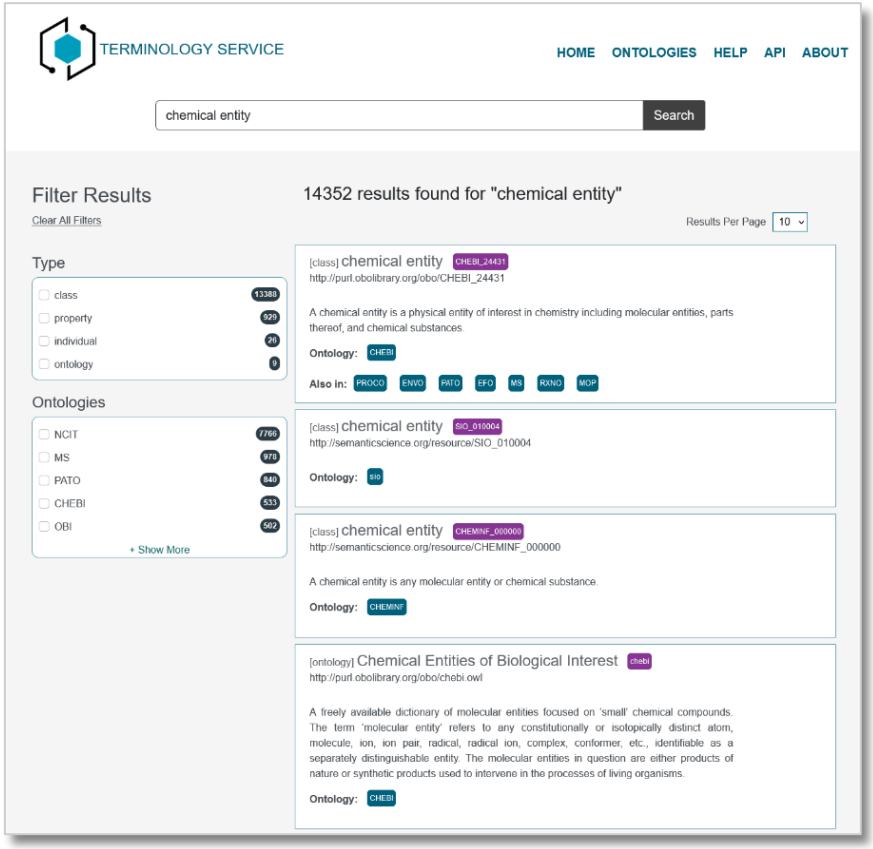
Metadata	dx	RRUFF	Prof Bocklitz
Acquisition			
method			
npoints	✅		
sampling procedure	✅		
laser wavelength		✅	✅
instrument settings		✅	
orientation		✅	
laser power			✅
accumulation time			
spectral range			✅
resolution	✅		

Raman: Max Information

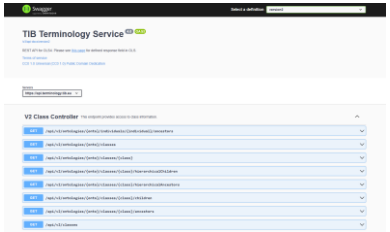
Harmonisation across resources - Ontologies



The NFDI4Chem Terminology Service (TS)



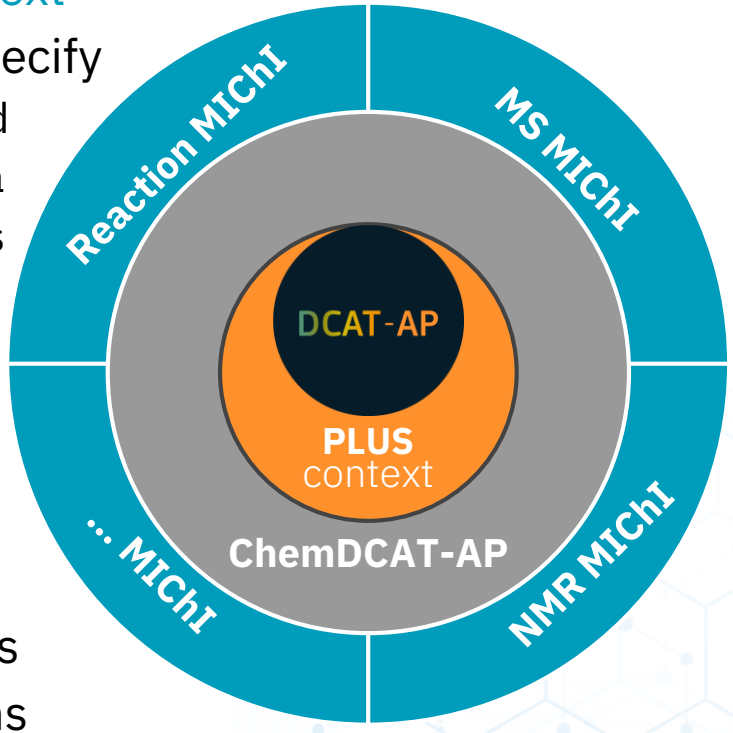
- Search in all indexed fields
  - E.g. abstract, label, definition
- Filter by type of term
  - Class, property, individual or ontology
- Filter by ontology
  - Sorted by number of hits
- API available for service integration



Harmonisation across resources – Metadata

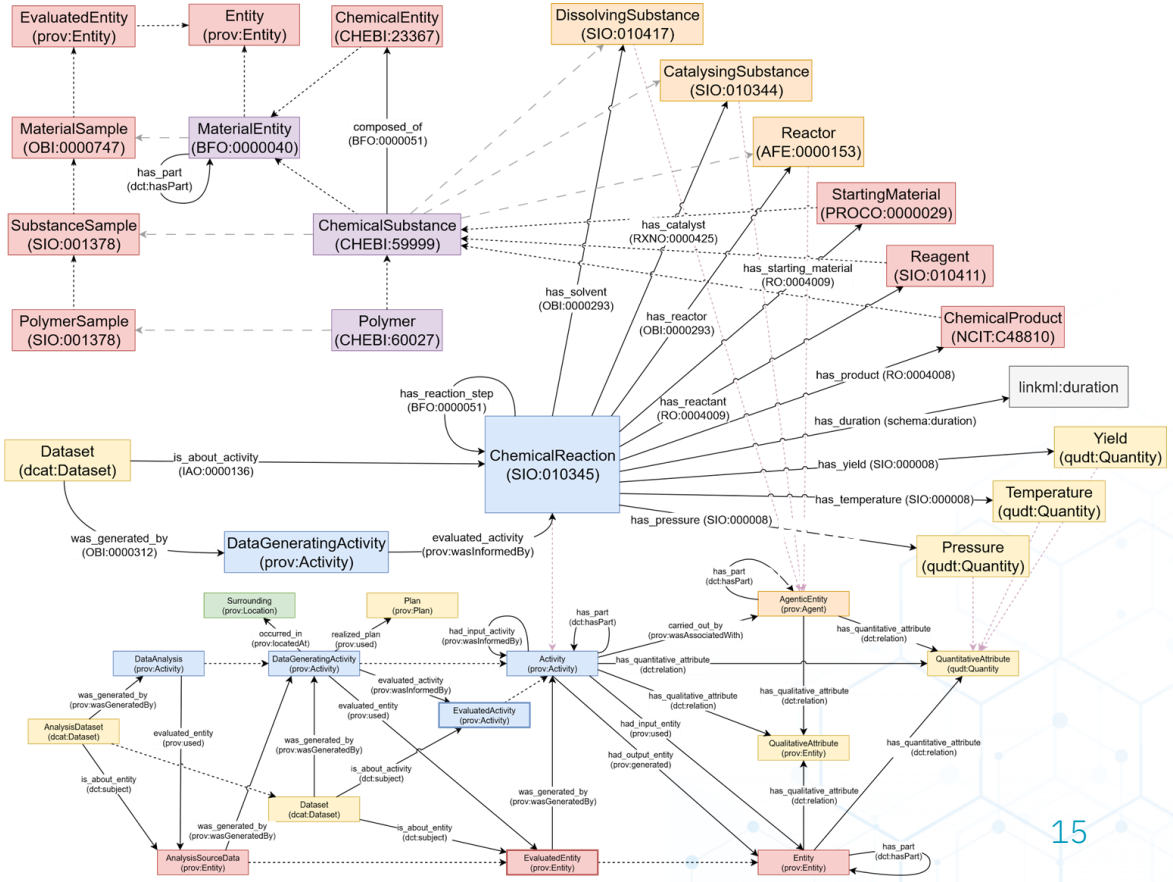
- DCAT-AP Plus Links to Use-case Specific context
  - Added classes, properties & constraints to specify
    - The entity or process being evaluated & described
    - The process(es) & tools used to generate the data
    - Entity, process & attribute types via terminologies
- ChemDCAT-AP

  DCAT-AP PLUS & restrictions shared in all chemistry specific contexts
- MIChI schemas
  - ChemDCAT-AP & specific MIChI requirements
  - Developed by domain experts & MIChI admins



ChemDCAT-AP

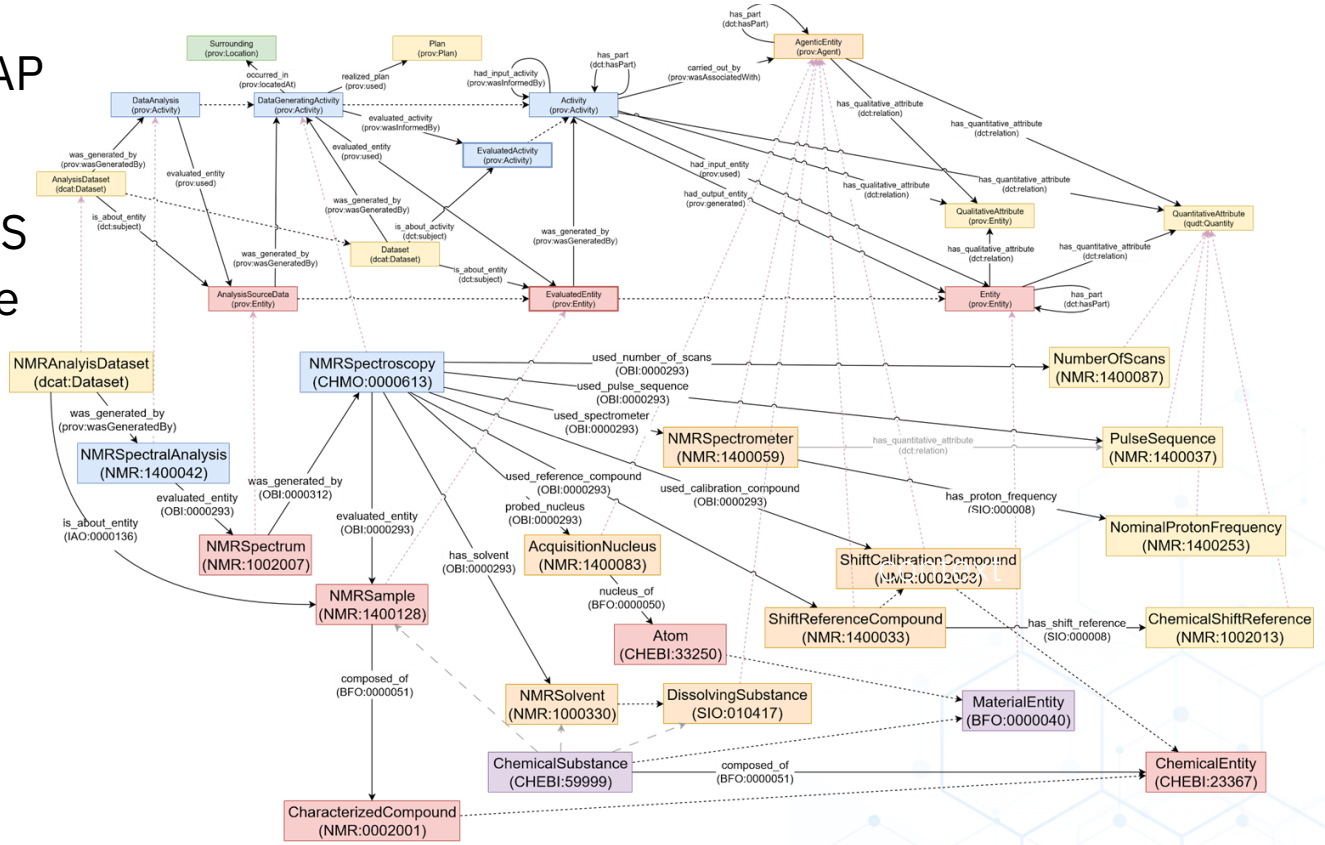
- Imports DCAT-AP PLUS
- Adds requirements shared by NFDI4Chem & NFDI4Cat
  - ChemicalEntity, Atom, Substance, Sample, Reaction, Reagent, Solvent, Catalyst, Yield, Pressure, ...
- Mapped to chemistry specific ontologies



15

MIChI formalized with ChemDCAT-AP

- Imports ChemDCAT-AP
- MIChI profiles must use the DCAT-AP PLUS patterns as a template
  - But can be mapped to more relevant BFO based ontologies
- A MVP formalization of the MARGARITAS (NMR MIChI)
  - Only the required elements formalized



MIChI & Metadata & Ontologies

Minimum Information for Chemical Investigations collect ***what*** to report

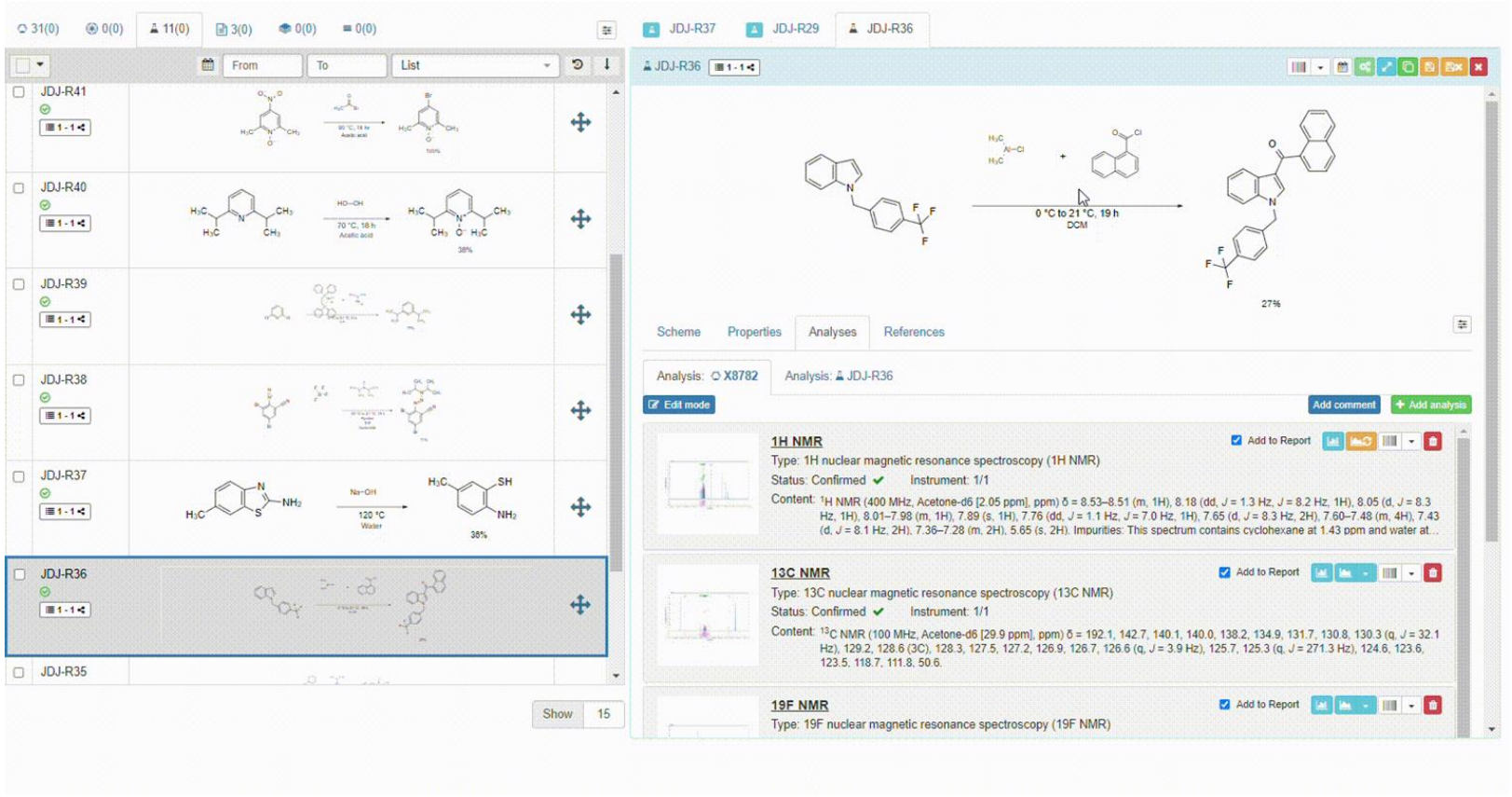
Metadata Schemas define ***how*** to report data.

Schemas and ontologies form the ***semantic foundation*** and enable harmonisation, metadata harvest & search

TERMINOLOGY SERVICE. NFDI4Chem

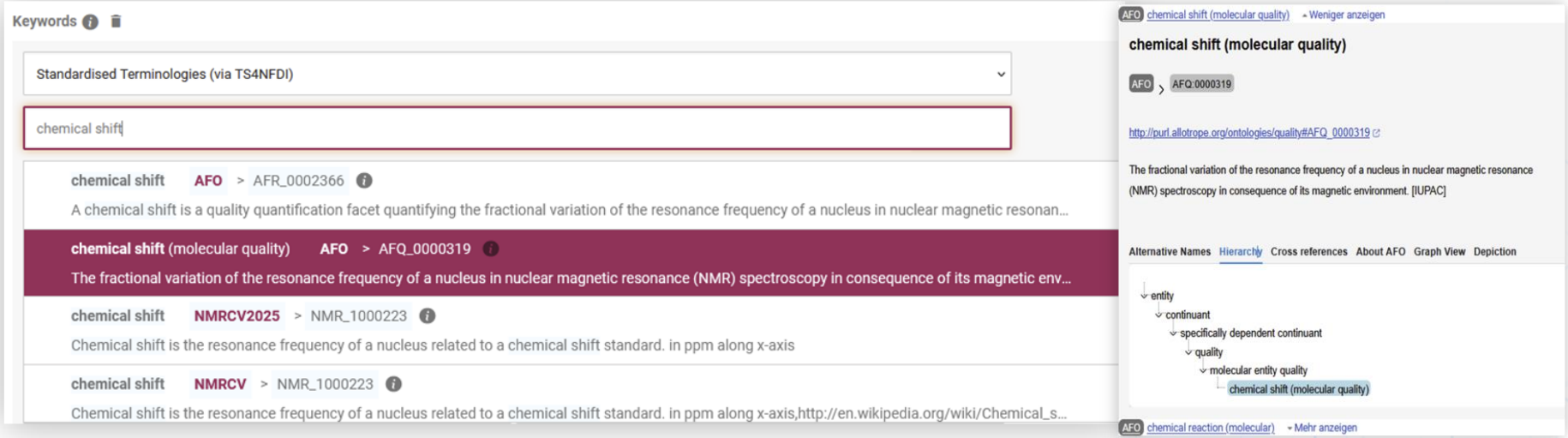
SEARCH SERVICE. NFDI4Chem

Chemotion ELN: Integration of ontologies from TS

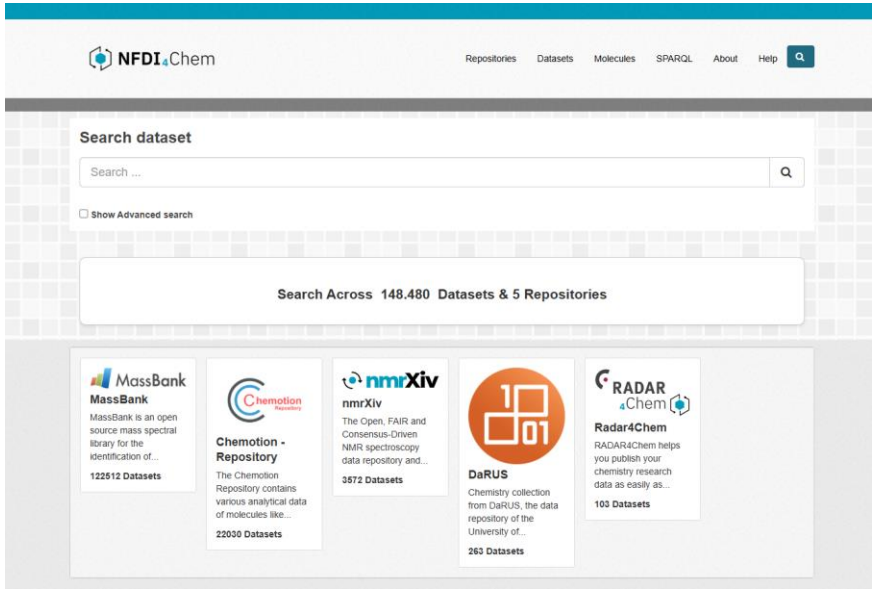


RADAR4Chem: Integration of Ontologies from TS

- Metadata editor: Suggestion lists of chemistry specific terms (e. g. NFDI4Chem ontology collection) are provided via type ahead.
- Landing page: Clear presentation of standardized terms / ontologies incl. links to further information.
- Findability of data: Additional filter options “Keywords” and “Ontology”on publication list.
- MD for “Keywords”: Fosters semantic interoperability using chemistry-relevant vocabularies.



Data Discovery with Search Service

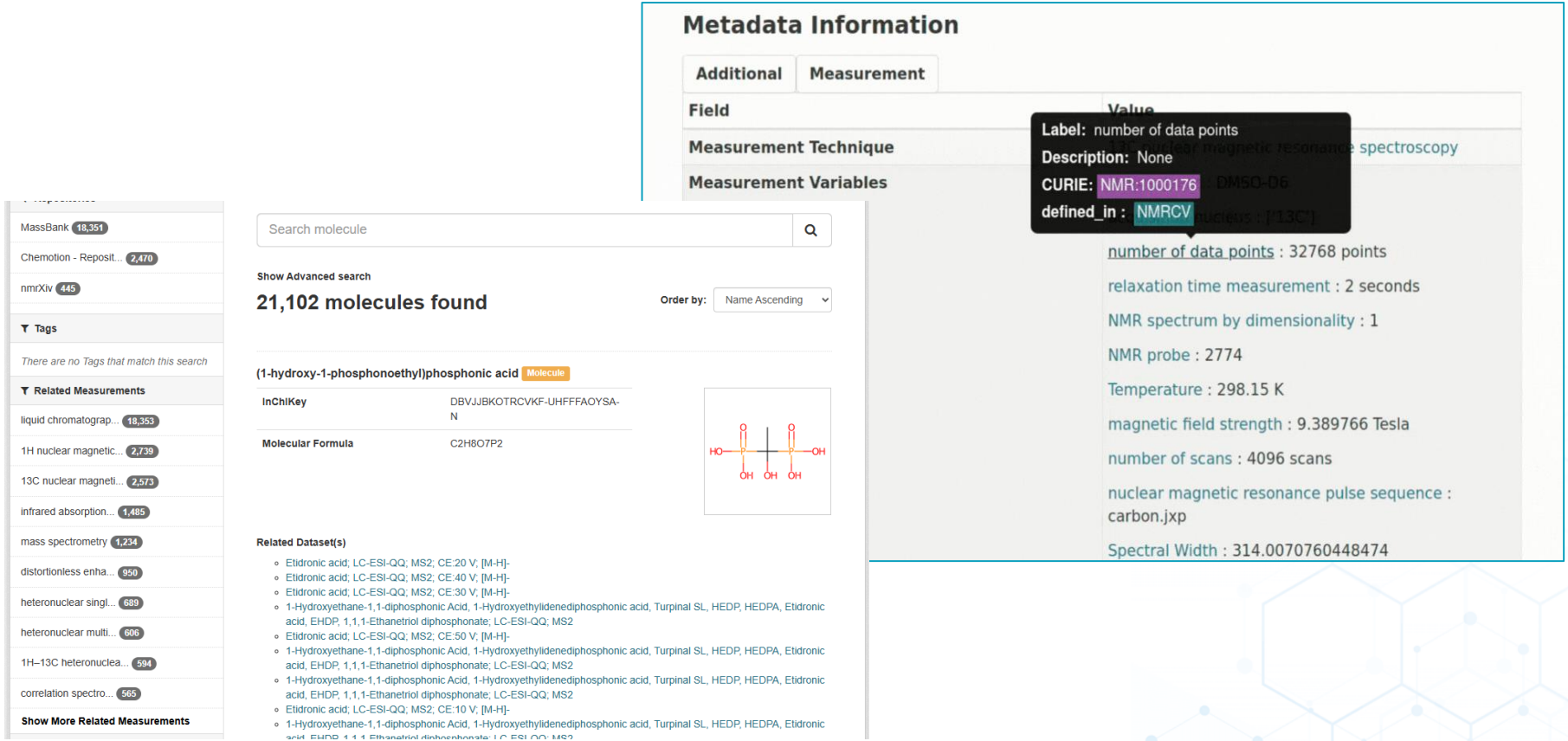


- The NFDI4Chem Search Service harvest major repositories of NFDI4Chem Federation
- Currently includes **>148.000 datasets**
- Enables searches **by molecules, analytical methods or chemical structure codes**
- Features are continuously adapted to latest **metadata schema** releases

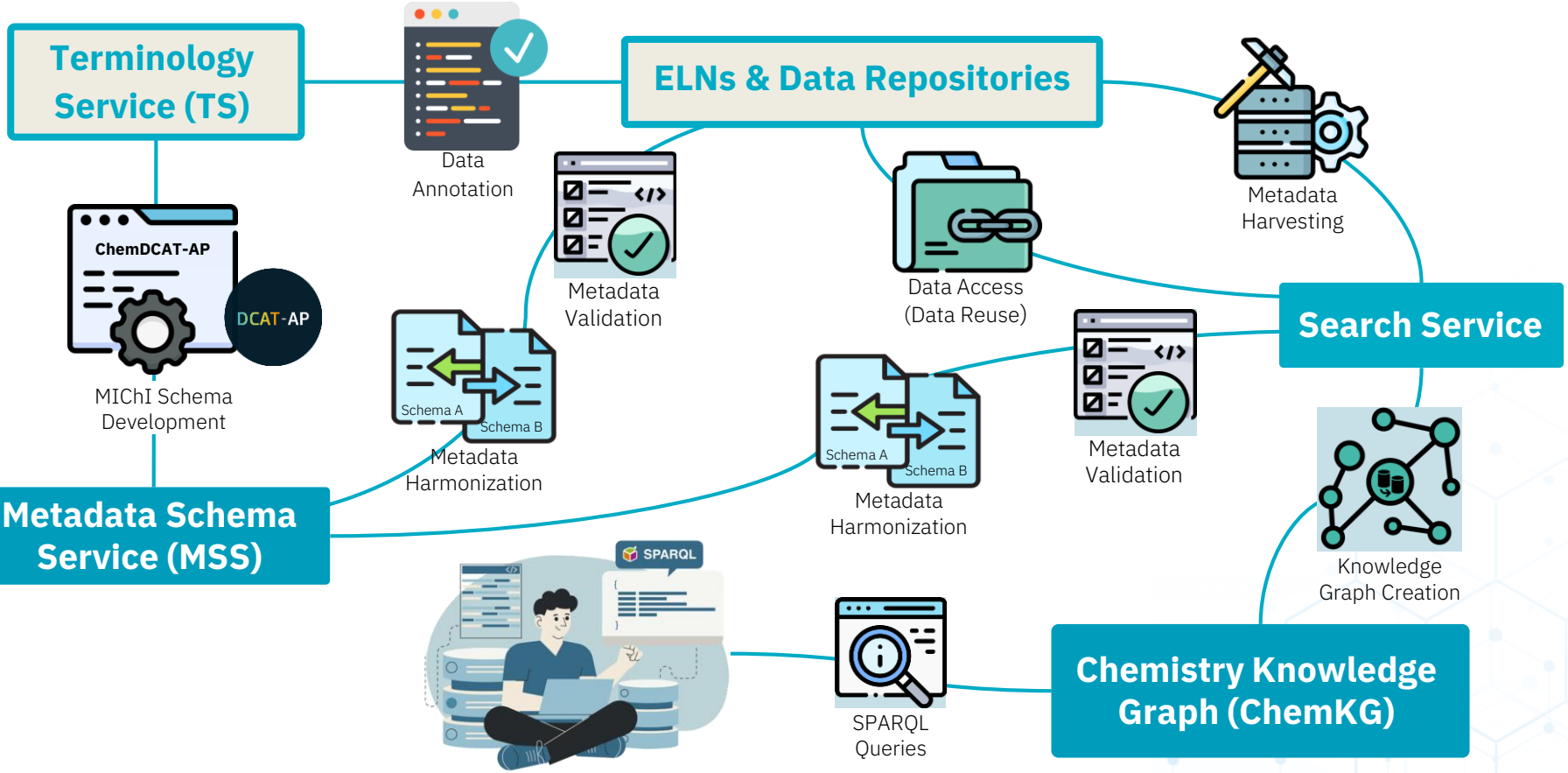
NFDI4Chem: NFDI4Chem Search Service Available: https://search.nfdi4chem.de/

148k Datasets

Semantic Foundation in Search Service



Next Steps - Semantic Data Hub



NFDI4Chem

ENHANCE
YOUR
DATA.