Karlsruhe Institute of Technology

Göttingen eResearch Toolbox Series - Electronic Note Keeping

Nicole Jung

INSTITUTE OF ORGANIC CHEMISTRY - Stefan Bräse Group Karlsruhe

Chemotion
funded by
DFG

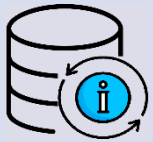
Access to research data: Key challenges

Digital Data

Almost no digital data

- Only written lab journals
- Not searchable
- Analytics ›› syntheses
- Processing

Databases



Commercial databases

- Limited access
- Limited availability of data

Open Access databases

Only few data

Journals

Limited content

- No primary data
- Almost no additional info
- Accessibility

Academia-specific challenges:

Strategic issues:

- Sustainable, interdisciplinary research
- project management vs community network
- Collecting data as a basic instrument for research
- Acceptance in the community

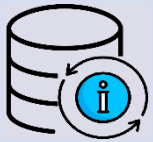
Access to research data: Key challenges

Digital Data

Almost no digital data

- Only written lab journals
- Not searchable
- Analytics ›› synthesis
- Processing

Databases



Commercial databases

- Limited access
- Limited availability of data

Open Access databases

Only few data

Journals

Limited content

- No primary data
- Almost no additional info
- Accessibility

Academia-specific challenges:

Strategic issues:

- Sustainable, interdisciplinary research
- project management vs community network
- Collecting data as a basic instrument for research
- Acceptance in the community

Technical issues:

- Infrastructure and processes differ
- Equipment and devices in general very old
- High flexibility necessary

Access to research data: Key challenges

ELN

- Allows digital storage
- Processing of structures

Repositories

- Free access
- No limitation (file types)
- Linking journal contents to additional data

Device Integration

- Dial-a-device (remote platform)
- API via external IoT (cubuslab server)
- ELN-data management

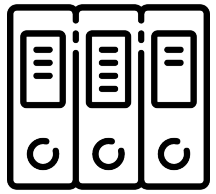
Key instrument: ELN implementation I

Electronic Lab Notebook (ELN)

Generic Functions
Adaptable to other fields of research
Management and systematic structure

Management and Organization

- Management for projects and subprojects (flexible structure)
- Organization according to lists and detail tabs
- User defined settings, BarCodes and QR-Codes
- Items are linked, ancestries are summarized and visualized

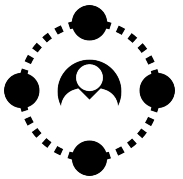


Reporting

- Reports for single elements and projects
- User defined detail level
- Settings for formatting

Share & Synchronize

- Share a selected group of information
- Synchronize to allow access to projects, support of collaboration
- User defined detail level
- Definition of groups for sharing and synchronization

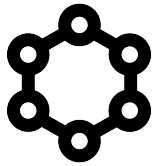


Key instrument: ELN implementation II

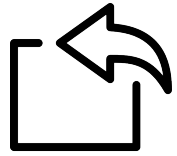
Electronic Lab Notebook (ELN)

Chemistry Specific Functions
Application in molecule based research
Processing of chemical data

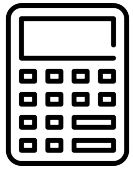
Structure Editor (Ketcher)
- Standard functions + support of polymers/surfaces
- User defined templates and general accepted templates
- Information function for structures and substructures
- Implementation of abbreviations



Export and Import functions
- Import and export of Excel and sdf
- User defined export scheme (detail level)



Calculations
- Calculations for elemental analysis (e.g. polymer supported material)
- Calculations of yield, control for „impossible values“



Search functions
- Text search and structure (substructure/similarity) search

Key instrument: ELN implementation II

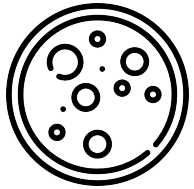
Electronic Lab Notebook (ELN)

Chemistry Specific Functions
Application in molecule based research
Processing of chemical data

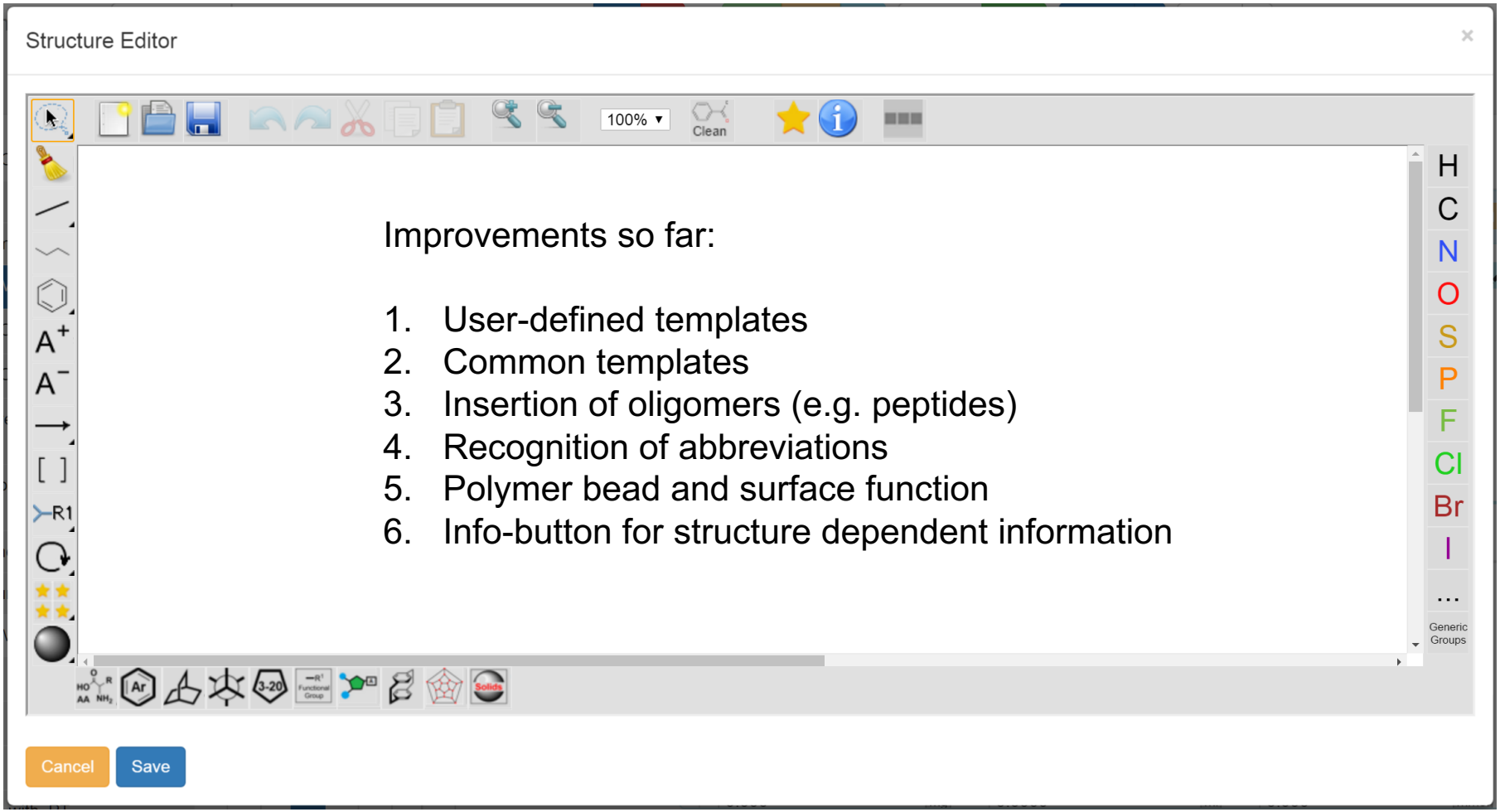
Additional Information: Database integration
- OpenBabel
- PubChem, SciFinder
- NMRbd
- Direct visualization in lists

Datasets and File attachment
- Description of metadata (dropdown and freetext)
- Labeling (approved, not approved)
- All file formats, no limitation

Functions for biology and chemical biology
- Wellplate function for sample management
- Organizer for biological assays and their description
- Correlation of molecules, wellplates and assays



Ketcher molecule editor



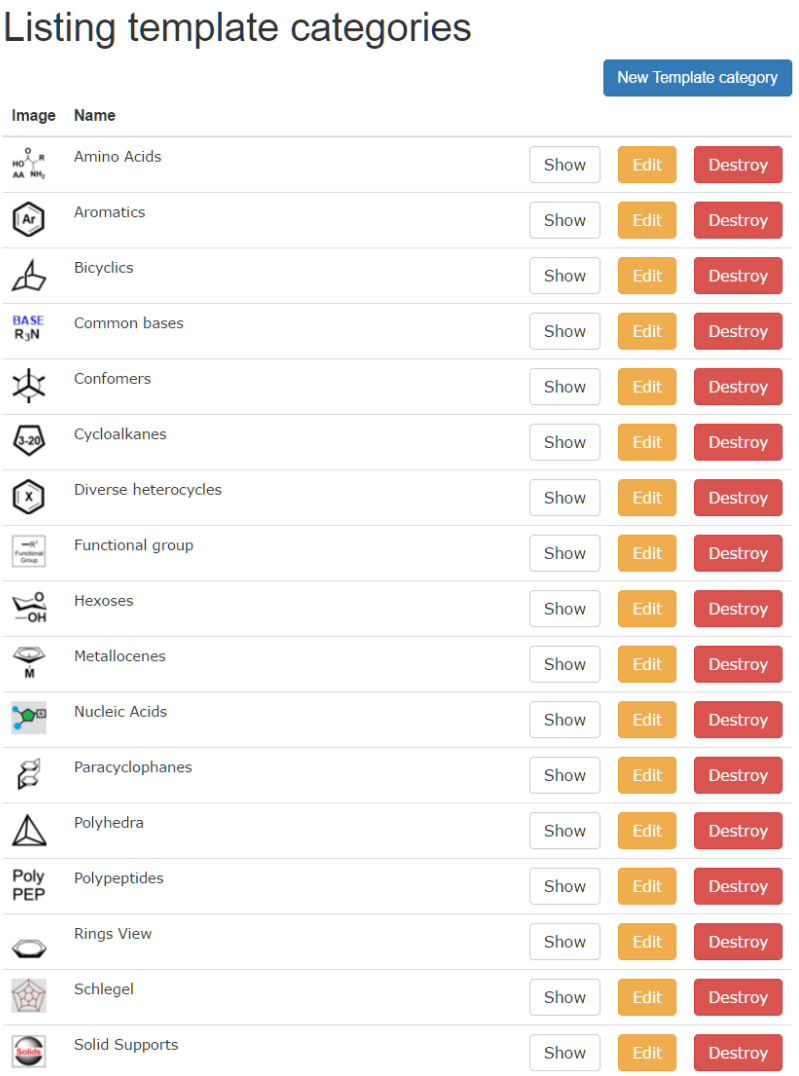
Ketcher molecule editor

Template categories
- Image selection
- Define and approve or reject

Templates
- Image selection
- Import SDF or Molfile

Abbreviations
- Name definition
- Define attachment point
- Use of molfile for database requests

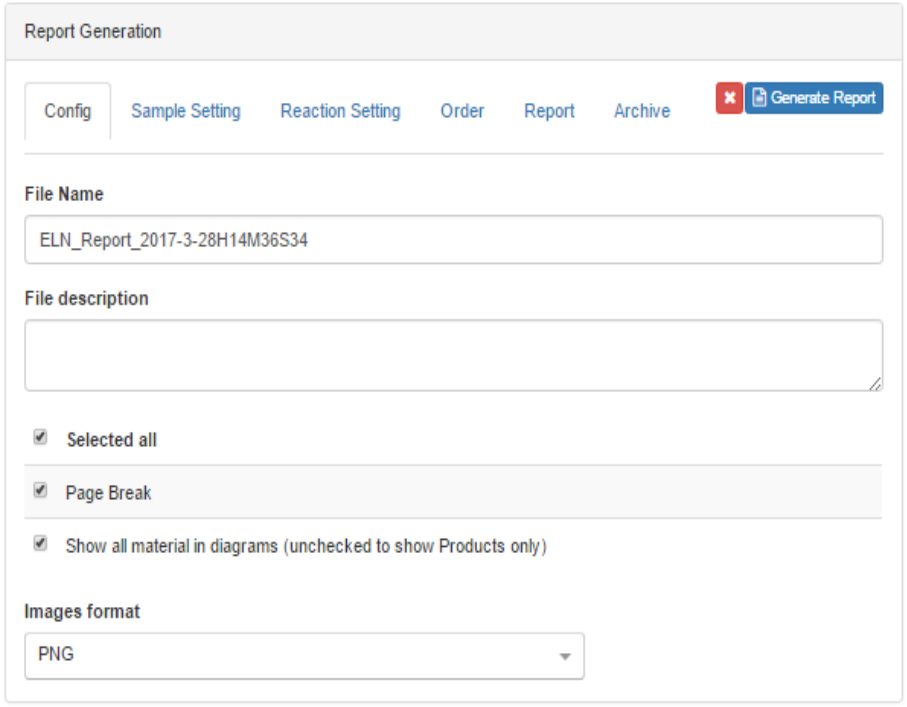
Monomer generation for oligomers
- Definition of name and abbreviation
- Definition of two attachment points
- Use of molfile for database requests
- Reverse order of monomers

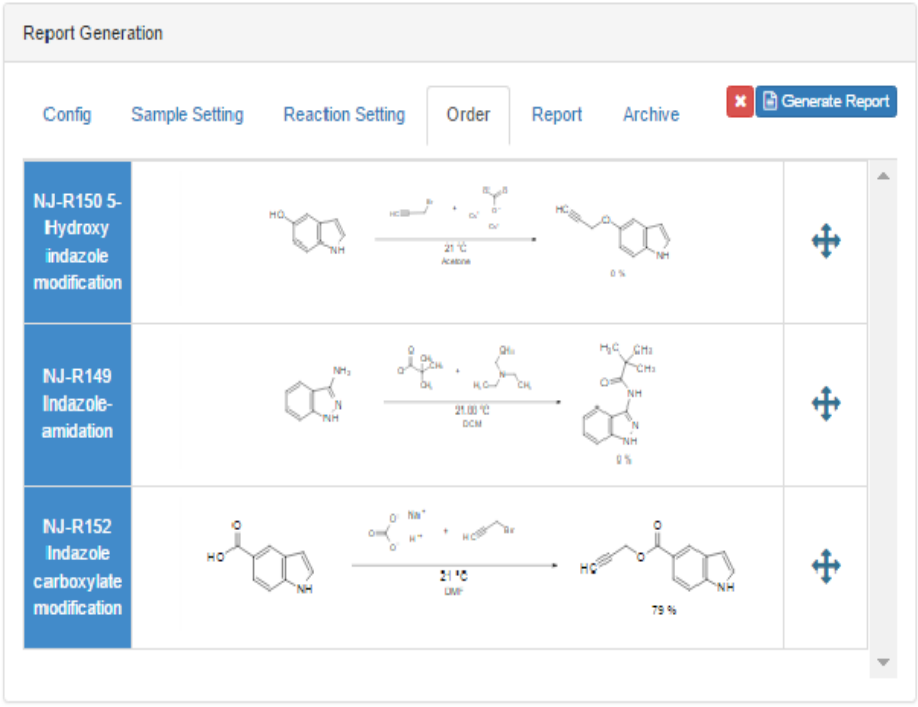


Reporting

Reporting function in brief:

- Project reports
- Supporting information for publications
- Exchange of information





Access to research data: Key challenges

ELN

- Allows digital storage
- Processing of structures

Repositories

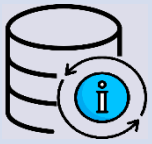
- Free access
- No limitation (file types)
- Linking journal contents to additional data

Device Integration

- Dial-a-device (remote platform)
- API via external IoT (cubuslab server)
- ELN-data management

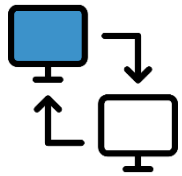
Disclosure of information: Repository

Repository



Basic Functions
Technology identical with ELN
Generic and field specific functions identical

Input and Transfer of Data
- Selection of data within ELN, determination of detail level
- Transfer of data via shared folder repository-ELN
- Visibility of data after additional confirmation
- Data for molecules and reactions



Characteristics
- Search functions based on FP2-fingerprints
- Metadata defined by ELN definitions
- DOI-generation, PubChem listing, Datacite indexing

Visibility and Advertisement
- Infrastructure for repositories
- Competitions with incentives (by KIT-Stiftung)

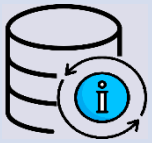
RIsources The Research Infrastructure Portal funded by DFG

re3data.org REGISTRY OF RESEARCH DATA REPOSITORIES

Disclosure of information: Repository

Electronic Lab Notebook (ELN)

Repository



External

Selection
Datasets to publish

Private Repository
Virtual DOI from metadata (molecule-dataset)

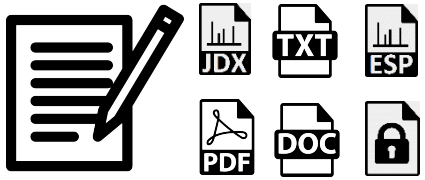
Open Repository
DOI – dataset/molecule url (TIB Hannover)

DataCite
Mint and register DOI

PubChem
Structure + DOI

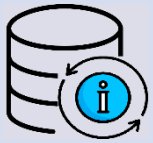
DOI-resolver

Molecule - Datasets



Disclosure of information: Repository

Repository

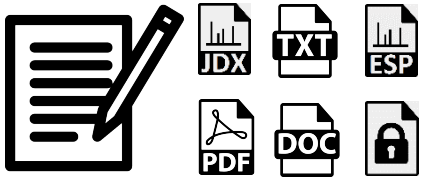


Private Repository
Virtual DOI from metadata (molecule-dataset)

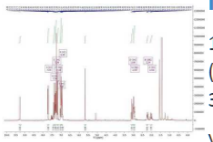
Open Repository
DOI – dataset/molecule url (TIB Hannover)

DOI-resolver

Molecule - Datasets



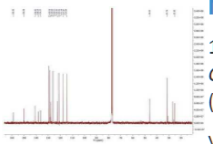
5 Datasets Add Dataset



NMR/1H/CDCl3/400

1H NMR ($CDCl_3$, 400 MHz), δ = 9.28 (s, 1 H, N*H*), 7.75–7.67 (m, 2 H, H_{Ar}), 7.43–7.34 (m, 2 H, H_{Ar}), 7.34–7.25 (m, 3 H, H_{Ar}), 7.23–7.16 (m, 2 H, H_{Ar}), 7.06–6.99 (m, 3 H, H_{Ar}), 6.98–6.90 (m, 2 H, H_{Ar}), 5.67 (s, 1 H, C*H*), 3.14–3.02 (m, 2 H, SCH_2), 3.00–2.92 (m, 2 H, SCH_2), 2.28–2.21 (m, 1 H, CH_2), 2.09–1.94 (m, 1 H, CH_2) ppm.

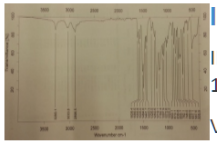
Virtual DOI: 10.14272/PBTUVUQCTONFGS-YYDJUVGSSA-N/NMR/1H/CDCl3/400



NMR/13C/CDCl3/100

13C NMR ($CDCl_3$, 100 MHz) δ = 158.9 (C_{Ar}O), 150.2 (C_{Ar}O), 140.9 (*C*N), 138.1 (C_{Ar}N), 136.3 (C_{Ar}), 129.7 (2 C_{Ar}), 128.5 (2 C_{Ar}), 128.1 (C_{Ar}), 126.1 (2 C_{Ar}), 122.3 (C_{Ar}), 121.2 (2 C_{Ar}), 117.5 (2 C_{Ar}), 114.4 (2 C_{Ar}), 46.0 (*S*CS), 31.7 (2 SCH_2), 25.4 (CH_2) ppm.

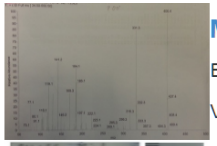
Virtual DOI: 10.14272/PBTUVUQCTONFGS-YYDJUVGSSA-N/NMR/13C/CDCl3/100



IR

IR (ATR) v = 3283, 3034, 2897, 1608, 1585, 1556, 1505, 1485, 1442, 1420, 1327, 1264, 1219, 1162, 1098, 1073, 1022, 900, 870, 837, 806, 782, 749, 690, 657, 588, 506 cm^{-1}.

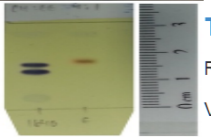
Virtual DOI: 10.14272/PBTUVUQCTONFGS-YYDJUVGSSA-N/IR



Mass/EI

EI MS (70 eV, 200 °C) m/z (%) = 406 [M^+] (100), 331 (86), 184.1 (53), 141 (58), 129 (31), 119 (37), 77 (22).

Virtual DOI: 10.14272/PBTUVUQCTONFGS-YYDJUVGSSA-N/Mass/EI



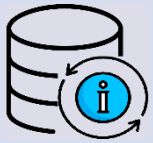
TLC

Rf = 0.48 (cyclohexane : ethyl acetate, 4:1)

Virtual DOI: 10.14272/PBTUVUQCTONFGS-YYDJUVGSSA-N/TLC

Disclosure of information: Repository

Repository

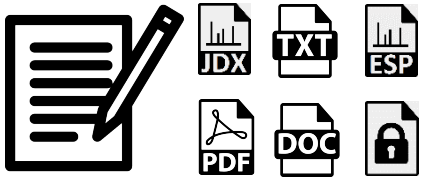


Private Repository
Virtual DOI from metadata (molecule-dataset)

Open Repository
DOI – dataset/molecule url (TIB Hannover)

DOI-resolver

Molecule - Datasets

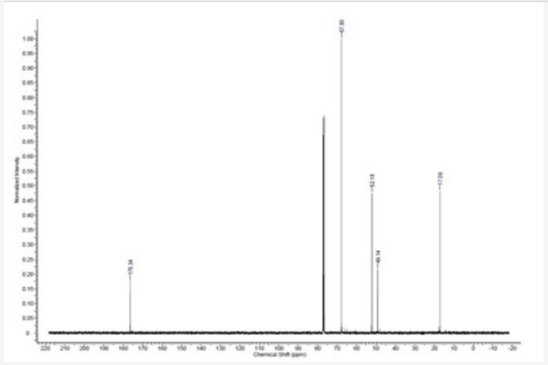


NMR/13C/CDCl3/125

Download as ZIP

Description: 13C NMR (125 MHz, CHLOROFORM-d, ppm) δ = 17.1, 49.1, 52.2, 67.8, 176.3.

Preview Jdx Viewer

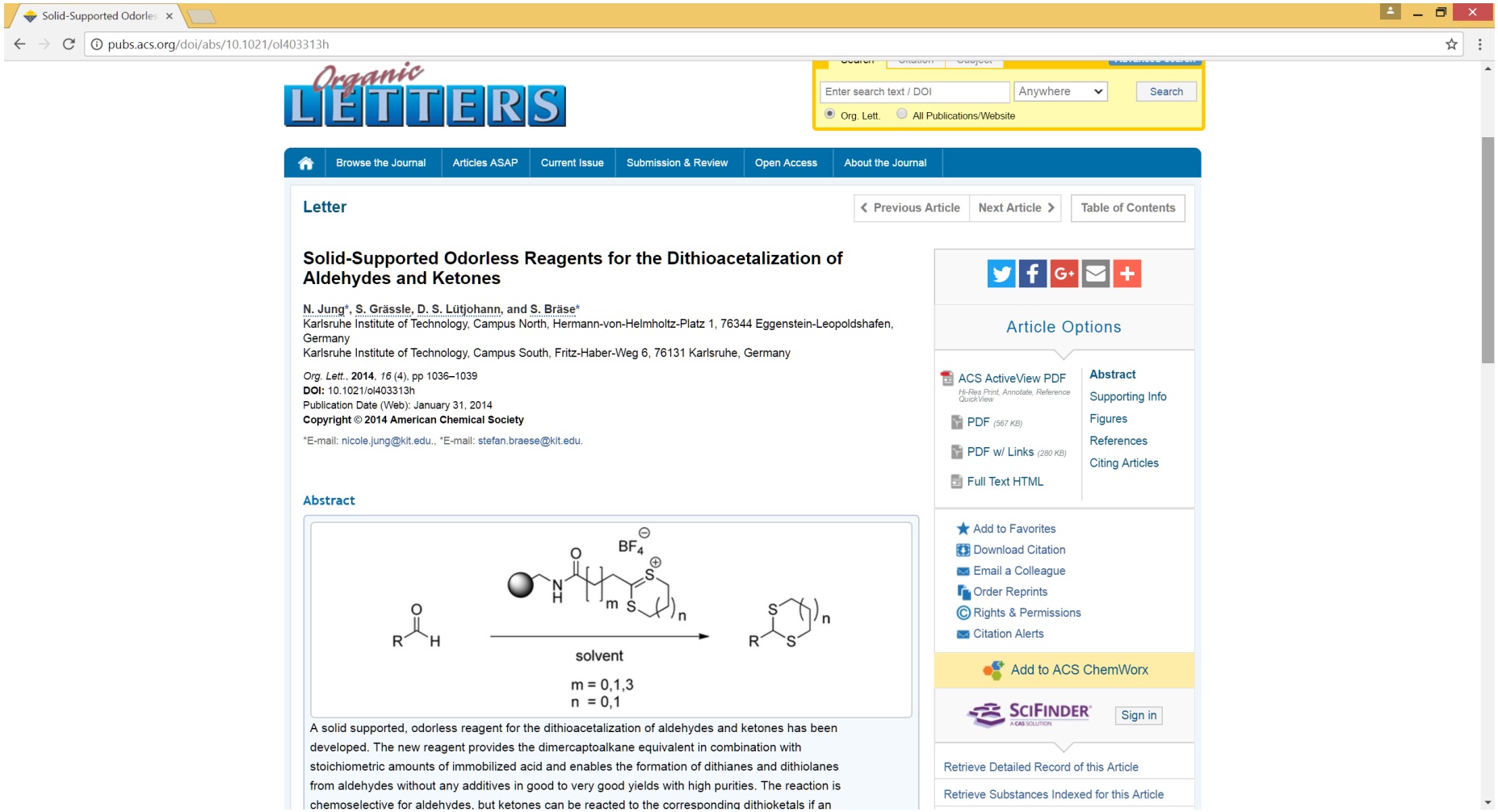


View attached files

To edit this dataset, please fork it first.

	1433_Carbon.jpg	72.58 KB
	SSU-1433.010.001.1R.esp	529.82 KB
	SSU-1433.010.001.1R.spc	263.03 KB
	SSU-1433.010.001.1R.jdx	264.14 KB

Disclosure of information: Repository



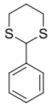
Disclosure of information: Repository

Solid-Supported Odorles | ol403313h_si_001.pdf | ol403313h 1..4

pubs.acs.org/doi/suppl/10.1021/ol403313h/suppl_file/ol403313h_si_001.pdf

2. Individual compounds and analysis

2-Phenyl-1,3-dithiane (2a), CAS: 5425-44-5; Smiles: C1(SCCCS1)c1ccccc1



Benzaldehyde (**1a**) (10.6 mg, 75% purity, 0.075 mmol, 1.00 equiv.) is dissolved in MeCN and 250 mg of the dithianylium resin **8** (0.81 mmol/g, 0.203 mmol, 2.7 equiv.) are added. The reaction is heated to 80 °C and is shaken for 8 h until complete conversion of the starting material is shown via TLC control (see TLC support). The resin is filtered according to work-up procedure A and conversion and purity have been checked via NMR spectroscopy. The obtained filtered product is further purified via a short column chromatography to get 14.7 mg (0.075 mmol) of the target compound in quantitative yield. – R_f = 0.66 (cyclohexane/ethyl acetate; 4:1). – ^{1}H NMR (250 MHz, $CDCl_3$, ppm), δ = 7.49–7.44 (m, 2 H), 7.38–7.28 (m, 3 H), 5.17 (s, 1 H), 3.16–2.87 (m, 4 H), 2.23–1.84 (m, 2 H). – ^{13}C NMR (62.5 MHz, $CDCl_3$, ppm), δ = 25.3, 32.3 (2 C), 51.6, 127.9 (2 C), 128.6, 128.9 (2 C), 139.2. – EI (m/z (%), 70 eV, 80 °C): 196.2 (100), 153.1 (20), 131.2 (70), 122.1 (84), 121.1 (94), 105.1 (37). – HRMS ($C_{10}H_{12}S_2$): calc 196.0380; found 196.0382. – IR (ATR, ν): 2928, 2890, 1648, 1519, 1494, 1450, 1422, 1411, 1274, 1233, 1170, 1026, 908, 882, 831, 723, 695, 673, 594, 502 cm^{-1}.

For the conversion in a larger scale, the reaction was repeated with 1.30 mmol of benzaldehyde, giving 221.6 mg (1.128 mmol) of the target 2-phenyl-1,3-dithiane **2a** in 87% yield.

Further supporting information is available via DOI-referencing to the *Chemotion*-repository:

DOI: 10.14272/GXKPARDRBFURON-UHFFFAOYSA-N

analytics	DOI	My* or others**
NMR		
^{1}H	10.14272/GXKPARDRBFURON-UHFFFAOYSA-N/NMR/1H/CDCL3/250	M
^{13}C	10.14272/GXKPARDRBFURON-UHFFFAOYSA-N/NMR/13C/CDCL3/62.5	M
MASS		
EI	10.14272/GXKPARDRBFURON-UHFFFAOYSA-N/Mass/EI	M
IR		
ATR	10.14272/GXKPARDRBFURON-UHFFFAOYSA-N/IR	M
TLC	10.14272/GXKPARDRBFURON-UHFFFAOYSA-N/TLC	M

My* = the spectra that are available via the chemotion repository have been measured and added by the authors; or others** = we refer herein to spectra that have been provided by other researchers

Disclosure of information: Repository

Solid-Supported Odorles | chemotion | ol403313h 1..4

Secure | https://chemotion.net/molecules/1

chemotion Repository Community Chemicals About

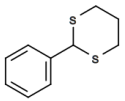
Samples
Molecules
Datasets

Published molecules

chemotion (X) 578

2-phenyl-1,3-dithiane

m



Formula: C10H12S2

Canonical Smiles: C1CSC(SC1)c1ccccc1

Inchi: InChI=1S/C10H12S2/c1-2-5-9(6-3-1)10-11-7-4-8-12-10/h1-3,5-6,10H,4,7-8H2

Inchikey: GXKPARDRBFURON-UHFFFAOYSA-N

Mass: 196.04

Crosslinks: PubChem

Publication (over 3 years ago)

Authors:

Nicole Jung

Simone Gräßle

Affiliation:

Karlsruhe Institute of Technology, Institute of Organic Chemistry (Stefan Bräse Group)

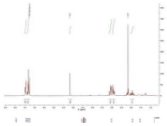
Karlsruhe Institute of Technology, Institute of Organic Chemistry (Stefan Bräse Group)

DOI: 10.14272/GXKPARDRBFURON-UHFFFAOYSA-N

Literature

Solid-Supported Odorless Reagents for the Dithioacetalization of Aldehydes and Ketones
N. Jung, S. Grässle, D. S. Lütjohann, S. Bräse, 2014, 'Solid-Supported Odorless Reagents for the Dithioacetalization of Aldehydes and Ketones', Organic Letters, vol. 16, no. 4, pp. 1036-1039

Characterization



NMR/1H/CDCL3/250

1H NMR (250 MHz, CDCl3, ppm), δ = 7.49–7.44 (m, 2 H), 7.38–7.28 (m, 3 H), 5.17 (s, 1 H), 3.16–2.87 (m, 4 H), 2.23–1.84 (m, 2 H).

DOI: 10.14272/GXKPARDRBFURON-UHFFFAOYSA-N/NMR/1H/CDCL3/250

Bug or Question? Click to report it on GitHub!

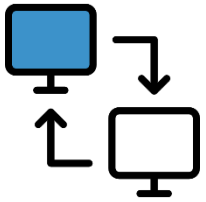
Device Integration: Part I, Remote access

Device Integration

Basic Concepts
Diverse solutions and continuous extension of basic structure: done

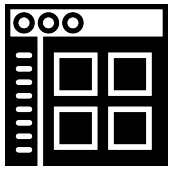
Dial a device (property: group)

- Server infrastructure managing remote access to devices implemented: e.g. HPLC, GC-MS, Raman spectrometer (10)
- Single infrastructure connecting devices Campus North and South
- Virtual machines



ELN device management (property: university)

- User interface for the submission of analysis requests
- NMR implemented (Bruker)

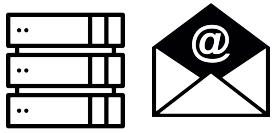


Device Integration: Part II, Data and Information

Device Integration

Basic Concepts
Diverse solutions and continuous extension of basic structure: ongoing

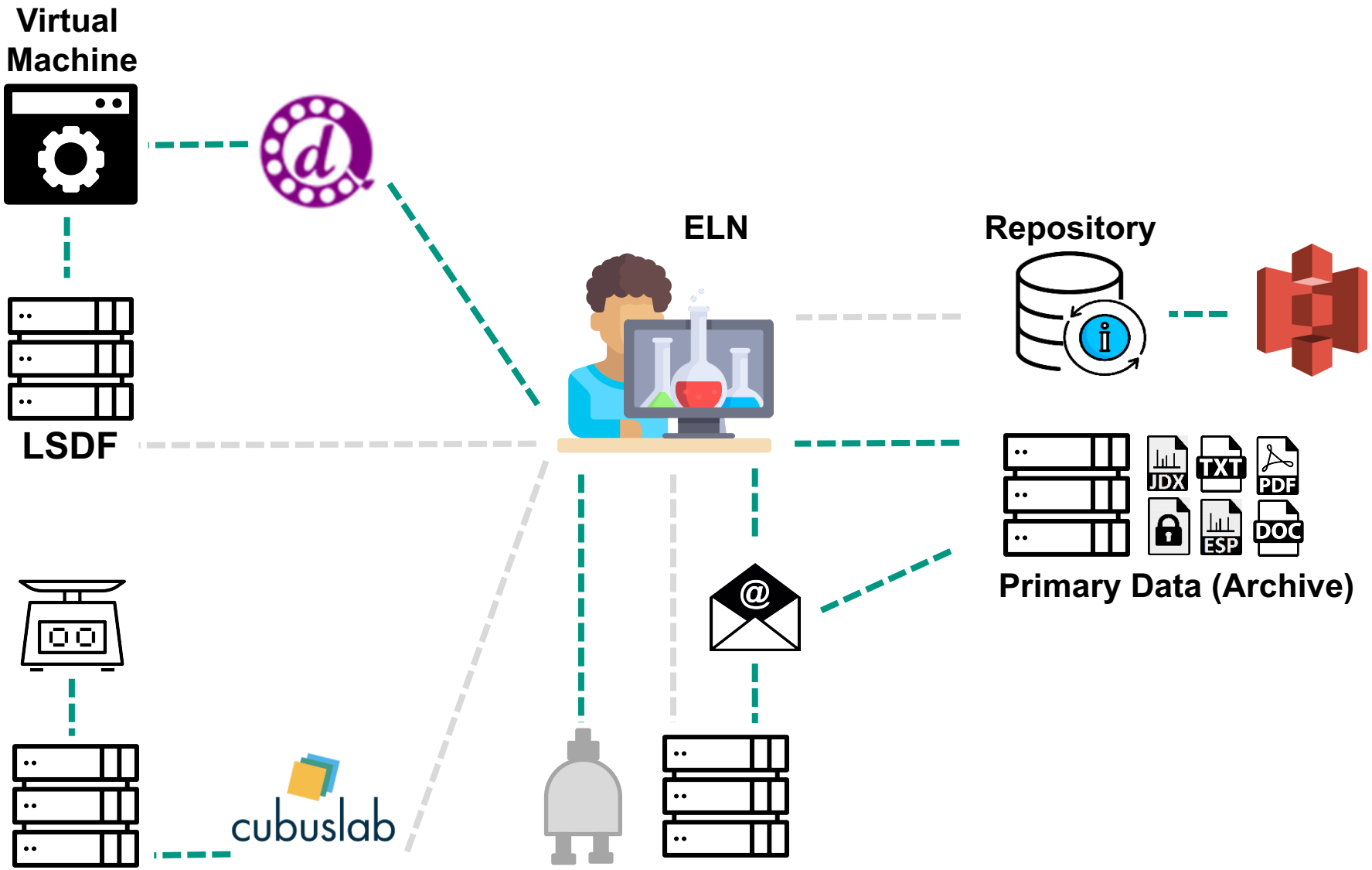
ELN data management
- Server requests (e.g. HPLC, GC)
- Mail-Support (NMR data)
- Primary data storage and archiving



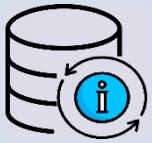
API to external IoT
- Support of cubuslab server
- primary application: balance integration

cubuslab

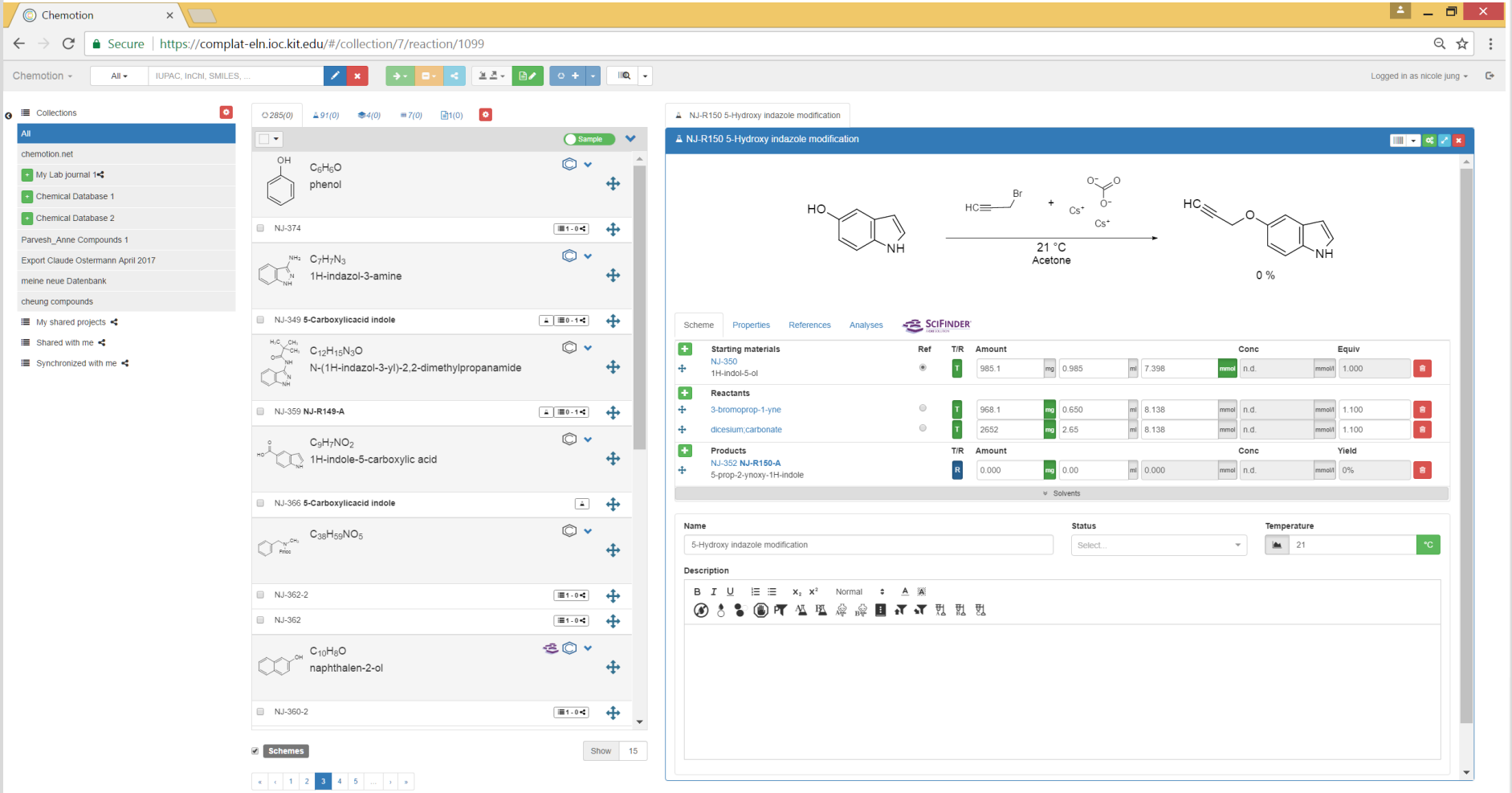
Device Integration: Part II, Data and Information

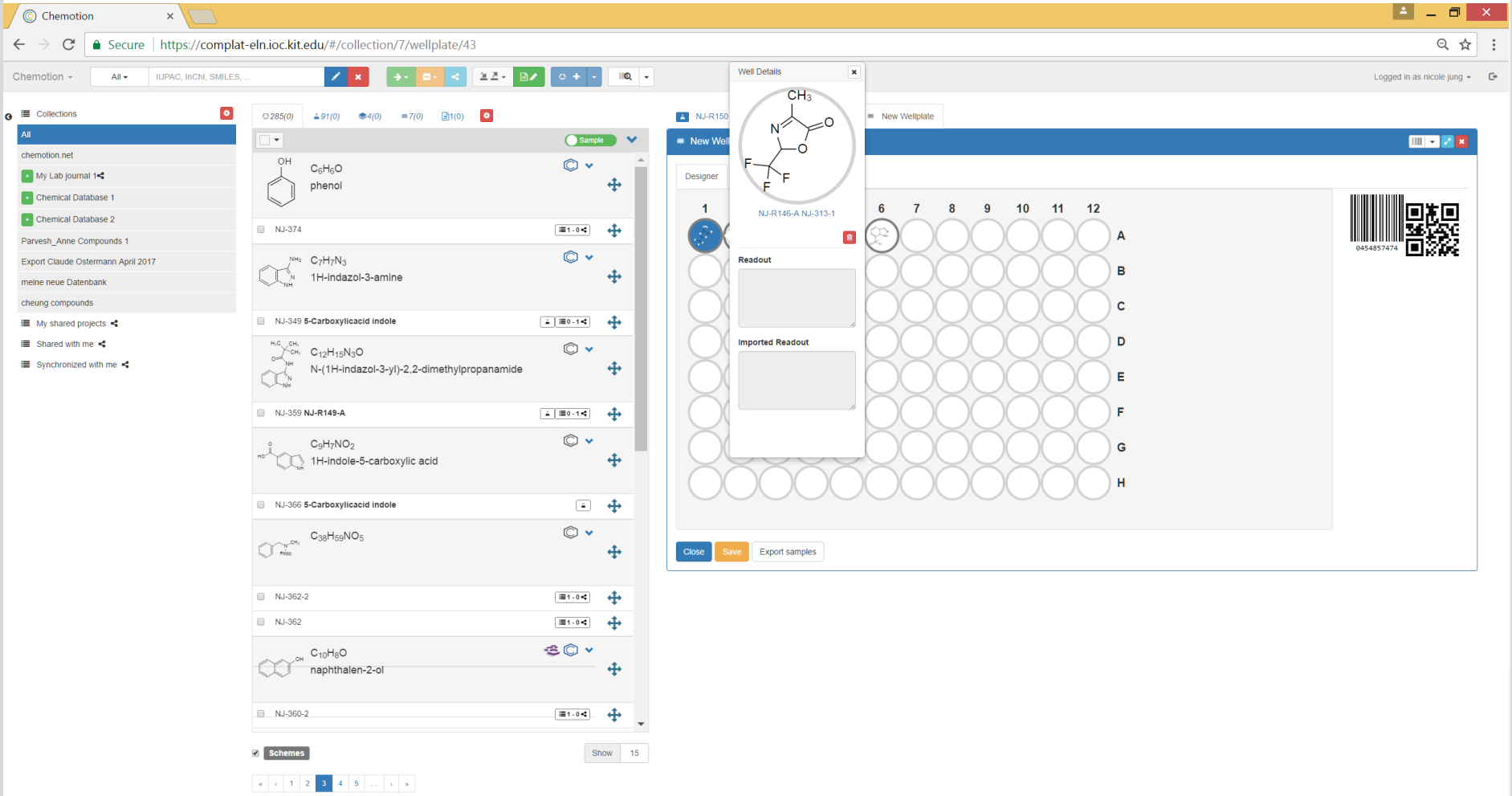


Challenges and questions

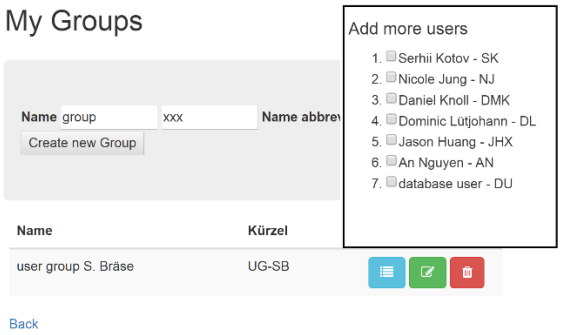
Electronic Lab Notebook (ELN) 	**Repository** 	**Device integration** 
Integration of additional information systems • OpenPhacts (server) • Reaxys • chemical safety data/PubChem	Establishment of automatic proof of analyses	ELN-interface for mass, IR, EA requests, Analysis manager
API to alternative archives • bwDataDiss • Radar	Acceptance of datasets for registration processes	Additional instruments (Cubuslab)

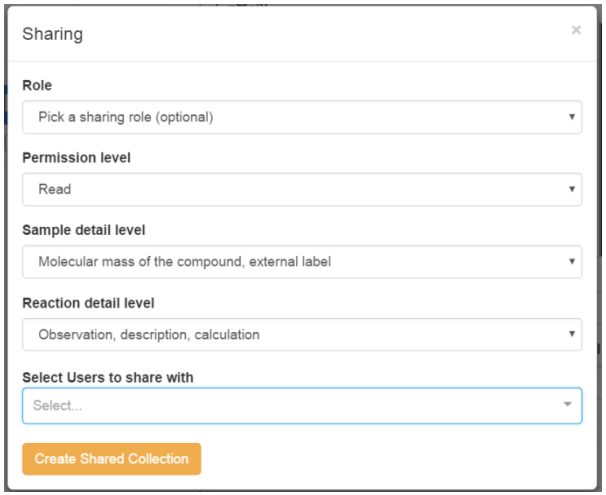
Advanced 	Deep learning support • Data analyses • Reaction prediction

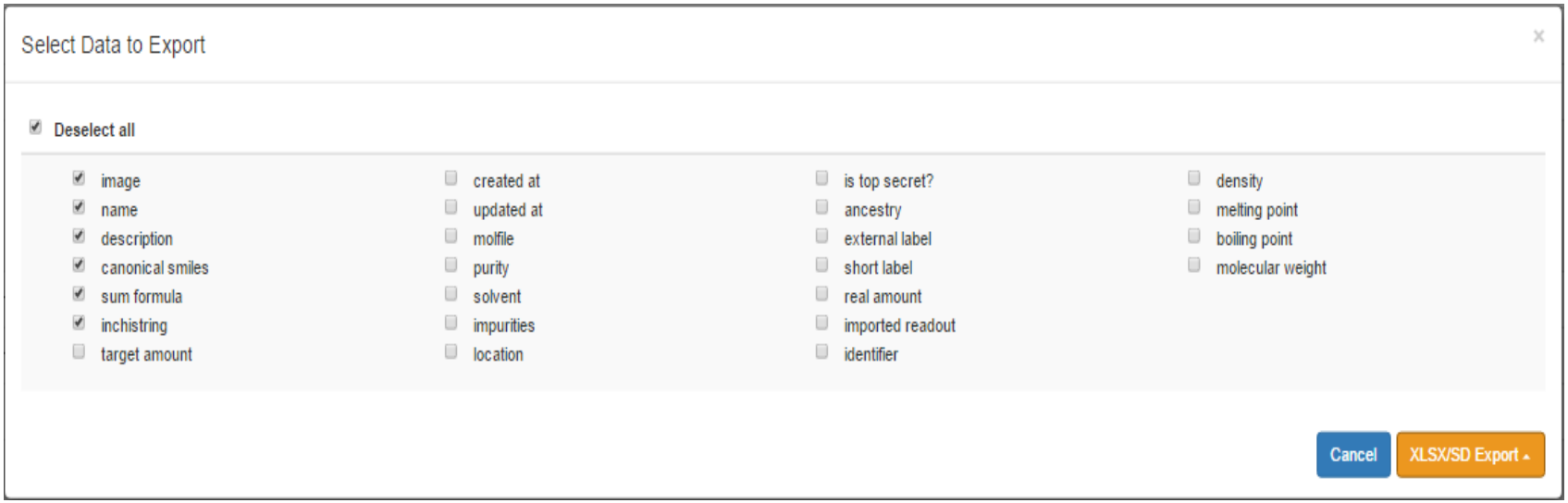

Chemotion
Secure https://complat-eln.ioc.kit.edu/#/collection/7/reaction/1099
Chemotion
All
IUPAC, InChI, SMILES, ...
Logged in as nicole jung
Collections
All
chemotion.net
My Lab journal 1
Chemical Database 1
Chemical Database 2
Parvesh_Anne Compounds 1
Export Claude Ostermann April 2017
meine neue Datenbank
cheung compounds
My shared projects
Shared with me
Synchronized with me
285(0)
91(0)
4(0)
7(0)
1(0)
Sample
C6H6O
phenol
NJ-374
C7H7N3
1H-indazol-3-amine
NJ-349 5-Carboxylicacid indole
C12H15N3O
N-(1H-indazol-3-yl)-2,2-dimethylpropanamide
NJ-359 NJ-R149-A
C9H7NO2
1H-indole-5-carboxylic acid
NJ-366 5-Carboxylicacid indole
C38H59NO5
NJ-362-2
NJ-362
C10H8O
naphthalen-2-ol
NJ-360-2
Schemes
Show 15
NJ-R150 5-Hydroxy indazole modification
NJ-R150 5-Hydroxy indazole modification
HO
NH
HC
Br
Cs+
O-
21 °C
Acetone
0 %
Scheme
Properties
References
Analyses
SciFinder
Starting materials
NJ-350
1H-indol-5-ol
Reactants
3-bromoprop-1-yne
dicesium;carbonate
Products
NJ-352 NJ-R150-A
5-prop-2-ynoxy-1H-indole
Ref
T/R
Amount
Conc
Equiv
Yield
Solvents
Name
5-Hydroxy indazole modification
Status
Select...
Temperature
21
°C
Description
Normal


Chemotion
Secure
https://complat-eln.ioc.kit.edu/#/collection/7/wellplate/43
Chemotion
All
IUPAC, InChI, SMILES, ...
Logged in as nicole jung
Collections
All
chemotion.net
My Lab journal 1
Chemical Database 1
Chemical Database 2
Parvesh_Anne Compounds 1
Export Claude Ostermann April 2017
meine neue Datenbank
cheung compounds
My shared projects
Shared with me
Synchronized with me
285(0)
91(0)
4(0)
7(0)
1(0)
Sample
C6H6O
phenol
NJ-374
C7H7N3
1H-indazol-3-amine
NJ-349 5-Carboxylicacid indole
C12H15N3O
N-(1H-indazol-3-yl)-2,2-dimethylpropanamide
NJ-359 NJ-R149-A
C9H7NO2
1H-indole-5-carboxylic acid
NJ-366 5-Carboxylicacid indole
C38H59NO5
NJ-362-2
NJ-362
C10H8O
naphthalen-2-ol
NJ-360-2
Schemes
Show
15
NJ-R150
New Wellplate
New Wel
Designer
Well Details
NJ-R146-A NJ-313-1
Readout
Imported Readout
1
6
7
8
9
10
11
12
A
B
C
D
E
F
G
H
0454857474
Close
Save
Export samples

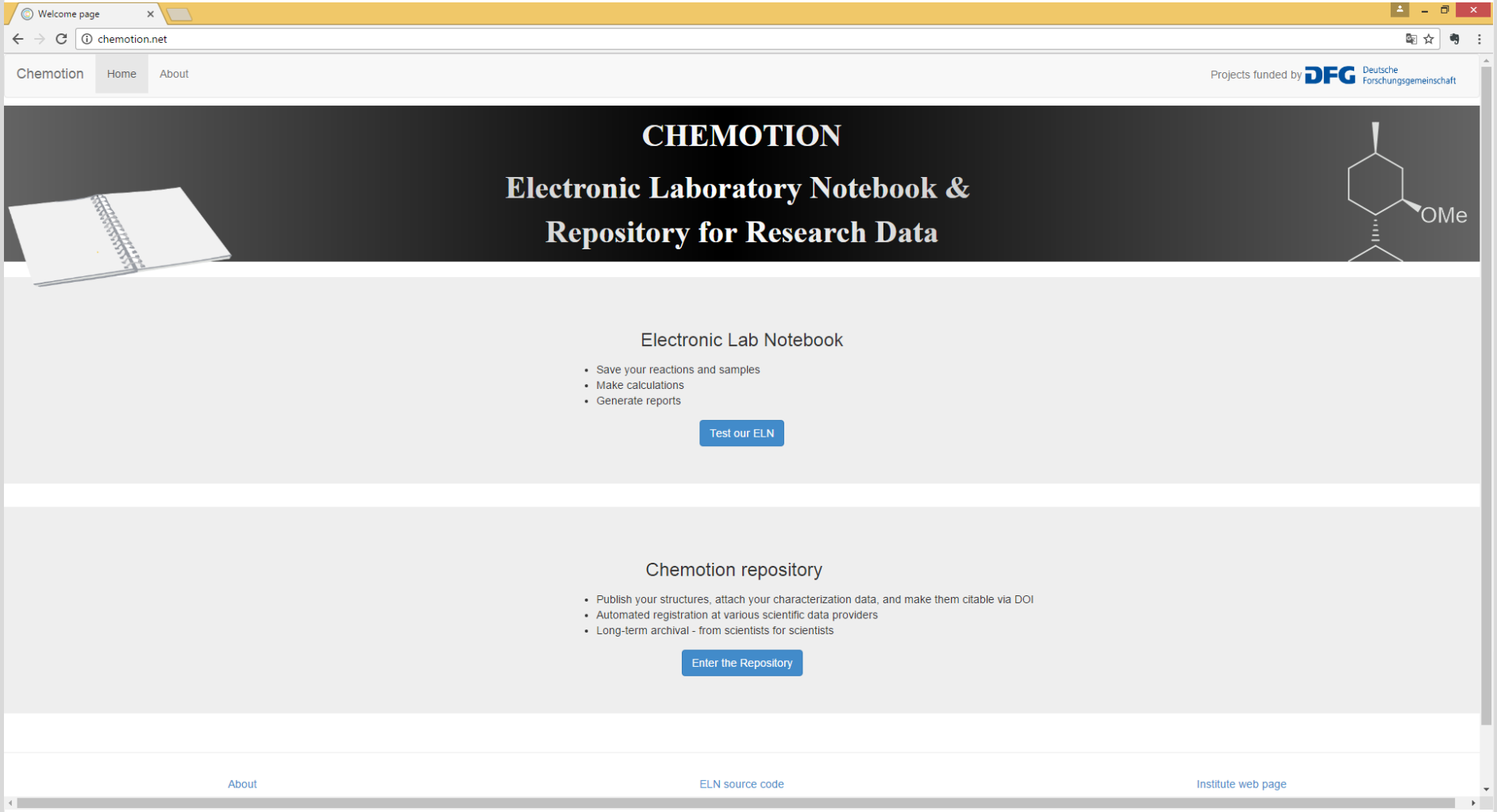
Sharing with researchers or the repository







Chemotion Projects online: www.chemotion.net



Project Partners and Supporters

Karlsruher Institut für Technologie

Institute of Organic Chemistry

re3data.org
REGISTRY OF RESEARCH DATA REPOSITORIES

SciFinder®

RIsources
The Research Infrastructure Portal
funded by DFG

RADAR

PubChem

KIT STIFTUNG

DataCite

Karlsruhe Library

cubuslab

NINJA
CONCEPT.com

Steinbuch Center of Computing

CYNORA
optoelectronic solutions

SIGMA-ALDRICH

Thermoanalytik Karger

KNF NEUBERGER

Gute Chemie

abcr

Fritz Henkel Stiftung

Thank you for your attention

Contact: nicole.jung@kit.edu
stefan.braese@kit.edu