

Forschungsdatenmanagement

Notwendige, aber nicht hinreichende Voraussetzung
für den wissenschaftlichen Erkenntnisgewinn

20. Elektronisches Laborbuch

Dr. habil. Till Biskup

Physikalische Chemie

Universität Rostock

12.07.2024





- ❏ Ein Laborbuch dient primär der Protokollierung der Datenerhebung und der Dokumentation erster Ergebnisse.
- ❏ Struktur(ierung) ist essentiell für die Nachnutzbarkeit. Struktur ist wichtiger für Informationen als Digitalität.
- ❏ ELNs dienen nicht der Ablage oder Verarbeitung von Daten. Schnittstellen ermöglichen aber die Informationsweitergabe.
- ❏ Es gibt nicht das eine ELN für alle Anwendungsfälle. ELNs müssen *immer* an die Abläufe angepasst werden.
- ❏ Interoperabilität zwischen ELNs ist ein Wunschtraum. Entscheidend sind Struktur und robuste Exportformate.

Problemstellung, Herausforderungen und Anforderungen

Konzept und Umsetzung

Anwendung auf die eigene Praxis

Vorteile gegenüber Alternativen

Problemstellung

- ▶ Strukturierte Dokumentation der Abläufe im Labor
- ▶ Vollständige Nachvollziehbarkeit, ggf. Reproduzierbarkeit

Antwort auf sechs Fragen

- ❓ Wer hat
- ❓ was
- ❓ mit wem
- ❓ wann
- ❓ wie
- ❓ und warum gemacht?

Herausforderungen

- ▶ Durchdachte digitale Abbildung der Prozesse und Abläufe für möglichst intuitive Bedienbarkeit
- ▶ Strukturierung und größtmögliche Maschinenverarbeitbarkeit ohne Einschränkung der notwendigen Flexibilität und Freiheit

Anforderungen

- ▶ plattformunabhängig, robust, resilient
- ▶ muss im Labor genauso wie vom Schreibtisch erreichbar sein
- ▶ Unterstützung von Formatierung und Medien (Bilder, ...)
- ▶ Mehrwert gegenüber dem analogen Laborbuch: (automatische) Querverweise (Links), Verfügbarkeit an mehr als einem Ort
- ▶ muss Vertraulichkeit (inkl. Zugriffskontrolle) und Datensicherung/Datenschutz gewährleisten können

Kontext

- ▶ wissenschaftlicher Kernaspekt
 - Nachvollziehbarkeit
- ▶ Station im Forschungsdatenlebenszyklus
 - Datenerhebung, Planung
- ▶ Bedeutung
 - Nachvollziehbarkeit der Datenerhebung
 - Ausgangspunkt für weitere Planungen

Beitrag

- ▶ digitale, strukturierte Dokumentation der Abläufe im Labor
 - intuitiv aufgrund intellektueller Durchdringung der Abläufe
- ▶ flexibles, von den Nutzenden weitestgehend anpassbares System

Problemstellung, Herausforderungen und Anforderungen

Konzept und Umsetzung

Anwendung auf die eigene Praxis

Vorteile gegenüber Alternativen

LabInform ELN: A lightweight and flexible electronic laboratory notebook for academic research based on the open-source software DokuWiki

*Mirjam Schröder, Till Biskup**



ChemRxiv, 2023, doi:10.26434/chemrxiv-2023-2tvct

2021-06

san
sampl
bl
tempera
spectrom
probeh
purp

Edit

Protocol

12:30

12:50

14:46

Comments

...

Analysis

The sample shows no ang
 774).

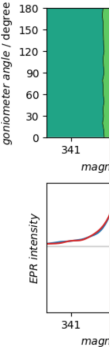


Figure 1: Angular-depde accumulations; MW: 0.14
frequency: 9.60 GHz. Sh
the signal  image as P

Documentation

 Sa772-02.info

cwEPR Info file - v. 0.1.3 (2016-01-18)

GENERAL

Filename: Sa772-02
Date start: 2021-06-11
Date end: 2021-06-11
Time start: 12:51:00
Time end: 14:46:00
Operator: Till Biskup
Label: Sa772 #02
Purpose: Goniometersweep

SAMPLE

Name: PTEG:Mo3-monoanion
ID: 772
Description: doped Film, drop-cast
Solvent:
Preparation: Emmy
Tube: 5 x 4 x 250 mm

EXPERIMENT

Type: field sweep
Runs: 1
Variable parameter: magnetic field
Increment: N/A

SPECTROMETER

Model: Magnettech MS-5000 [11-0096]
Software: Magnettech ESRStudio 1.54.1

MAGNETIC FIELD

Field sweep time: N/A

Typischer Aufbau einer Laborbuchseite

- ▶ Metadaten am Kopf
 - Schlüssel-Wert-Paare
 - zur Aggregation von Informationen in Überblickstabellen genutzt
 - bei Änderungen neue Laborbuchseite anlegen
- ▶ Protokoll
 - chronologisch, typischerweise minutengenau
 - oft mit Querverweisen auf Messungen vorher und/oder nachher
- ▶ Kommentare
 - Einschätzungen, Verweise auf die Literatur, ...
- ▶ Analyse
 - nachträglich außerhalb erzeugt (z.B. ASpecD) und eingebunden
- ▶ Dokumentation
 - Metadaten *während* der Datenerhebung: typischerweise Infofile

Create new entry



You need to have created an entry for a **sample** (and a **batch** the sample originates from) in order to fill out the form below. The sample number refers to the ID of the respective sample and is used for cross-linking within the LabInform ELN.

General parameters for the measurement

Date *

Sample name

Sample ID

Temperature (in K) *

Experiment * cwEPR

Purpose *

cwEPR experiment parameters

Spectrometer * Bruker EMX

Probehead * Bruker 4119HS-W1

Submit form

Create entry

Schlüsselkonzepte

- ▶ Eine Messung findet immer auf einer Probe statt.
- ▶ Eine Probe ist aus einem Batch abgeleitet.

Konsequenzen

- ▶ Vor einer Messung muss eine Probe, vor einer Probe ein Batch angelegt werden.



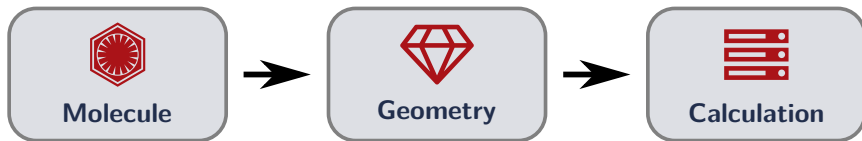
- ☞ Batches und Proben werden automatisch durchnummeriert und automatisch miteinander und mit der Messung verknüpft.

Schlüsselkonzepte

- ▶ Eine Rechnung findet immer auf einer Geometrie statt.
- ▶ Eine Geometrie gehört immer zu einem Molekül.

Konsequenzen

- ▶ Vor einer Rechnung muss eine Geometrie, vor einer Geometrie ein Molekül angelegt werden.



- ☞ Moleküle und Geometrien werden automatisch durchnummeriert und automatisch miteinander und mit der Rechnung verknüpft.

PNDIT2

ID: 348
Substanz: PNDIT2
Datum: 2018-06-29
Bezugsquelle: ~~Reinstoffe-Werkstoffe GmbH / M&C~~
Chiffre: RM 167.1
Solvent: Toluol
Zustand: da
Lagerort: -70°C

Bearbeiten

Bearbeiten

Proben

ID	Bezeichnung	Substrat	Datum	Vermessen
505	PNDIT2		2018-06-30	ja
499	PNDIT2	Glas (Ilmasil PS)	2018-06-30	ja

Bearbeiten

Bemerkungen

- 5.6 mg eingewogen und in 1120 µL Toluol gelöst
- $\Rightarrow c = 5 \text{ mg/mL}$

PNDIT2

ID: 505
Bezeichnung: PNDIT2
Datum: 2018-06-30
Batch: [PNDIT2](#)
Konzentration: 5 mg/mL
Vermessen: ja
Zustand: da
Lagerort: LN2

Bearbeiten

Bearbeiten

Bemerkungen

- 150 µL aus Batch [348](#) in ein ESR-Röhrchen abgefüllt und weggefroren
- Lagert in Dewar 1, Hänger 7, TB 43

Bearbeiten

Messungen

Laborbucheintrag	Kommentar
trepr/2018/07/02_1059	m01: 680 nm, 2 mJ/Puls, OK

Was ist ein ELN? Und was ist es *nicht*?

Versuch einer sinnvollen Definition



essentielle Bestandteile eines ELN

- ▶ Ersatz für das analoge, papierbasierte Laborjournal
- ▶ Zugriff sowohl vom Labor als auch vom Schreibtisch (→ webbasiert)
- ▶ Bilder und Medien
- ▶ strukturierte Inhalte (Schlüssel-Wert-Paare)
- ▶ vorlagenbasierte Erstellung neuer Seiten
- ▶ Anpassbarkeit der Vorlagen über die Nutzerschnittstelle
- ▶ tabellarischer Überblick über Experimente/Messungen mit Möglichkeit der Sortierung und Filterung
- ▶ Export aller Inhalte in ein generisches Format

Was ist ein ELN? Und was ist es *nicht*?

Versuch einer sinnvollen Definition



optionale Bestandteile eines ELN

- ▶ Inventar für Proben
- ▶ Schnittstelle zu Auswertungsroutinen, um Ergebnisse von Analysen automatisiert ins Laborbuch einzutragen

keine Bestandteile eines ELN

- ▶ Datenspeicher/Repositorium
- ▶ Forschungsdatenkatalog
- ▶ Management der Messaufbauten^(LIMS?)
- ▶ allgemeines Projektmanagement inkl. Anträge und Publikationen
- ▶ Inventar für Verbrauchsmaterialien^(LIMS?)
- ▶ Workflows für die Datenauswertung

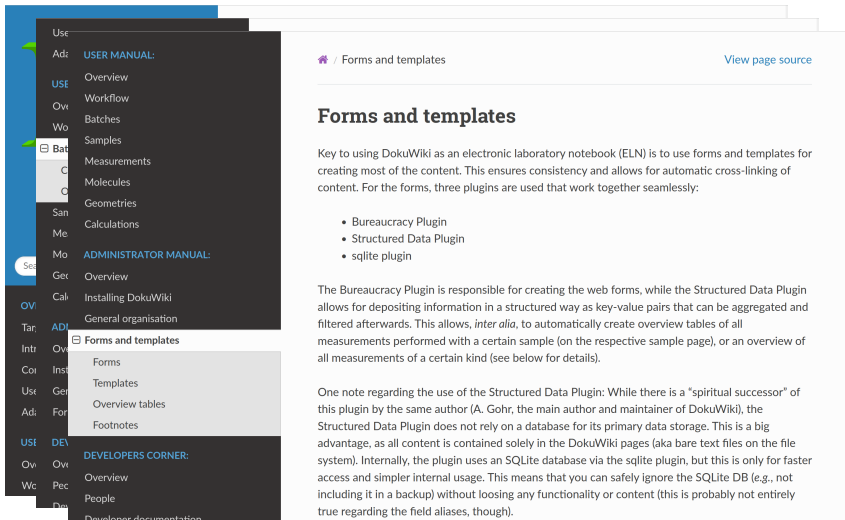
Problemstellung, Herausforderungen und Anforderungen

Konzept und Umsetzung

Anwendung auf die eigene Praxis

Vorteile gegenüber Alternativen

- ▶ eigene Abläufe intellektuell durchdringen und verbalisieren
 - zwingende Voraussetzung für die Digitalisierung
 - gute Gelegenheit, bisherige Abläufe kritisch zu hinterfragen
 - ▶ Abgleich mit Strukturen, die LabInform ELN vorgibt
 - Batch → Probe → Messung, Molekül → Geometrie → Rechnung
 - Passt was davon? Lässt sich was verwenden/erweitern?
 - typische Erweiterungen: Synthesen, Probengefäße, ...
 - ▶ Vorlagen für die eigenen Anwendungsfälle entwickeln
 - generelle Struktur sollte erhalten bleiben:
strukturierte Kopfdaten mit Links, Protokoll, Beschreibung
 - Vorlagen lassen sich *aus dem Wiki heraus* bearbeiten/anpassen
 - Vorlagen für Bereiche einer Seite sind ebenfalls möglich
- ☞ aktuell ca. 10 individuelle LabInform ELN-Installationen

The image is a screenshot of a web browser displaying the DokuWiki 'Forms and templates' page. On the left, a dark sidebar menu is open, showing a tree structure of the site's content. The 'Forms and templates' item is highlighted. The main content area on the right has a light blue header with a breadcrumb trail and a 'View page source' link. The title 'Forms and templates' is in a large, bold, dark font. Below the title, a paragraph explains the key to using DokuWiki as an electronic laboratory notebook (ELN). A bulleted list follows, listing three plugins: Bureaucracy Plugin, Structured Data Plugin, and sqlite plugin. Another paragraph describes the Bureaucracy Plugin's function. A final paragraph provides a note about the Structured Data Plugin, mentioning its use of a SQLite database. The sidebar menu includes sections for 'USER MANUAL', 'ADMINISTRATOR MANUAL', and 'DEVELOPERS CORNER', each with sub-items like 'Overview', 'Workflow', 'Batches', etc.

Use

Ad: **USER MANUAL:**

- Overview
- Workflow
- Batches
- Samples
- Measurements
- Molecules
- Geometries
- Calculations

Se: **ADMINISTRATOR MANUAL:**

- Overview
- Installing DokuWiki
- General organisation

OV: **Forms and templates**

- Forms
- Templates
- Overview tables
- Footnotes

US: **DEVELOPERS CORNER:**

- Overview
- People
- Developer documentation

/ Forms and templates [View page source](#)

Forms and templates

Key to using DokuWiki as an electronic laboratory notebook (ELN) is to use forms and templates for creating most of the content. This ensures consistency and allows for automatic cross-linking of content. For the forms, three plugins are used that work together seamlessly:

- Bureaucracy Plugin
- Structured Data Plugin
- sqlite plugin

The Bureaucracy Plugin is responsible for creating the web forms, while the Structured Data Plugin allows for depositing information in a structured way as key-value pairs that can be aggregated and filtered afterwards. This allows, *inter alia*, to automatically create overview tables of all measurements performed with a certain sample (on the respective sample page), or an overview of all measurements of a certain kind (see below for details).

One note regarding the use of the Structured Data Plugin: While there is a "spiritual successor" of this plugin by the same author (A. Gohr, the main author and maintainer of DokuWiki), the Structured Data Plugin does not rely on a database for its primary data storage. This is a big advantage, as all content is contained solely in the DokuWiki pages (aka bare text files on the file system). Internally, the plugin uses an SQLite database via the sqlite plugin, but this is only for faster access and simpler internal usage. This means that you can safely ignore the SQLite DB (e.g., not including it in a backup) without losing any functionality or content (this is probably not entirely true regarding the field aliases, though).

Problemstellung, Herausforderungen und Anforderungen

Konzept und Umsetzung

Anwendung auf die eigene Praxis

Vorteile gegenüber Alternativen

Eine Reihe quelloffener bekannterer ELNs

▶ Chemotion

- ursprünglich von und für synthetische Chemiker entwickelt
- von einer Uni-Arbeitsgruppe entwickelt (keine Softwareentwickler)
- will sehr viel sein: ELN, Repositorium, Datenauswertung, ...

▶ elabFTW

- hochgradig modular, muss aber komplett angepasst werden
- explizit nicht hierarchisch: ausschließlich Tags
- im Wesentlichen ein Entwickler

▶ openBIS

- Fokus auf Abläufe in der Biologie
- ist sehr viel mehr als ein ELN
- von einer Uni-Arbeitsgruppe entwickelt (keine Softwareentwickler)

☞ gute Kandidaten, wenn sie zu den eigenen Anforderungen passen

robust, minimale Anforderungen

- ▶ keine Datenbank: alles in Textdateien im Dateisystem
- ▶ robuste und bewährte Wiki-Engine (DokuWiki)
- ▶ ausdrucksstarkes Wiki-Markup, einfach erlernbar
- ▶ einfache Exportmöglichkeiten (u.a. PDF, ODT)
- ▶ generische API zum Zugriff von außen (erlaubt Automatisierungen)
- ▶ Konfiguration und Anpassung fast ausschließlich browserbasiert

Trennung von Zuständigkeiten

- ▶ Fokus: Dokumentation der Abläufe im Labor
- ▶ keine Ablage von Daten
 - LabInform Datensafe als Repositorium für warme Forschungsdaten
- ▶ keine Datenverarbeitung
 - ASpecD-Framework zur lückenlos protokollierten Datenverarbeitung



- ❏ Ein Laborbuch dient primär der Protokollierung der Datenerhebung und der Dokumentation erster Ergebnisse.
- ❏ Struktur(ierung) ist essentiell für die Nachnutzbarkeit. Struktur ist wichtiger für Informationen als Digitalität.
- ❏ ELNs dienen nicht der Ablage oder Verarbeitung von Daten. Schnittstellen ermöglichen aber die Informationsweitergabe.
- ❏ Es gibt nicht das eine ELN für alle Anwendungsfälle. ELNs müssen *immer* an die Abläufe angepasst werden.
- ❏ Interoperabilität zwischen ELNs ist ein Wunschtraum. Entscheidend sind Struktur und robuste Exportformate.