

# Von der Laborbank zum globalen Austausch von Forschungsdaten

Erfahrungen aus Chemie und Biologie

Christoph Steinbeck

European Bioinformatics Institute, Cambridge  
Friedrich-Schiller-Universität, Jena

# Biologie

Seit den frühen 90'ern  
herrscht ein engagierter  
Geist des offenen Teilens



Seit 1988



Seit 2013



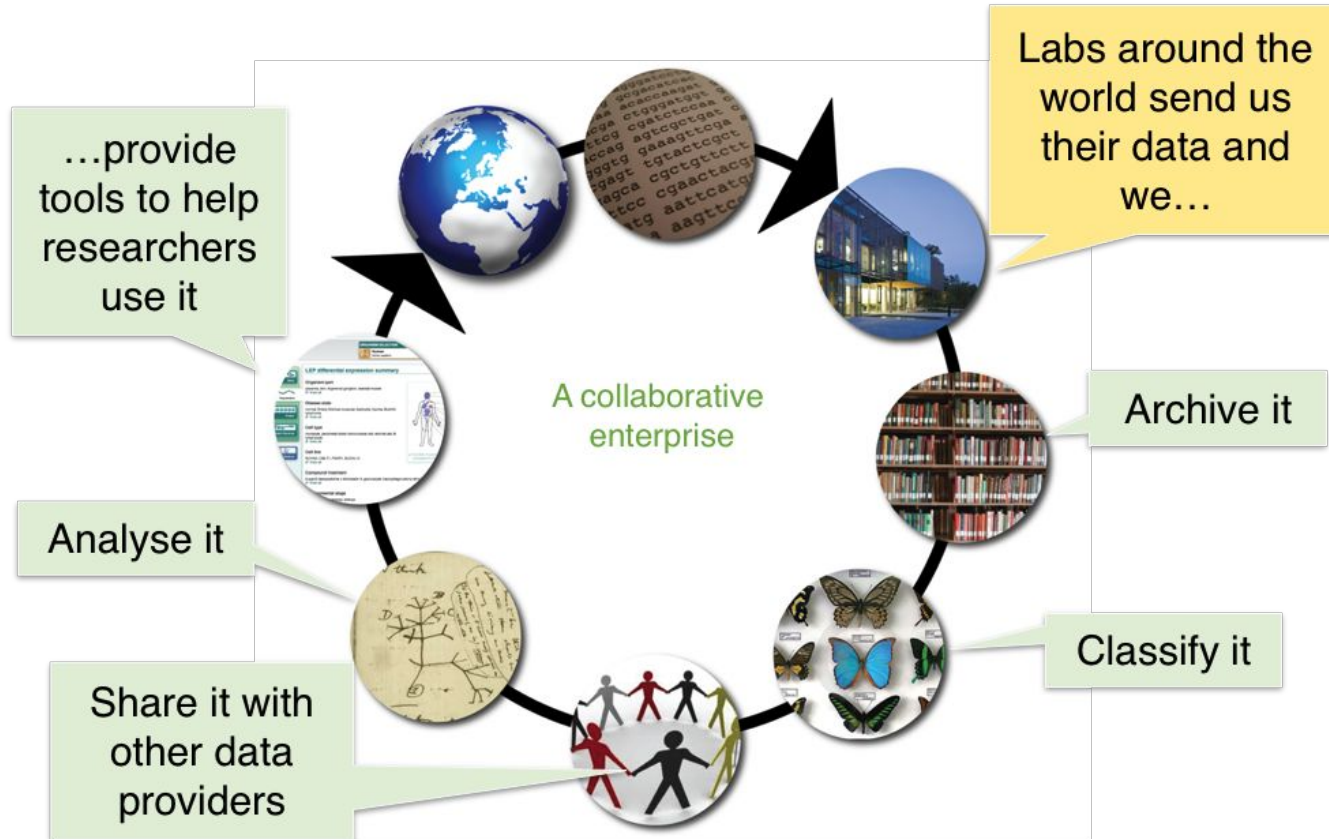
Seit 1986

EMBL-EBI



Seit 1994

# What EBI repositories do



# Biologie



Open Source



Offene Datenformate



Ontologien und kontrollierte Vokabularien



Mindestinformationsstandards (MI)

# Chemie

Datenbanken und Datenstandards  
fest in kommerzieller Hand



**MDL**<sup>®</sup>  
*Information Systems, Inc.*



**SCIFINDER**<sup>®</sup>  
A CAS SOLUTION



Reaxys<sup>®</sup>



**DAYLIGHT**  
Chemical Information Systems, Inc.

# Chemie

Zwei große offene Erfolgsgeschichten ...

Pub**C**hem

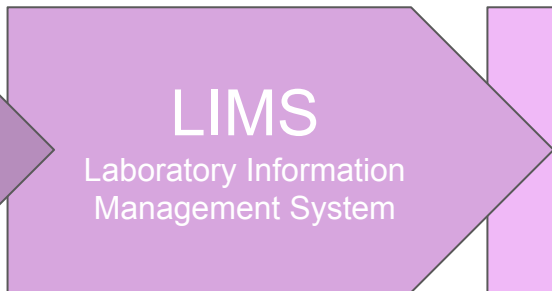
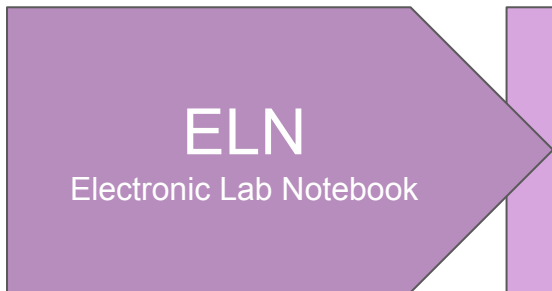


... aber keine global agierenden Forschungsdatenarchive



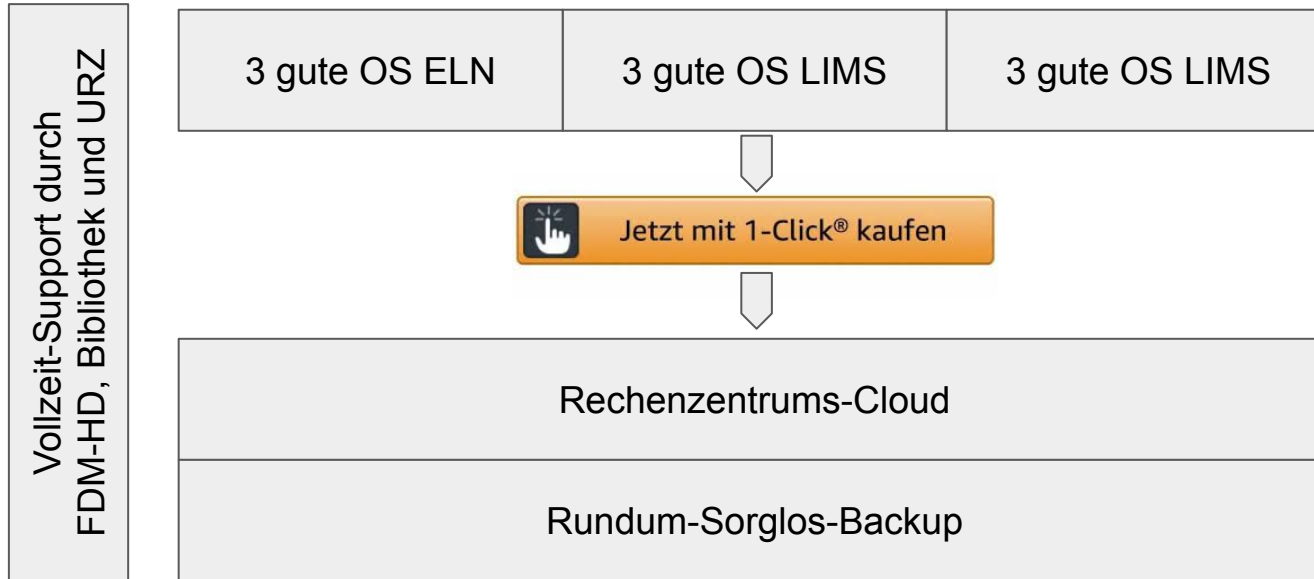
# Biologie und Chemie

Die fehlenden Bausteine:





# Was kann die Universität tun?





# NFDI4Chem

Sonja Herres-Pawlis, Oliver Köpler, Christoph Steinbeck

1. NFDI-Konferenz, 13.-14. Mai 2019, Bonn

# Die NFDI im Allgemeinen

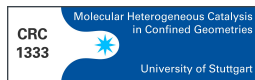
---



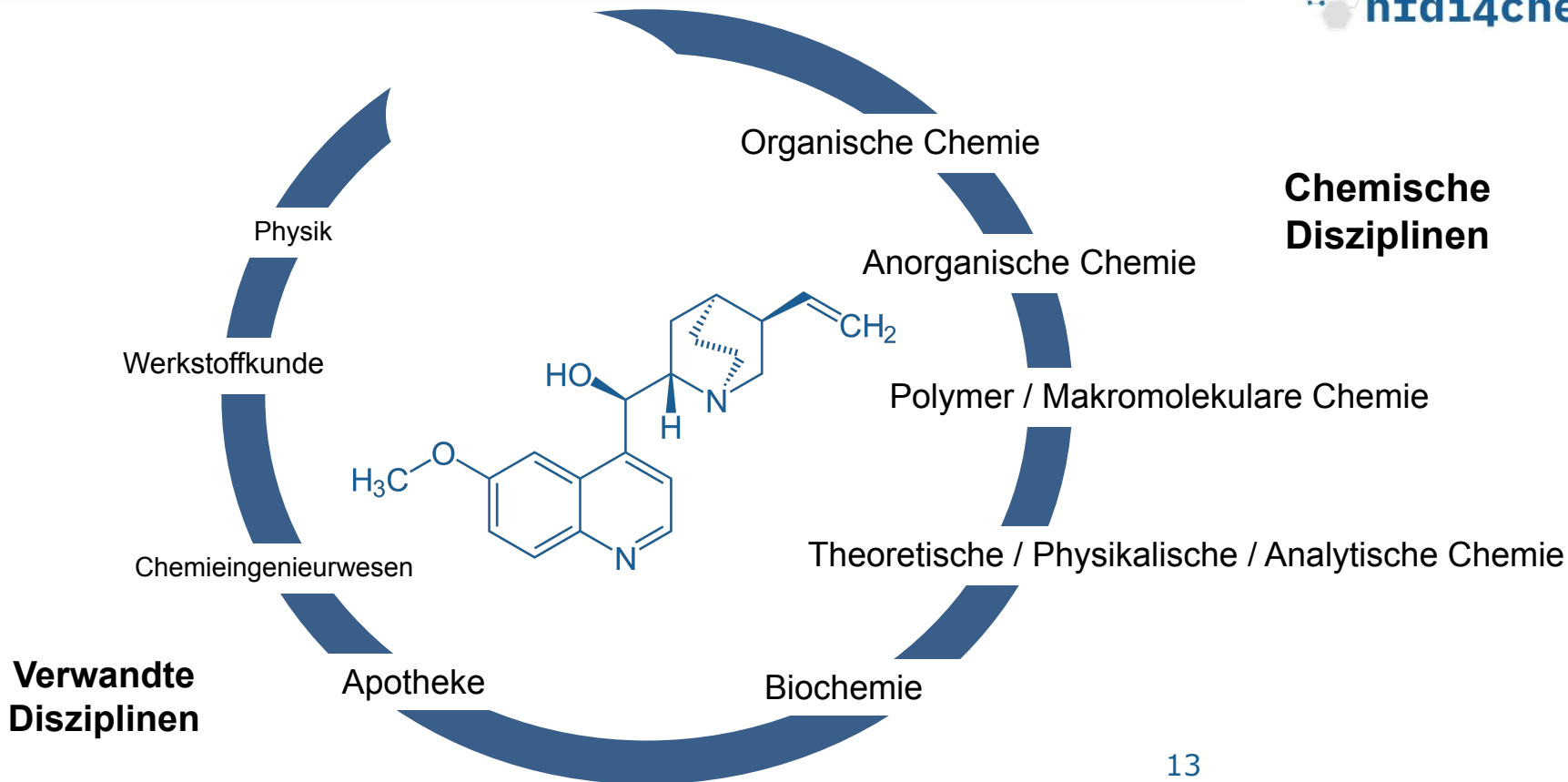
Bund und Länder finanzieren eine nationale Forschungsdateninfrastruktur (NFDI) für Deutschland

- Bis zu 85 Mio Euro Fördermittel **pro Jahr** im Endausbau
- Gefördert werden 30 Konsortien über alle Bereichen der Wissenschaft
  - Erste Konferenz Mai 2019 mit 57 Konsortien
- Erste Anträge von etwa 5-10 Konsortien im Oktober 2019
- Erste und zweite Förderperiode:
  - 5 Mio pro Konsortium und Jahr für 5 Jahre
- Start der Arbeiten Oktober 2020

# Wer wir sind - Basisbewegung



# Unsere wissenschaftliche Gemeinschaft



# Hauptziele

---



**Ziel 1:** Verbinden vorhandener Datenrepositories, Entwicklung fehlendes Forschungsdatenrepositories, Verknüpfung mit internationalen Repositories.



**Ziel 2:** Mindestinformationsstandards (MI) für Daten und maschinenlesbare Metadaten, offene Datenstandards, um die FAIR-Grundsätze für Forschungsdaten zu unterstützen.



**Ziel 3:** Förderung elektronischer Laborbücher (ELN), Tools und APIs zwischen Instrumentierung und Software für eine eingebettete digitale Informationsarchitektur. Erfassen von Forschungsdaten in gut kommentierter elektronischer Form zum frühestmöglichen Zeitpunkt im Forschungsprozess.



# Hauptziele

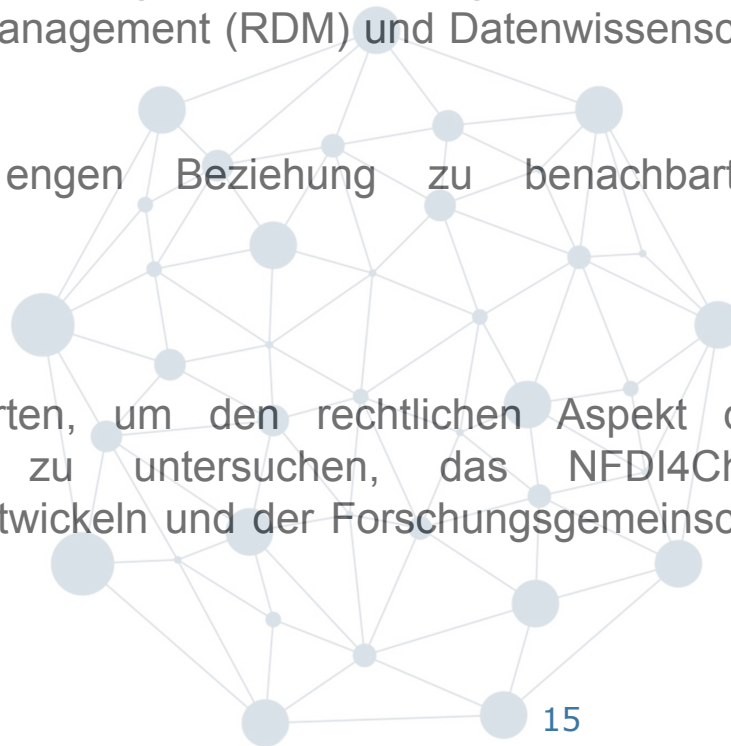
---



**Ziel 4:** Sensibilisierung für FAIR-Datenmanagement, Initiierung von Prozessen zur Integration von Forschungsdatenmanagement (RDM) und Datenwissenschaft in Lehrpläne.

**Ziel 5:** Aufrechterhaltung einer engen Beziehung zu benachbarten NFDI-Konsortien.

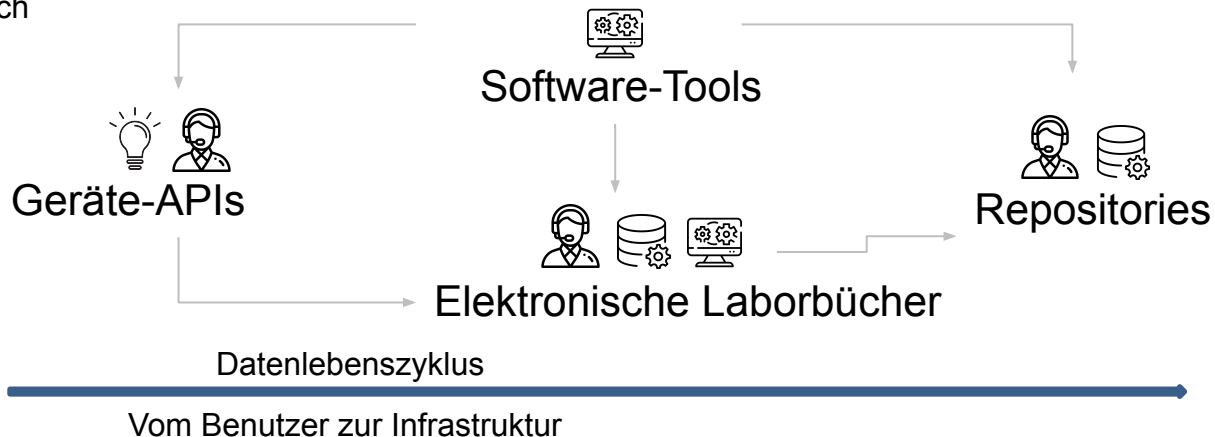
**Ziel 6:** Zusammenarbeit mit Experten, um den rechtlichen Aspekt des FAIR-Forschungsdatenmanagements zu untersuchen, das NFDI4Chem entsprechend zu entwerfen und zu entwickeln und der Forschungsgemeinschaft Ratschläge zu geben.



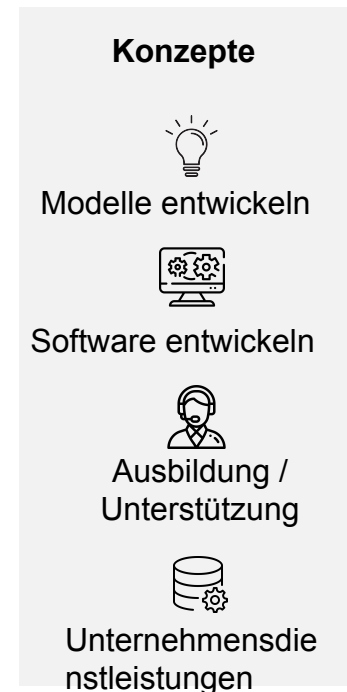
# Digitaler Wandel in der Chemie



Spezifisch



Generisch





# Datenaspekte

## Datenformate

Offene Standards  
Offene Formate  
*gegen*  
Heterogenität  
Eigentum

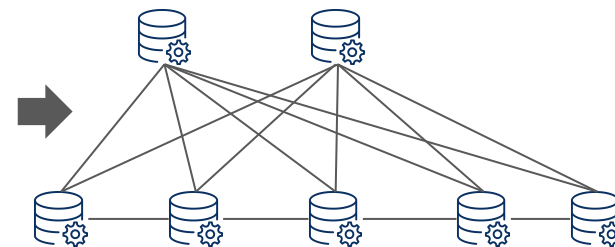
## Metadaten

Offene  
Standards  
Offene Formate

## Vokabeln

Abbildung  
heterogener Daten-  
und  
Metadatenformate  
auf  
domänenspezifische  
Vokabulare

## FAIR Prinzipien



Generisch



Standards und  
Formate



Gemeinschaften



Legale Aspekte

# Internationales Netzwerk

---



1. **Sensibilisierung für das Forschungsdatenmanagement in der Chemie.**
2. **Fokus auf Moleküle und Daten für experimentelle und theoretische Charakterisierung und Reaktion.**



Sonja Herres-Pawlis

RWTH Aachen



Oliver Koepler

Leibniz-Informationszentrum  
Wissenschaft und Technologie  
(TIB)



Christoph Steinbeck  
Friedrich-Schiller-Universität Jena