

NFDI4Chem: Shaping Digital and Cultural Change in Chemistry



Christoph Steinbeck*

Friedrich-Schiller-University Jena

Nicole Jung

Karlsruhe Institute of Technology (KIT)

Matthias Razum

FIZ Karlsruhe - Leibniz Institute for Information
Infrastructure

Johannes Liermann

Johannes Gutenberg University Mainz

Beilstein Open Science Symposium 2021

Oliver Koepler*

TIB Leibniz Information Centre of Science and
Technology

Sonja Herres-Pawlis & Alexander Hoffmann

RWTH Aachen University

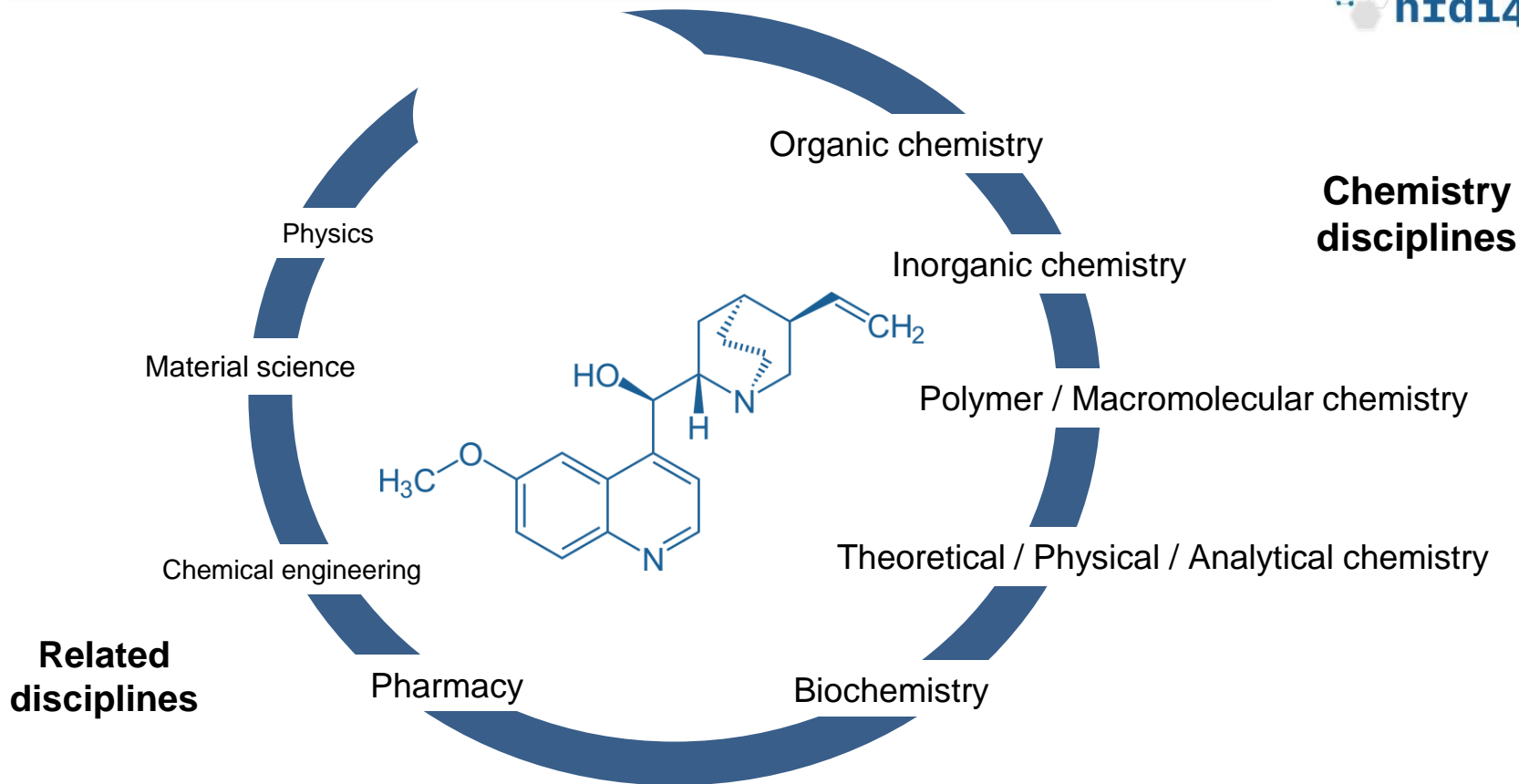
Felix Bach

Karlsruhe Institute of Technology (KIT)

Steffen Neumann

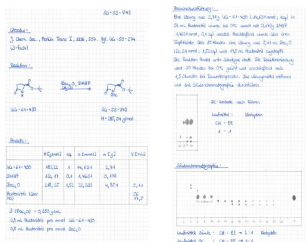
Institut für Pflanzenbiochemie, Halle

Our scientific community



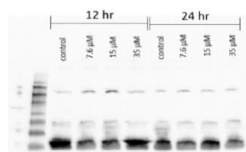
Molecules and related (meta)data

Reactions



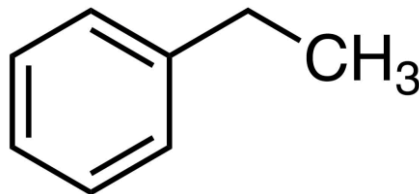
Experimental Description
Starting compounds /
Reagents
Conditions ...

Biological activity



Organism
Conditions
Activity
...

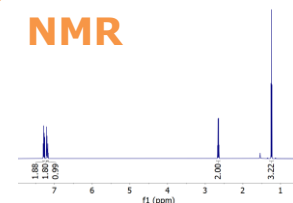
Properties



Name Ethylbenzene
Formula C₈H₁₀
CAS 100-41-4
InChI Key YNQLUTRBYVCPMQ-UHFFFAOYSA-N
...

Assignment

NMR



Nucleus
Pulse sequence
Solvent ...

MS



Ionisation method
Voltage ...

Spectra

Status Quo

Analogue / Digital

10-11-194

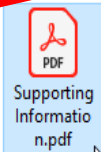
Abstract: ...

10-11-194

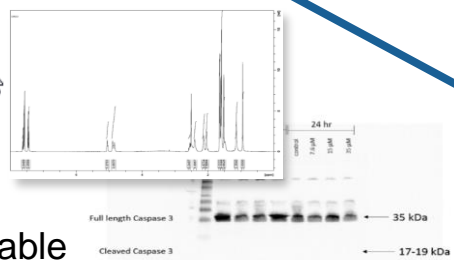
10-11-194

Parameter	Value	Unit	Value	Unit
10-11-194	48.12	A	14.514	2.34
10-11-194	42.14	0.4	4.624	0.19
10-11-194	14.12	0.1	12.124	1.34

Supporting Information n.pdf



Supporting Information n.pdf



Digital



UnFAIR

Not machine-readable

Forces and Needs



Leitlinien zur Sicherung
guter wissenschaftlicher Praxis

Kodex

DFG

Practical questions in the organisation of a working group:



Where are the NMR data from 10 years ago?



Where are the synthesis details from 15 years ago?



Has anyone in the group already made compound xy?

What do the colleagues say....



Critical questions from colleagues:

- ? How much time does it cost my students?
- ? How much money does it cost me?
- ? Do I really have to do this?

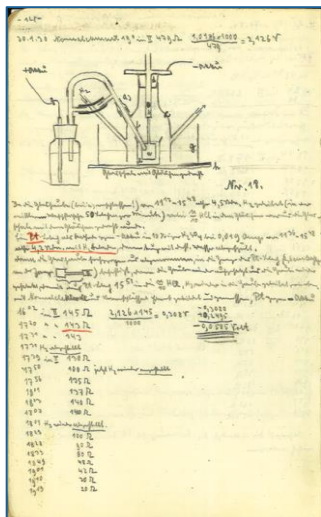
More critical statements from colleagues:

- The supporting information is sufficient.
- No one else does this...
- It is a better time invest when my students are not sitting too much at the computer.

Our Vision



1927



Digitalisation

Standards

Community



FAIR DATA

Experiment & Data collection

Data processing

Data Analysis

Data preserved

Data re-use

Data publication

Key objectives



Objective 1: Connect existing **data repositories**, fill in missing research data repositories, and link them to international repositories.



Objective 2: **Minimum information (MI) standards for data and machine-readable metadata**, open data standards, in order to support the FAIR principles for research data.



Objective 3: Foster **Electronic Laboratory Notebooks (ELN)**, tools and APIs between between instrumentation and software towards a embedded, digital information architecture. Capture research data in well-annotated electronic form at the **earliest possible point in time in the research process.**



Key objectives



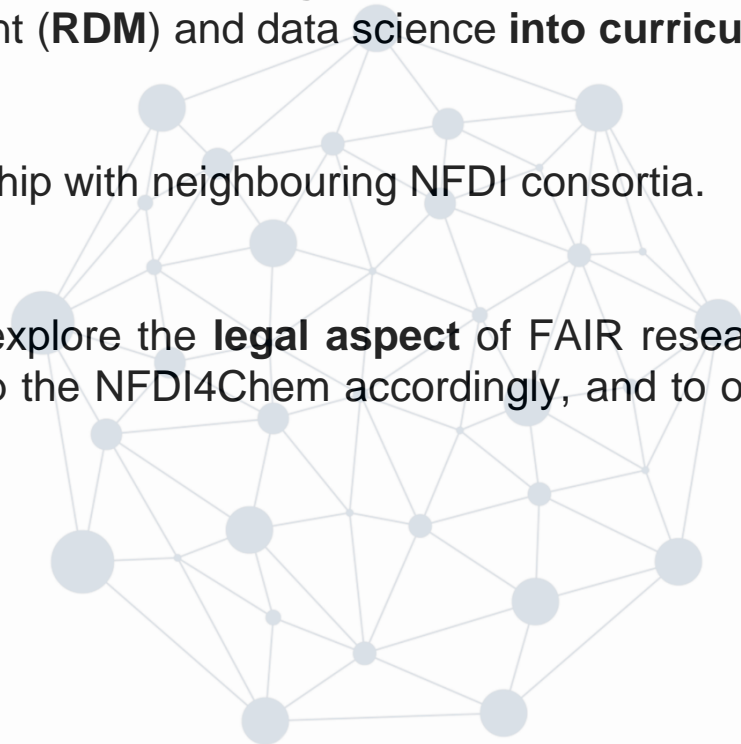
Objective 4: Create **awareness** for **FAIR data management**, initiate processes to **integrate** research data management (**RDM**) and data science **into curricula**.



Objective 5: Maintain a close relationship with neighbouring NFDI consortia.



Objective 6: Engage with experts to explore the **legal aspect** of FAIR research data management, design and develop the NFDI4Chem accordingly, and to offer advice for the research community.



Involvement of the Community

Community requirements

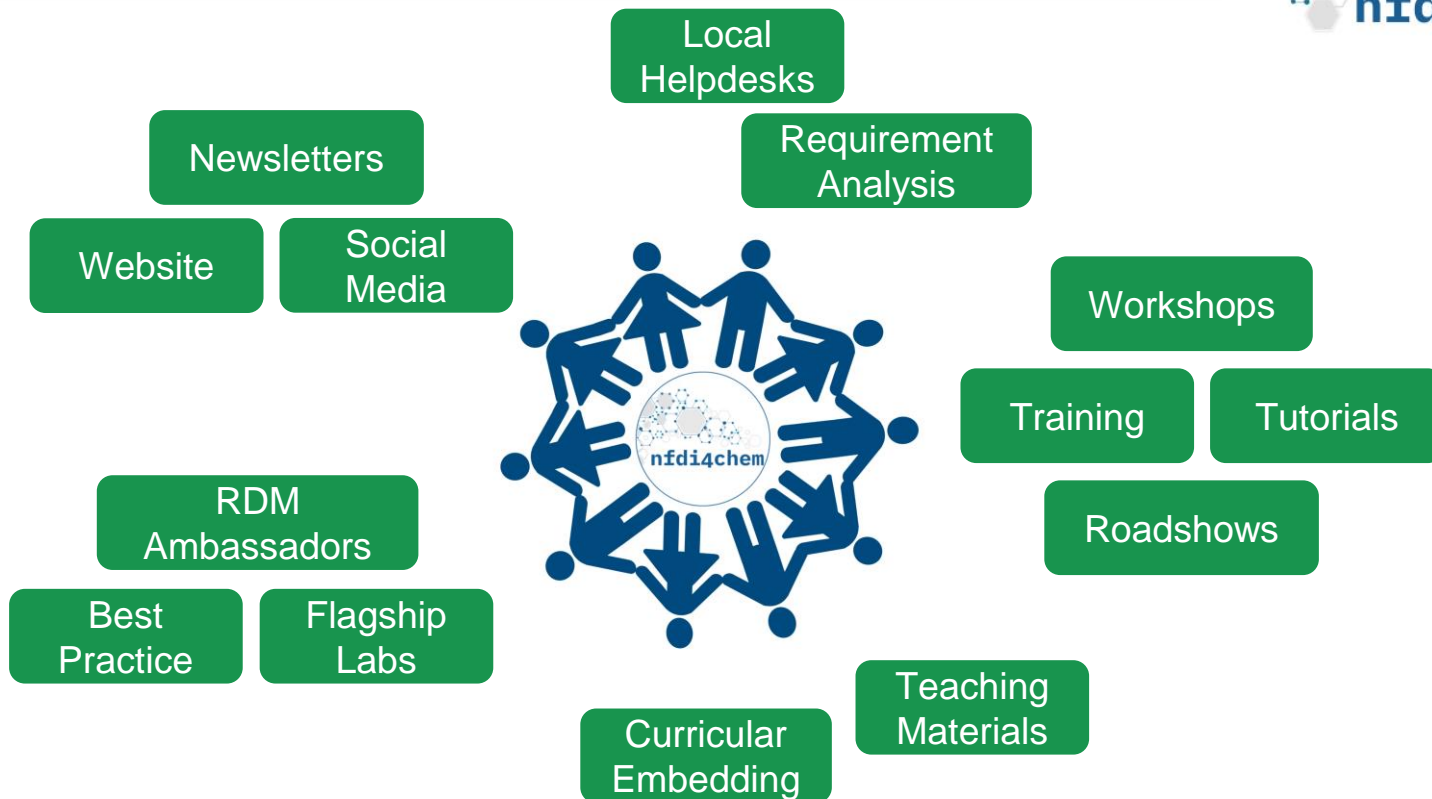


Fostering cultural change

Raising RDM awareness

RDM
Infrastructure

Involvement of the Community



Best practice

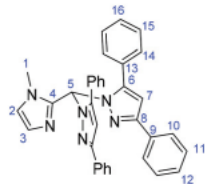
Dalton Transactions

PAPER



Cite this: DOI: 10.1039/d1dt00832c

<https://www.chemotion-repository.net/welcome>



^1H NMR (CD_2Cl_2 , 400 MHz): δ = 7.80 (d, 3J = 7.8 Hz, 4H, H-10), 7.67 (s, 1H, H-5), 7.39 (t, 3J = 7.4 Hz, 4H, H-11), 7.32 (d, 3J = 7.3 Hz, 4H, H-12 + H-16), 7.27 (t, 3J = 7.3 Hz, 4H, H-15), 7.12 (d, 3J = 7.1 Hz, 4H, H-14), 6.94 (d, 3J = 3.1 Hz, 2H, H-2 + H-3), 6.65 (s, 2H, H-7), 3.70 (s, 3H, H-1) ppm. ^{13}C $\{^1\text{H}\}$ NMR (CD_2Cl_2 , 100 MHz): δ = 151.8 (C-8), 146.4 (C-6), 142.5 (C-4), 133.7 (C-9), 129.5 (C-13), 129.4 (C-14), 129.1 (C-15), 129.1 (C-12 + C-16), 128.5 (C-11), 128.0 (C-3), 126.3 (C-10), 124.3 (C-2), 106.3 (C-7), 70.0 (C-5), 34.9 (C-1) ppm. HRMS (ESI $^+$, MeOH): m/z (found) = 533.24353 (100%), 534.24664 (37%), 535.25024 (7%), 536.25299 (1%); m/z (calc.) = 533.24537 (100%, $^{12}\text{C}_{35}^{1}\text{H}_{29}^{14}\text{N}_6^+$), 534.24872 (38%, $^{12}\text{C}_{34}^{13}\text{C}^1\text{H}_{29}^{14}\text{N}_6^+$), 535.25207 (7%, $^{12}\text{C}_{33}^{13}\text{C}_2^1\text{H}_{29}^{14}\text{N}_6^+$), 536.25543 (1%, $^{12}\text{C}_{32}^{13}\text{C}_3^1\text{H}_{29}^{14}\text{N}_6^+$). IR (ATR, neat), $\tilde{\nu}$ [cm^{-1}]: 1605 (vw), 1551 (w), 1459 (w), 1437 (w), 1410 (vw), 1298 (vw), 1281 (vw), 1259 (w), 1202 (w), 1138 (vw), 1076 (w), 1027 (w), 1007 (w), 957 (w), 916 (w), 868 (w), 841 (w), 834 (w), 816 (m), 803 (w), 767 (m), 759 (vs), 752 (s), 704 (m), 691 (vs), 678 (m), 666 (m), 573 (m), 524 (w), 432 (vw).

Additional information on the NMR of the target compound including original data files is available via Chemotion Repository: <https://dx.doi.org/10.14272/LSGGPBVWWQPOY-UHFFFAOYSA-N.1>



ROYAL SOCIETY OF CHEMISTRY

View Article Online

View Journal

Copper nitrenes and amination

ai C. Göbgen,
mann and

as intermediates in the copper catalyzed amination of alkenes, copper complexes were synthesized at low temperatures from the (pyrazolyl)methane family. The copper complex with $^5\text{PhINTs}$ in dichloromethane.

Chemotion www.chemotion-repository.net/mydb/collection/2171/sample/49359

Repository - All ILPAC, InChI, SMILES, RIn

Chemotion 015(0) 40(0)

Scheme-only reactions

My Published Elements

Pending Publications

Embargoed Publications

My Collections

All

ELN Gate

My Data

FT-17 C28H28N6

FT-22 C26H27Cu2F12N10P2

FT-15 C12H11N5
2-[di(pyrazol-1-yl)methyl]pyridine

FT-7 C38H28N8
1-[(3,5-diphenylpyrazol-1-yl)-(1-methylimidazol-2-yl)methyl]-3,5-diphenylpyrazole

FT-6 C38H28ClN8
2-[bis(3,5-diphenylpyrazol-1-yl)methyl]-4-chloropyridine

FT-5 C28H22CuF6N6P
acetonitrile,2-[bis(3-phenylpyrazol-1-yl)methyl]pyridine,copper(1+),hexafluorophosphate

C27H24Cu2F12N12P2
copper(1+),1-(3,5-diphenylpyrazol-1-yl)-(1-methylimidazol-2-yl)methyl-3,5-diphenylpyrazole

FT-5 C28H22CuF6N6P
acetonitrile,2-[bis(3-phenylpyrazol-1-yl)methyl]pyridine,copper(1+),hexafluorophosphate
627.003260 g/mol
Exact mass: 626.064373 g/mol

Properties Analyses QC & curation Literature Results

Molecule Stereo Abs Stereo Rel
 any any Top secret

Name External label Solvent

Amount Boiling point
 mg mg mg °C

Density Molarity Purity Melting point
 g/ml g/ml °C

Description

Location

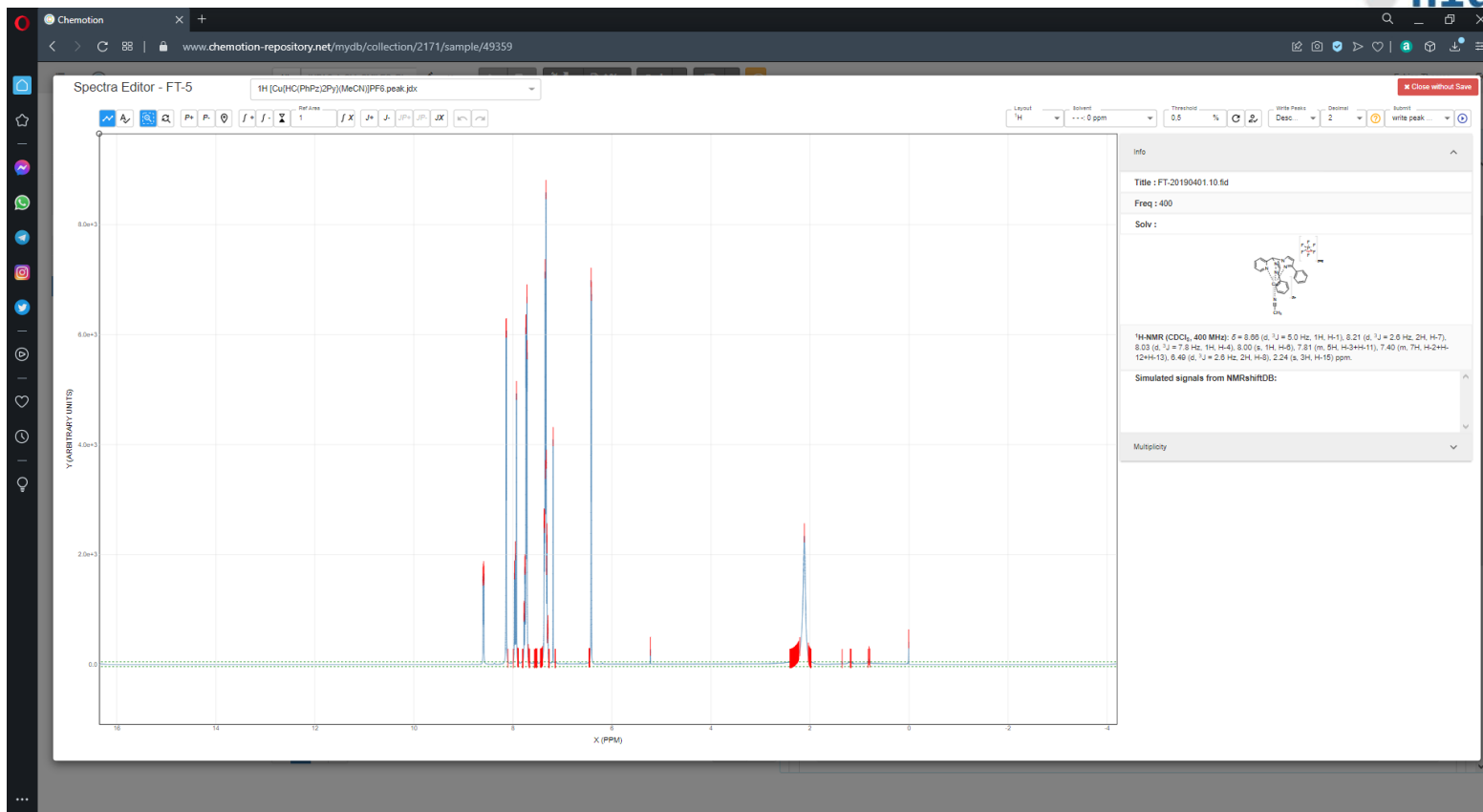
Elemental composition >

Chemical identifiers >

Clear Save Download Analysis

Show 15

Chemotion Repo



Integration into Curricular Teaching



Curricular recommendations of GDCh just came out!

→ Anchoring data literacy and research data management modules already in the B.Sc. studies

BUT: long march through the institutions...



Empfehlungen
der GDCh-Studienkommission

zum
Bachelorstudium Chemie
an Universitäten

Integration into Curricular Teaching



More direct and faster way by hidden integration:

Example 1: in a Master lecture

Master lecture (2h per week) in the free area of the Master studies @RWTH

- Sustainable coordinative polymerisation catalysis
- 80-100 students, 50 take the exam
- Explaining chemistry with case studies
- Dissecting the RDM of the case studies (good and bad examples)
- Integrating videos on the basics of RDM from RWTH library/NFDI4Chem

Integration into Curricular Teaching



More direct and faster way by hidden integration:

Example 2: lab stage in advanced inorganic chemistry

- 5th semester in bachelor studies @RWTH
- ~ 130 students
- Introduction of the Electronic Laboratory Notebook Chemotion
- Integrating videos on the basics of research data management, FAIR principles, data management plan, metadata and InChI and SMILES

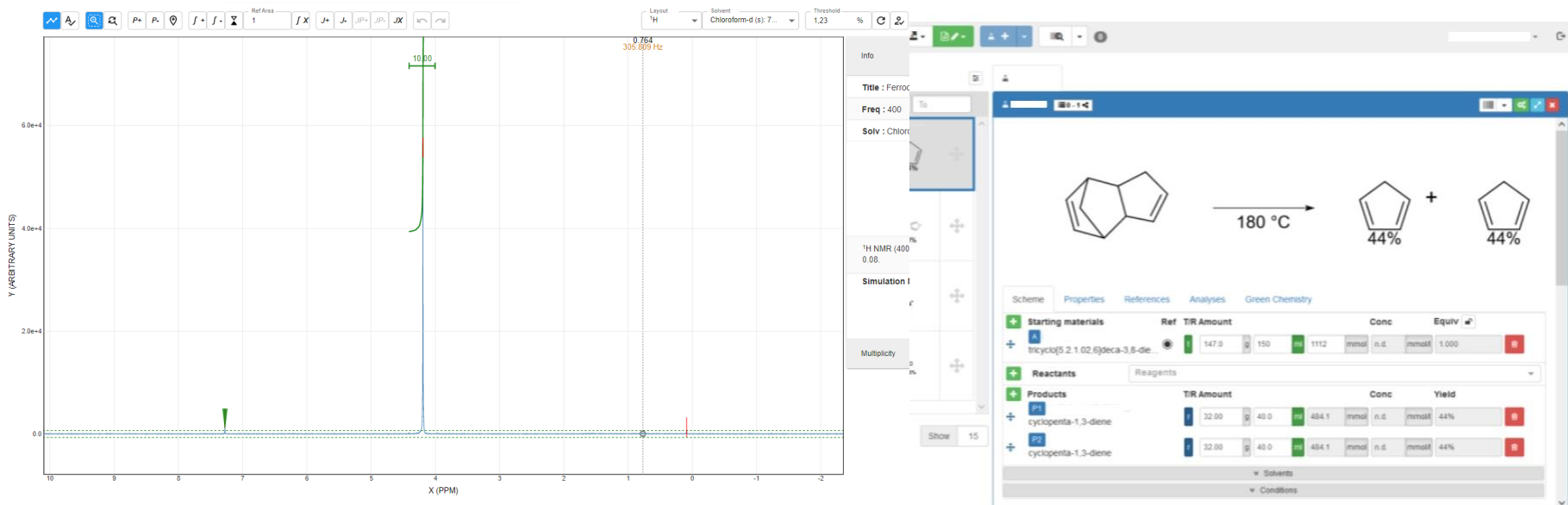


Integration into Curricular Teaching



More direct and faster way by hidden integration:

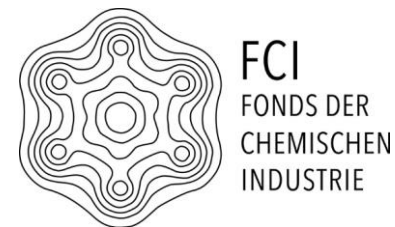
Example 2: lab stage in advanced inorganic chemistry



„FAIRThesis“



- The teaching project "FAIRThesis" is supported by the Fonds der Chemischen Industrie (FCI).
- Focus on the digitization of research data obtained in research internships, Bachelor and Master theses and the processing of FAIR research data (Findable, Accessible, Interoperable and Reusable) in the electronic laboratory notebook Chemotion.
- Installation of Chemotion on a server available via VPN of the RWTH Aachen University
- In addition, students will be provided with in-depth knowledge of research data management and electronic laboratory notebook. We are supported by Dr. Nicole Jung of the Karlsruhe Institute of Technology in the implementation of the project.
- Available starting in January 2022, Contact @RWTH: Dr. Alexander Hoffmann



FAIR4Chem Award



What's it about?

Award is given for published chemistry research datasets that best meet the [FAIR principles](#).

How are datasets evaluated?

Available FAIR assessment tools and a jury will evaluate the best dataset.



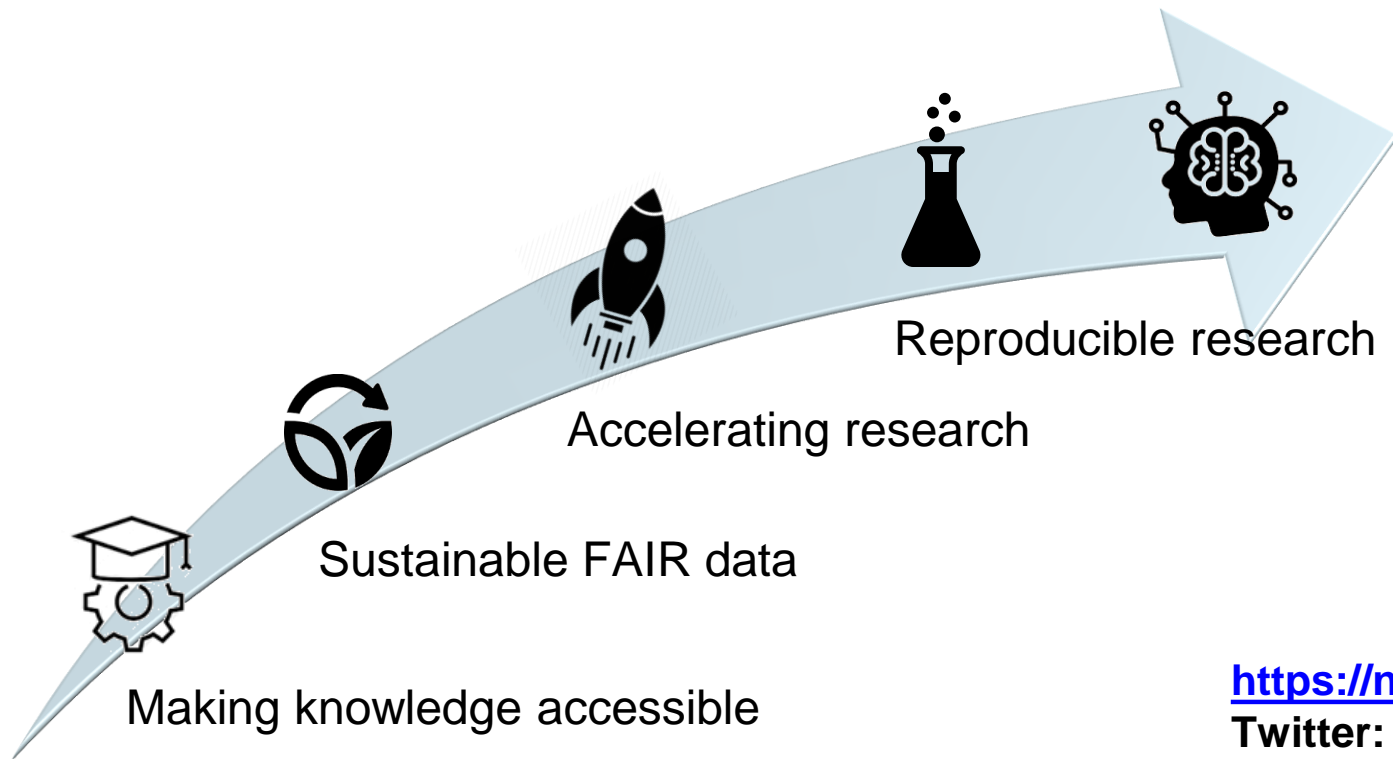
What's the prize?

Prize money of 500 € (financed by FCI) and talk about winning dataset at JCF Frühjahrssymposium 2022.

How to participate?

Submit the link to the dataset between October 15th and December 15th 2021 via [FAIR4Chem Award webpage](#).

A Quantum Leap in Chemistry RDM



Enabling
machine-learning
on the next level

<https://nfdi4chem.de/>

Twitter: @NFDI4chem

LinkedIn: NFDI4Chem