Erfassung und Speicherung von Forschungsdaten im Fachbereich Chemie

Bereitstellung moderner Forschungsinfrastrukturen durch ein elektronisches Laborjournal mit Repositorium-Anbindung

INSTITUTE OF ORGANIC CHEMISTRY - Stefan Bräse Group Karlsruhe

Institute of Organic Chemistry

Summary of the activities of the Compound Platform:

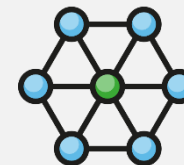- chemistry (service and research) • storage and distribution • programming

Chemistry
- Design
- Synthesis
- Purification
- Analyses

Materials Science
- Upscaling of reactions

Biology and Botany
- Formatting
- Distribution

Additional Institutes
- KIT-Library (BIB)
- Steinbuch Center for Computing (SCC)

Requirement
Basic research
Biologically interesting compound building blocks for materials, helper molecules
Design of the synthesis

Synthesis

Proof by analysis

Application

Storage
(Data and Compound)

New Requirement

Institute of Organic Chemistry

Summary of the activities of the Compound Platform:

- chemistry (service and research) • storage and distribution • programming

Chemistry
- Design
- Synthesis
- Purification
- Analyses

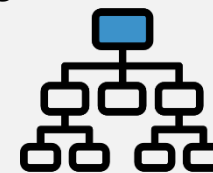### Materials Science
- Upscaling of reactions

Biology and Botany
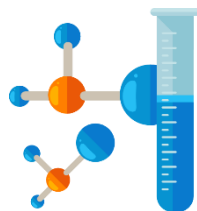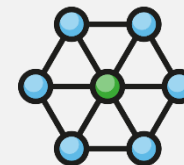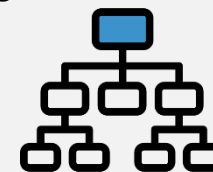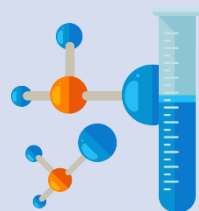- Formatting
- Distribution

Additional Institutes
- KIT-Library (BIB)
- Steinbuch Center for Computing (SCC)

Requirement
Basic research
Biologically interesting compound building blocks for materials, helper molecules
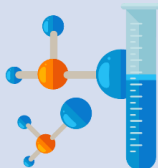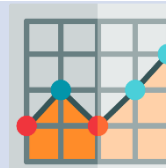Design of the synthesis

➡

Synthesis

➡

Proof by analysis

Application

Storage
(Data and Compound)

New Requirement

Data in Organic Chemistry

Synthesis

Lab-Documentation

Chemicals, solvents, reaction scheme, description, observation, calculations, equipment

Device information

Values, amounts, volumes, time, temperature, conditions

Additional information

Additional info, images

Analysis

Server data (device)

NMR data, spectroscopic data

Desktop data (device)

(GC-, HPLC-)Chromatogramms

Additional Data

Crystal structures, mass spectra, EA, IR- UV-spectra

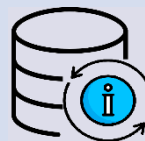# Access to research data: Key challenges

Digital Data

Databases

Journals

Almost no digital data
- Only written lab journals
- not searchable
- Analytical information
  not connected to synthesis
- Main datasets not processed

Commercial databases
- Good search functions
- Limited access
- Limited availability of data

Open Access databases
Only few data

Limited content
- No primary data
- Almost no additional info
- accessibility

Access to research data: Key challenges

Digital Data

Almost no digital data
- Only written lab journals
- not searchable
- Analytical information not connected to synthesis
- Main datasets not processed

ELN
- Allows digital storage
- Processing of structures

Databases

Commercial databases
- Good search functions
- Limited access
- Limited availability of data

Open Access databases
Only few data

Repository
- Free access
- No limitation (file types)

Journals

Limited content
- No primary data
- Almost no additional info
- accessibility

Repository
- Linkage of journal contents to additional data

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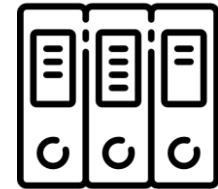
Reporting
- Reports for single elements and projects
- User defined detail level
- Settings for formatting

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

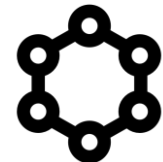# Key instrument: ELN implementation II

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
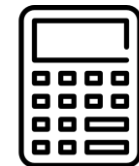- Implementation of abbreviations

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)
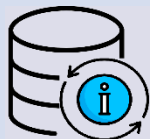- Labeling (approved, not approved)
- All file formats, no limitation

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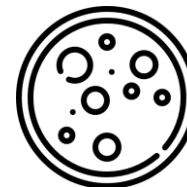
Disclosure of information: Repository

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
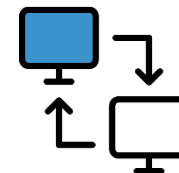
Characteristics
- Search functions based on FP2-fingerprints
- Metadata defined by ELN definitions
- DOI-generation, PubChem listing, Datacite indexing
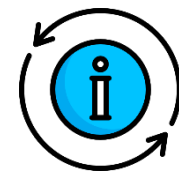
Visibility and Advertisement
- Infrastructure for repositories
- Competitions with incentives (by KIT-Stiftung)

Disclosure of information: Repository

Electronic Lab Notebook (ELN)	Repository	External

Selection
Datasets to publish

Private Repository
Virtual DOI from metadata (molecule-dataset)

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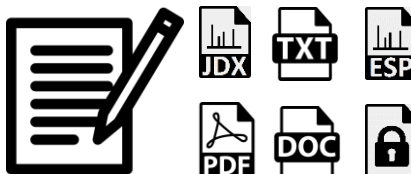DOI – dataset/molecule url (TIB Hannover)

DataCite — Mint and register DOI

PubChem
Structure + DOI

DOI-resolver

Molecule - Datasets
JDX TXT ESP
PDF DOC

Challenges and questions

Electronic Lab Notebook (ELN)	Repository	Advanced
Integration of additional information systems • OpenPhacts (server) • Reaxys	Establishment of automatic proof of analyses	Deep learning support • Data analyses • Reaction prediction
API to alternative archives • bwDataDiss • Radar	Acceptance of datasets for registration processes	
Device Integration • Dial-a-device		
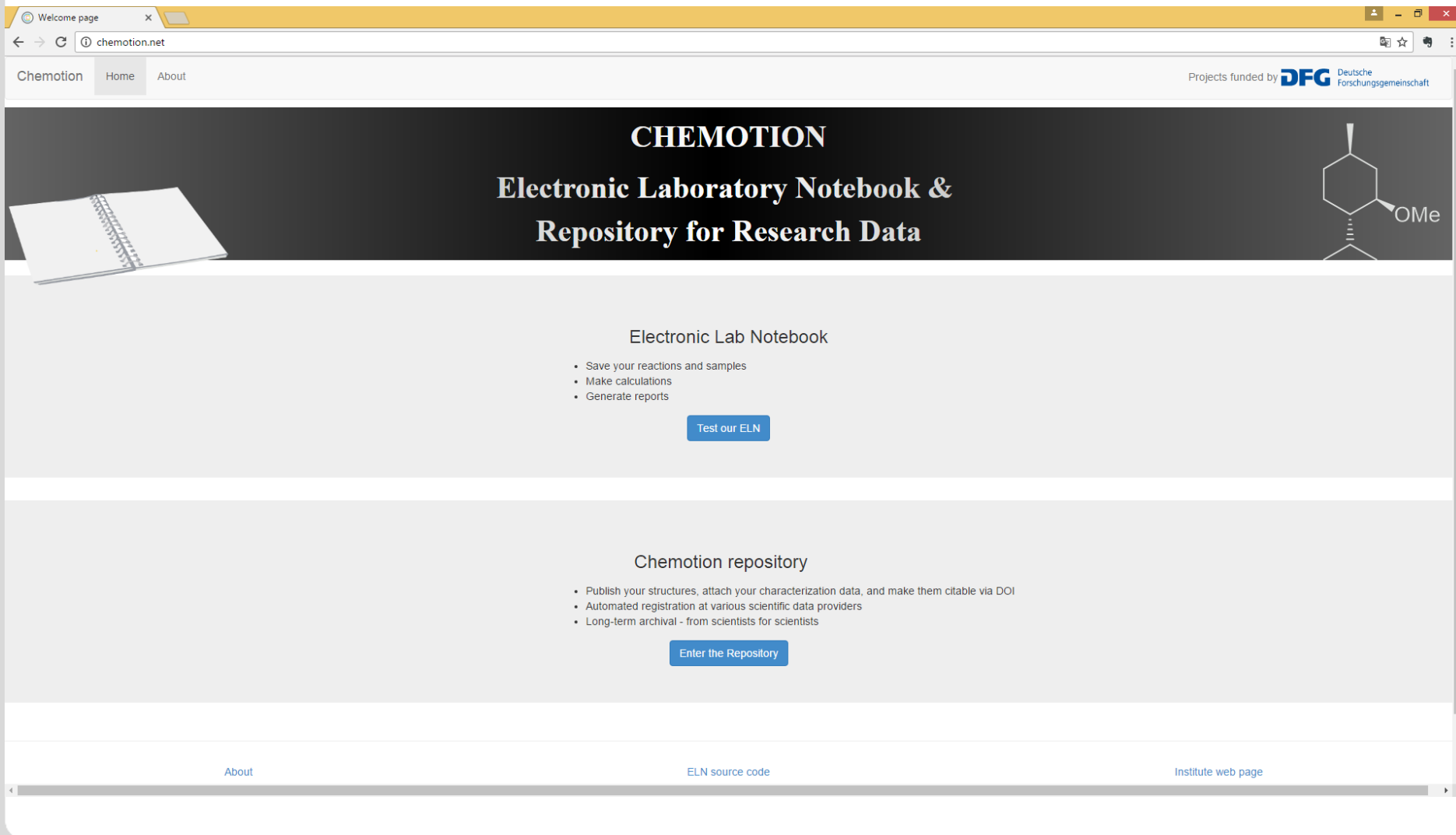
Chemotion Projects online: www.chemotion.net

Thank you for your attention
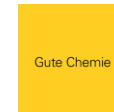
Icons made by *Freepik*, *Madebyoliver*, *Anatoli*, *Dinosoft*, *Vectors market*, *Zlatko Najdenovski,* and *Gregor Cresnar* from www.flaticon.com