

Behind the Scenes at **Chemotion**: A Popular **Research Software Project**

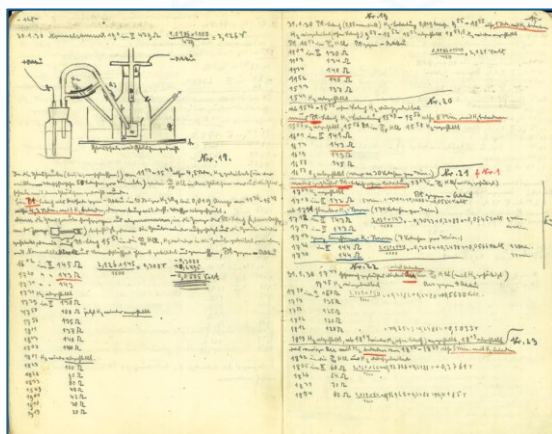
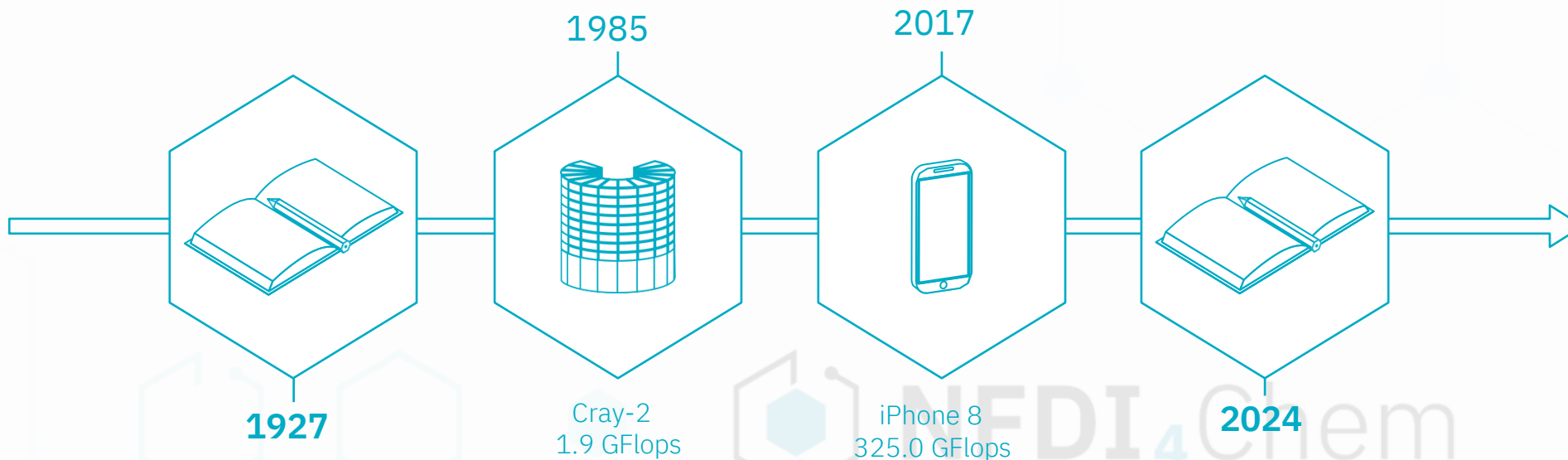


NFDI₄Chem

ENHANCE
YOUR
DATA.

Dr. **Shashank** S. Hariviyasi
Research Software Engineer
Karlsruhe Institute of Technology

Status Quo



SG-SS-V43

Übster:
J. Chem. Soc., Perkin Trans. I, 1936, 557. Hg. UG-SS-274 (3-fach)

Reaktion:

UG-64-430 $\xrightarrow{\text{SO}_2\text{O}_2, \text{DHPF}}$ SG-SS-V43
M: 285,34 g/mol

Result:

	H Etmol	eq	n Etmol	m Etmol	V Etmol
UG-64-430	185,32	1	14,634	2,74	
DHPF	122,49	0,1	1,4634	0,148	
SO ₂ O ₂	248,17	1,52	22,633	4,854	
Reaktionsm. (Lsg.)					2,41

$\delta(\text{SO}_2\text{O}_2) = 0,3850 \text{ g/mL}$
 2,5 mL Reaktionsm. pro mmol UG-64-430
 0,5 mL Reaktionsm. pro mmol SO₂O₂

Vorbereitung:
 Eine Lösung aus 2,94 g UG-64-430 (14,634 mmol, eq) in 56 mL Reaktionsm. wurde bei 0°C unter Rühren mit 0,148 g DHPF (1,4634 mmol, 0,1 eq) versetzt. Reaktionsm. wurde über einen Tropftrichter über 10 Minuten in eine Lösung aus 5,41 mL SO₂O₂ (22,64 mmol, 1,52 eq) und 1,5 mL Reaktionsm. gegeben. Die Reaktion findet unter Schütteln statt. Die Reaktionsmischung wird 30 Minuten bei 0°C gerührt und anschließend nach 1,5 Stunden bei Raumtemperatur. Das Lösungsmittel entfernen und die Säulenchromatographie durchführen.

Säulenchromatographie:

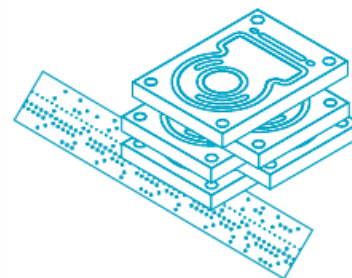
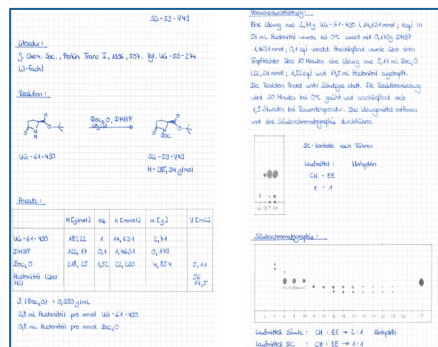
DC-Kontrolle nach Fäulen
 Laufmittel: CH₂Cl₂ / EE
 1 : 1

Säulenchromatographie:

Laufmittel: CH₂Cl₂ / EE → 2 : 1
 Laufmittel: CH₂Cl₂ / EE → 1 : 1

Status Quo

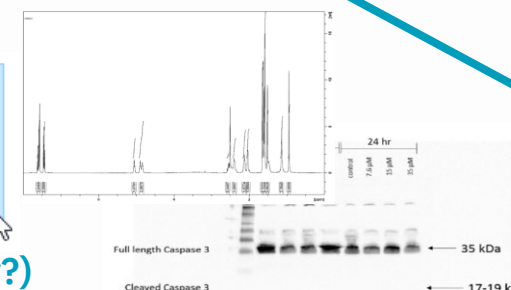
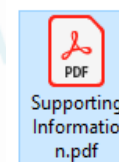
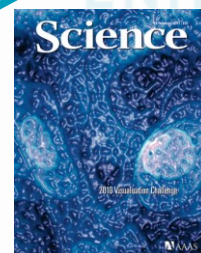
Analogue / Digital



Digital

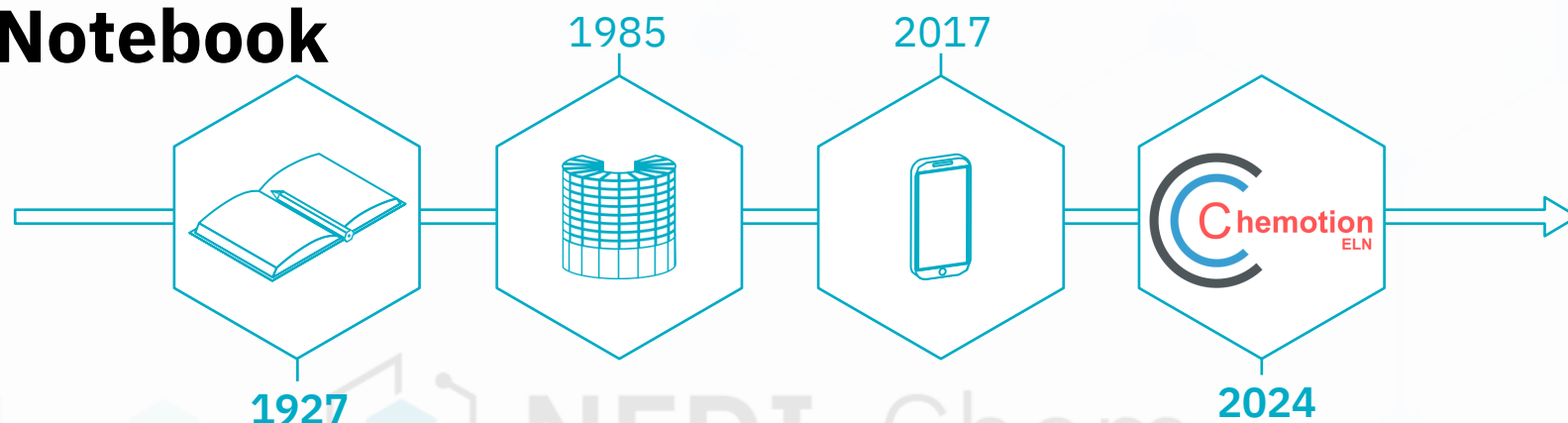


Findable
Accessible
Interoperable
Reusable



Chemotion

- An **Electronic Lab Notebook** for the Future

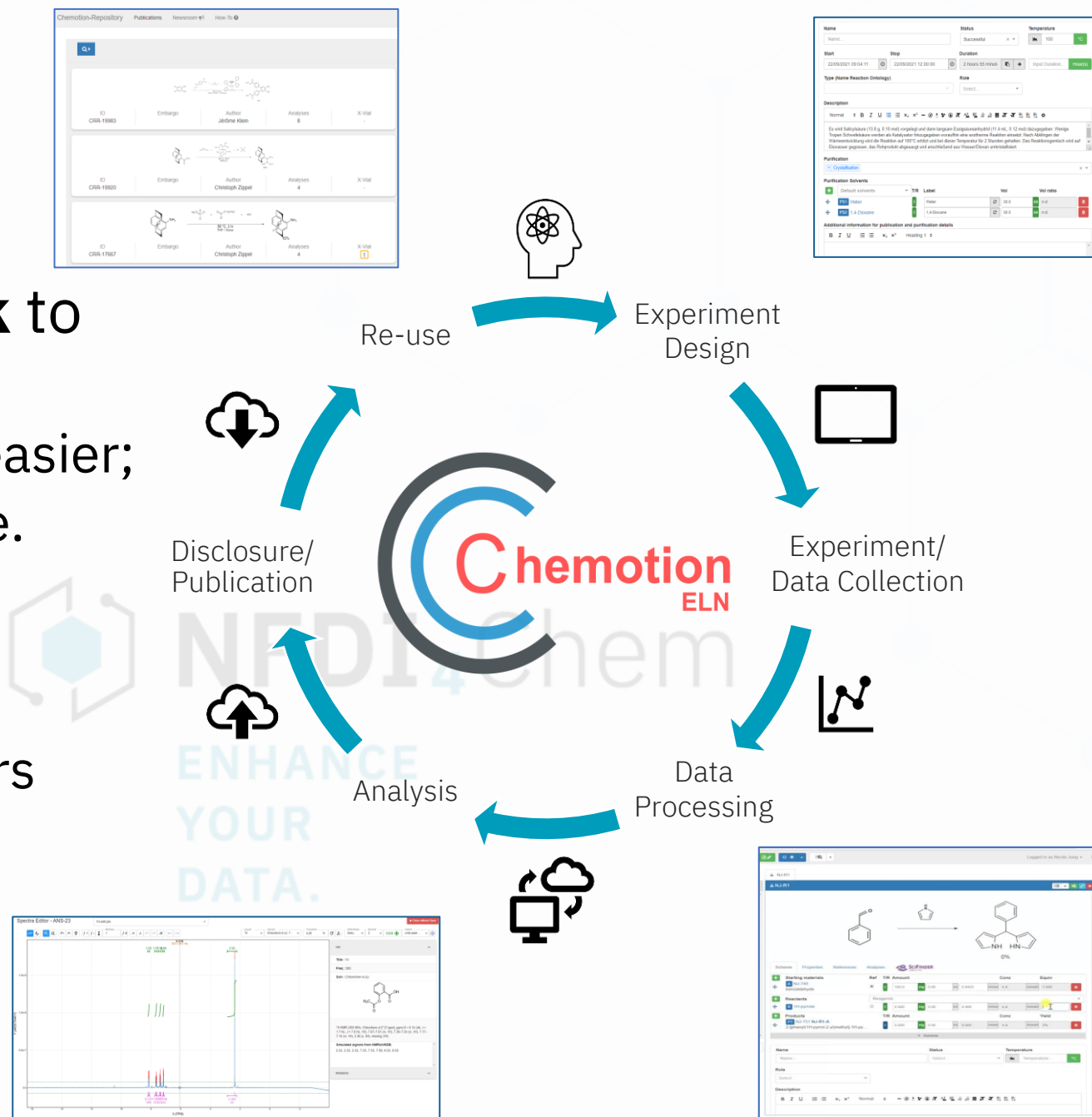


NFDI₄Chem

ENHANCE
YOUR
DATA.

Chemotion

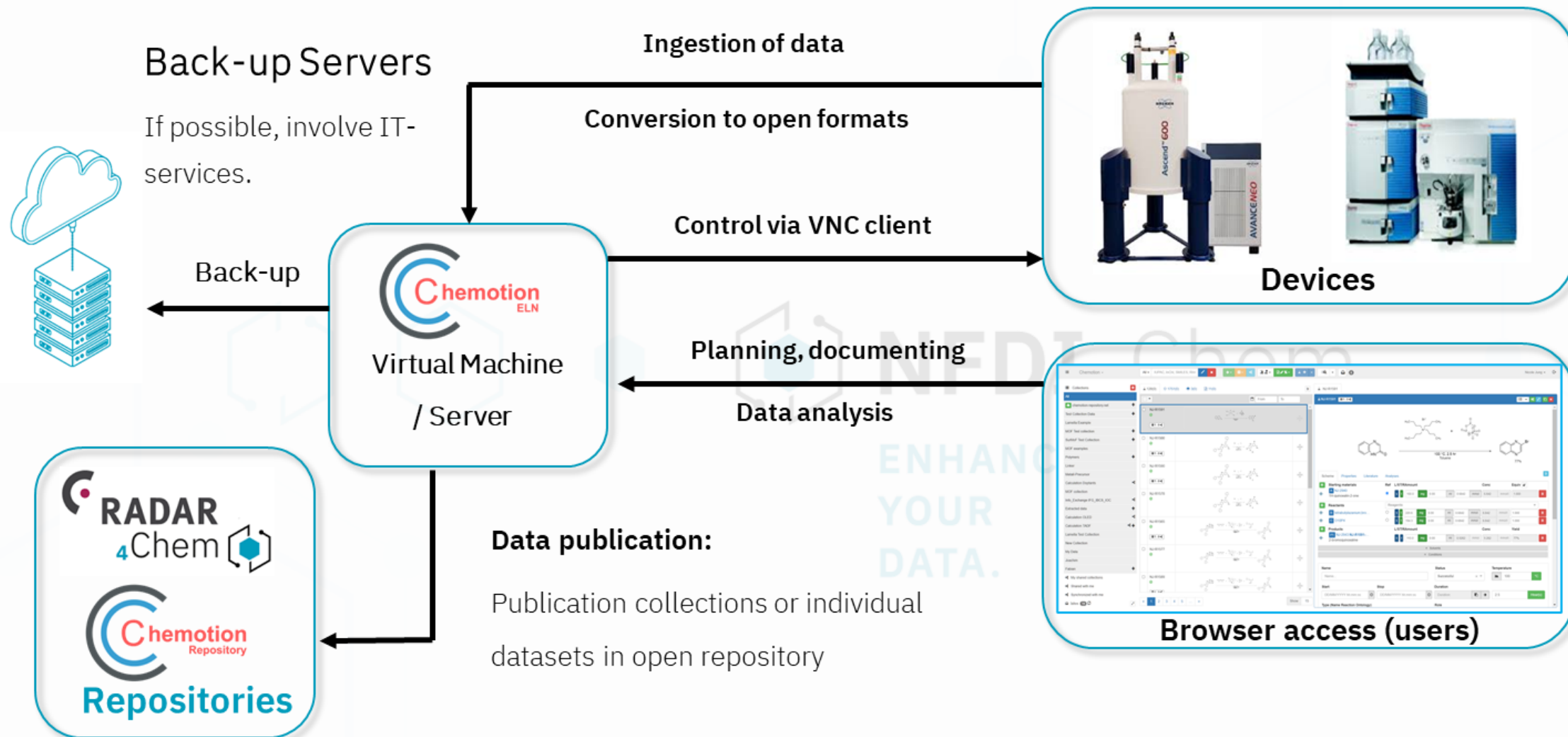
- An **Electronic Lab Notebook** to
 - Make data FAIR and
 - Research Data Management easier;
 - built on Open Source software.
- USPs:
 - By researchers, for researchers
 - Integrates with repositories
 - Promotes common standards



Timeline: 2012 vs. Present

- Idea:
 - ELN + Repo – common codebase
 - Implementation
 - Ruby on Rails
 - Postgres
 - Web frontend
 - Funding
 - Karlsruhe Institute of Technology
 - State of Baden-Württemberg
 - German Research Foundation (DFG)
 - Personnel
 - 3 FTEs
 - 2 PhD students (part-time)
- Towards:
 - Modularity, with API
 - Microservices, for dedicated tasks
 - Funding
 - Mainly DFG through NFDI4Chem
 - Plus, all on the left
 - Personnel (only counting developers)
 - 3-4 FTEs @ IBCS @ KIT
 - 2-3 FTEs @ SCC @ KIT
 - 4 part-time @ other collaborations
 - 1 FTE + 1 PhD for ML
 - External developers

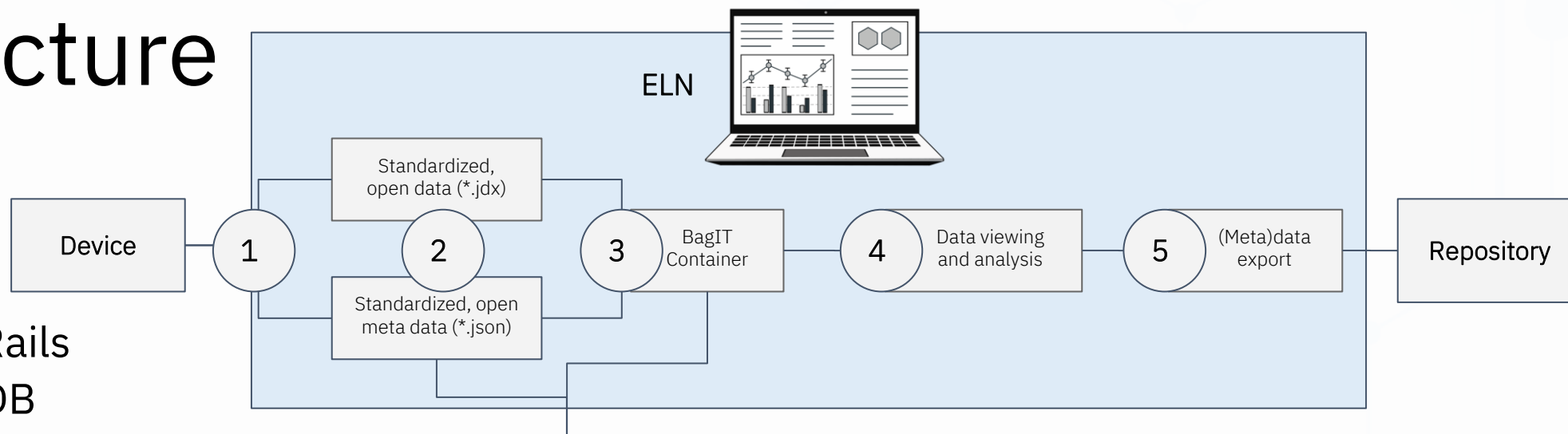
Usual Setup



Architecture

- Core

- ReactJS
- Ruby on Rails
- PostgresDB



e.g. ChemConverter

- Services

- ChemConverter (file format converter)
 - Python
- ChemSpectra (spectra viewer)
 - Python & JS

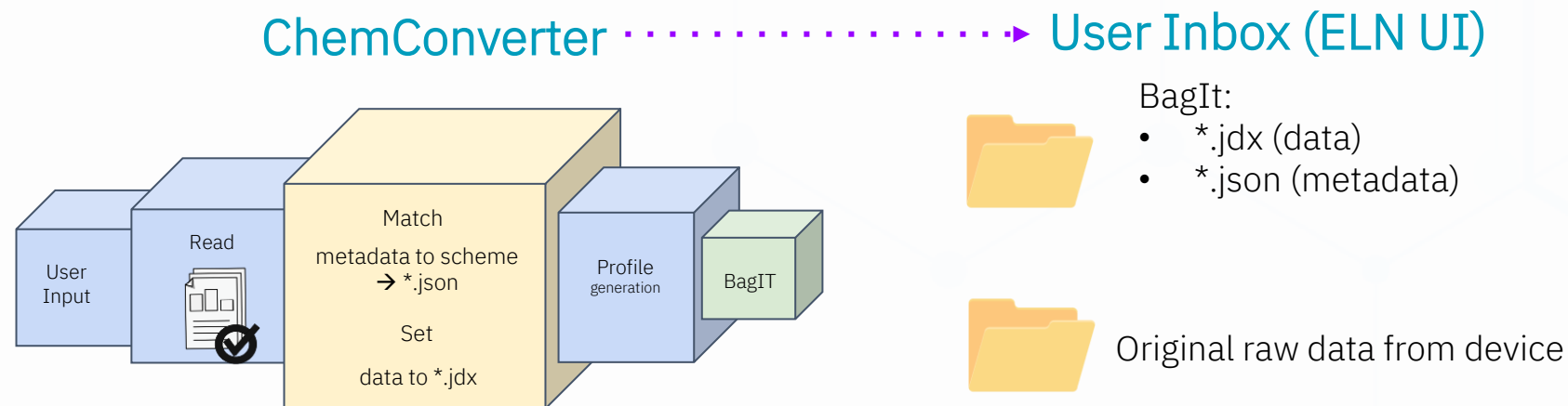
- Plugs into external services and devices

- NMRium (spectra viewer)
- Ketcher & ChemDrawJS (chemical structure editors)
- And many more

NFDI4Chem
ENHANCE
YOUR
DATA.

Architecture

- Core
 - ReactJS
 - Ruby on Rails
 - PostgresDB
- Services
 - ChemConverter (file format converter)
 - Python
 - ChemSpectra (spectra viewer)
 - Python & JS
- Plugs into external services and devices
 - NMRium (spectra viewer)
 - Ketcher & ChemDrawJS (chemical structure editors)
 - And many more

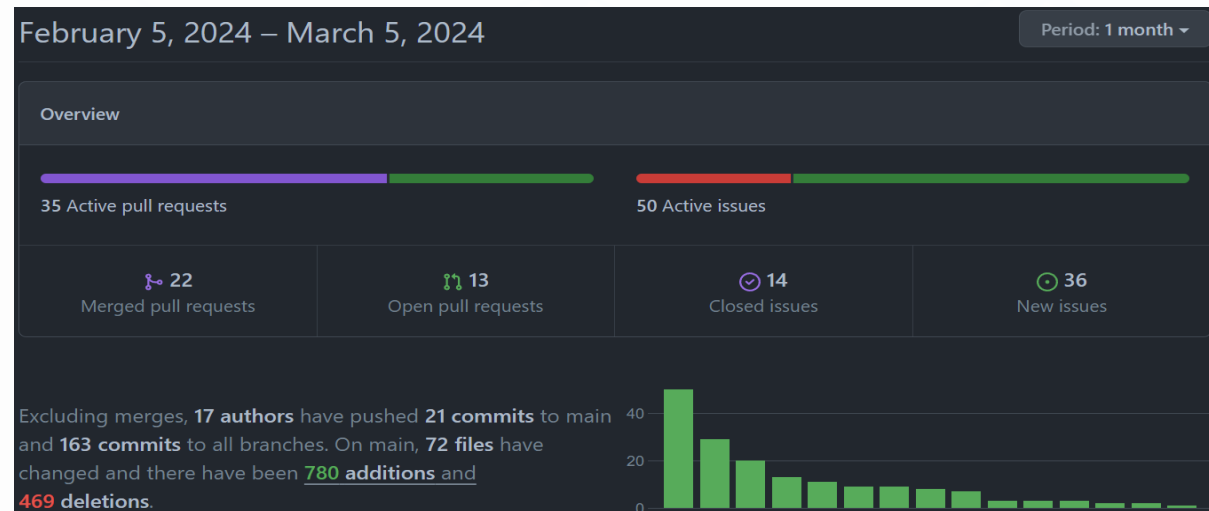


NFDI4Chem

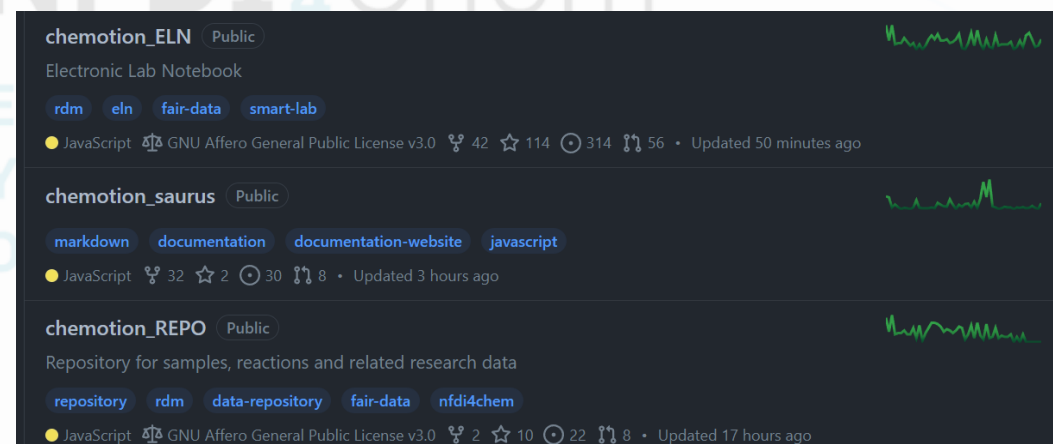
ENHANCE
YOUR
DATA.

Development

- Mostly in-house
 - Dedicated and experienced team
 - Long-term decision-making
 - Shifting priorities and refactoring
- Outsourcing – Why?
 - Bring flexibility, *can* work with unclear requirements
 - Provide feedback, sometimes even critical
 - **Lesson:** in-house team is valuable
 - Have dedicated expertise, e.g. Mobile App
- Effective management
 - Partitioning of project
 - Product ownership
 - Stability over speed



github.com/ComPlat/chemotion_ELN



github.com/orgs/ComPlat/repositories

Deployment

- Phased-out “native installations”
- Varying levels of expertise
 - Role of IT services
- Containerized
 - Docker images
 - `docker-compose.yml`
 - CLI tool (Go)
- Hand-holding during setup
 - Can be complicated
 - Networking & backup
 - Firewall and domain name setup
 - Networked storage
 - Access to external services

```
version: "3.9"
services:
  db:
    image: postgres:13
    restart: unless-stopped
    hostname: db
    environment:
      - POSTGRES_USER=postgres
      - POSTGRES_PASSWORD=postgres
    volumes:
      - chemotion_db:/var/lib/postgresql/data/
    networks:
      - chemotion

  msconvert:
    image: ptrxyz/chemotion:msconvert-1.8.2
    restart: unless-stopped
    hostname: msconvert
    volumes:
      - spectra:/shared:rw
    networks:
      - chemotion

  spectra:
    image: ptrxyz/chemotion:spectra-1.8.2
    restart: unless-stopped
    hostname: spectra
    volumes:
      - spectra:/shared:rw
    depends_on:
      - msconvert
    networks:
      - chemotion
```

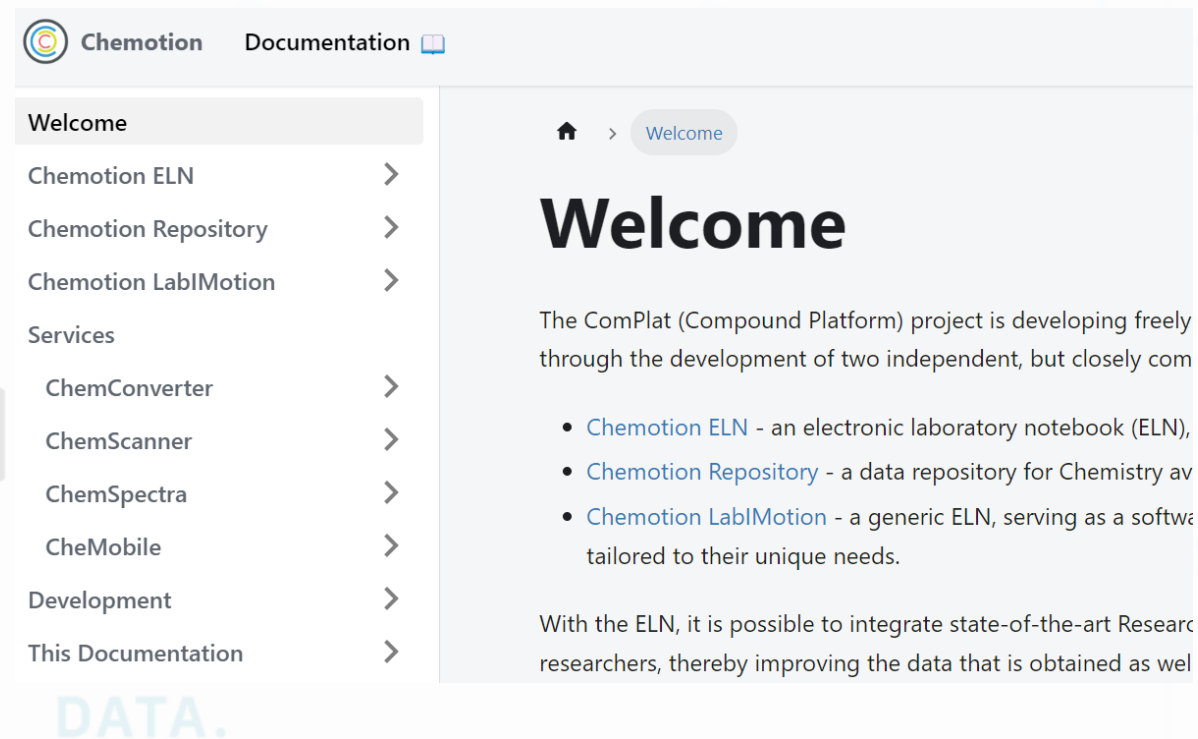
```
  eln:
    image: ptrxyz/chemotion:eln-1.8.2
    restart: unless-stopped
    depends_on:
      - db
      - spectra
    volumes:
      - ./shared/pullin:/shared
      - ./shared/backup:/backup
      - chemotion_data:/chemotion/data/
      - chemotion:/chemotion/app
    ports:
      - 4000:4000
    networks:
      - chemotion

  ketchersvc:
    image: ptrxyz/chemotion:ketchersvc-1.8.2
    restart: unless-stopped
    environment:
      - CONFIG_KETCHER_URL=http://eln:4000/ketcher/
    networks:
      - chemotion

  converter:
    image: ptrxyz/chemotion:converter-1.8.2
    restart: unless-stopped
    environment:
      - SECRET_KEY=imsupersecretandwanttobechanged
    volumes:
      - ./services/converter/profiles:/srv/chemotion/profiles
      - ./services/converter/datasets:/srv/chemotion/datasets
    networks:
      - chemotion
```

Helpdesk, in addition to extensive documentation

- One-stop solution
 - Intro / Demo / Workshop requests
 - Testing instances
 - Installation, deployment and Maintenance
 - Errors and bugs
 - Feature requests and feedback
- Not possible over email
 - However users *want* to email us!



Helpdesk, in addition to extensive documentation

- One-stop solution
 - Intro / Demo / Workshop requests
 - Testing instances
 - Installation, deployment and Maintenance
 - Errors and bugs
 - Feature requests and feedback
- Ticketing System
 - Zammad
 - 2 FTEs working part-time
 - ca. 300 closed tickets in 1.5 years
 - Bilingual support
 - Internal filtering and forwarding
 - Response time of 0-7 days, depending on complexity

Closed Tickets	525
Tickets Chemotion ELN	25
Tickets First Level Support	2
Tickets LabIMotion	0
My assigned Tickets	3
Unassigned & Open	14
My pending reached Tickets	0
Open	26
Pending reached	0
Escalated	0

👁️ Spam →

Closed Tickets		
1	2	3
4		
<input type="checkbox"/>	#	TITLE
<input type="checkbox"/>	50639	Umzug Chemotion data [REDACTED]
<input type="checkbox"/>	50638	[REDACTED]
<input type="checkbox"/>	50637	Test-Instanz AG [REDACTED]
<input type="checkbox"/>	50636	Chemotion
<input type="checkbox"/>	50635	RE: Chemotion
<input type="checkbox"/>	50634	chemotion - fresh eln startup ...
<input type="checkbox"/>	50633	Concern
<input type="checkbox"/>	50632	[REDACTED]
<input type="checkbox"/>	50631	We're Funding Business Ideas...
<input type="checkbox"/>	50629	Chemotion and ELN
<input type="checkbox"/>	50627	Electronic Notebook
<input type="checkbox"/>	50623	Problems using chemCLI and ...
<input type="checkbox"/>	50620	Problems using chemCLI and ...
<input type="checkbox"/>	50617	Contact Team Chemotion
<input type="checkbox"/>	50616	Problem mit Docker Installation
<input type="checkbox"/>	50615	Question about Chemotion Te...
<input type="checkbox"/>	50614	FW: Fragen zu Chemotion
<input type="checkbox"/>	50613	Integration von NMRium in Ch...

Challenges

- Hiring
 - Competition with Industry
 - Project-based funding can seem unstable
 - International, diverse
 - open and English-speaking
 - Looking for skills
 - Certifications as well as experiences

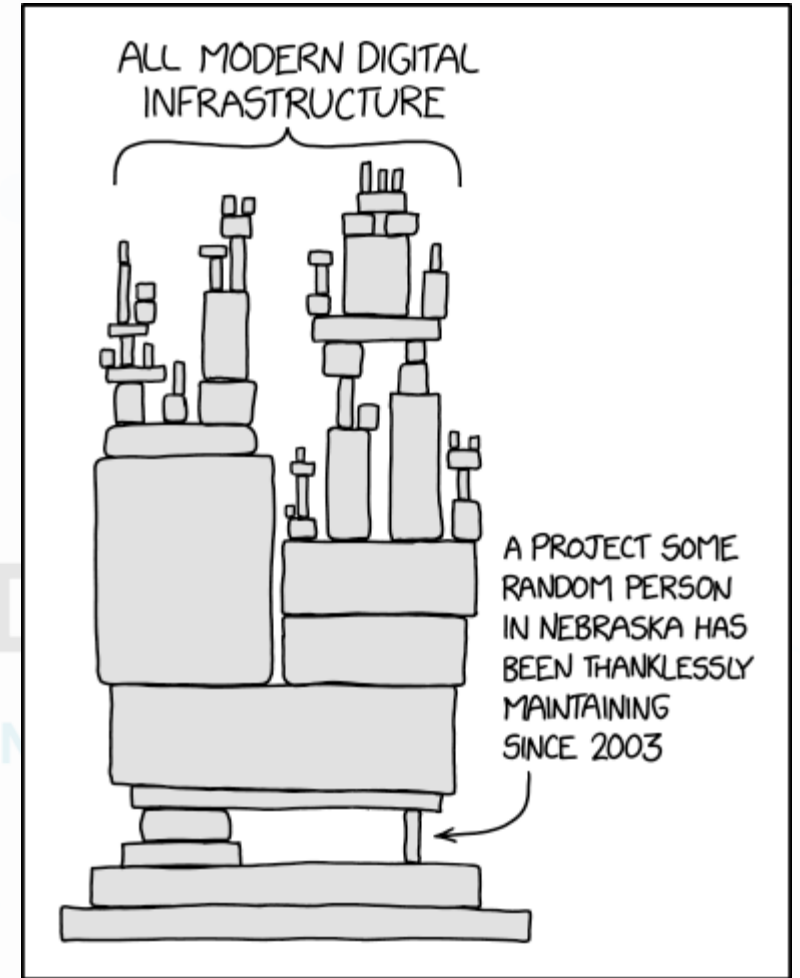


Dilbert.com DilbertCartoonist@gmail.com

10-1-12 ©2012 Scott Adams, Inc./Dist. by Universal Uclick

Challenges

- Hiring
 - Competition with Industry
 - Project-based funding can seem unstable
 - International, diverse
 - open and English-speaking
 - Looking for skills
 - Certifications as well as experiences
- Funding
 - Guaranteed basic funding (from KIT)
 - DFG projects: 3 to 5 years
 - Long-duration is preferred
 - Constant effort



xkcd.com/2347

State of the App

- 2012: We were lucky!
 - Time was ripe: RDM, resources & automation
 - Had critical mass (40 users)

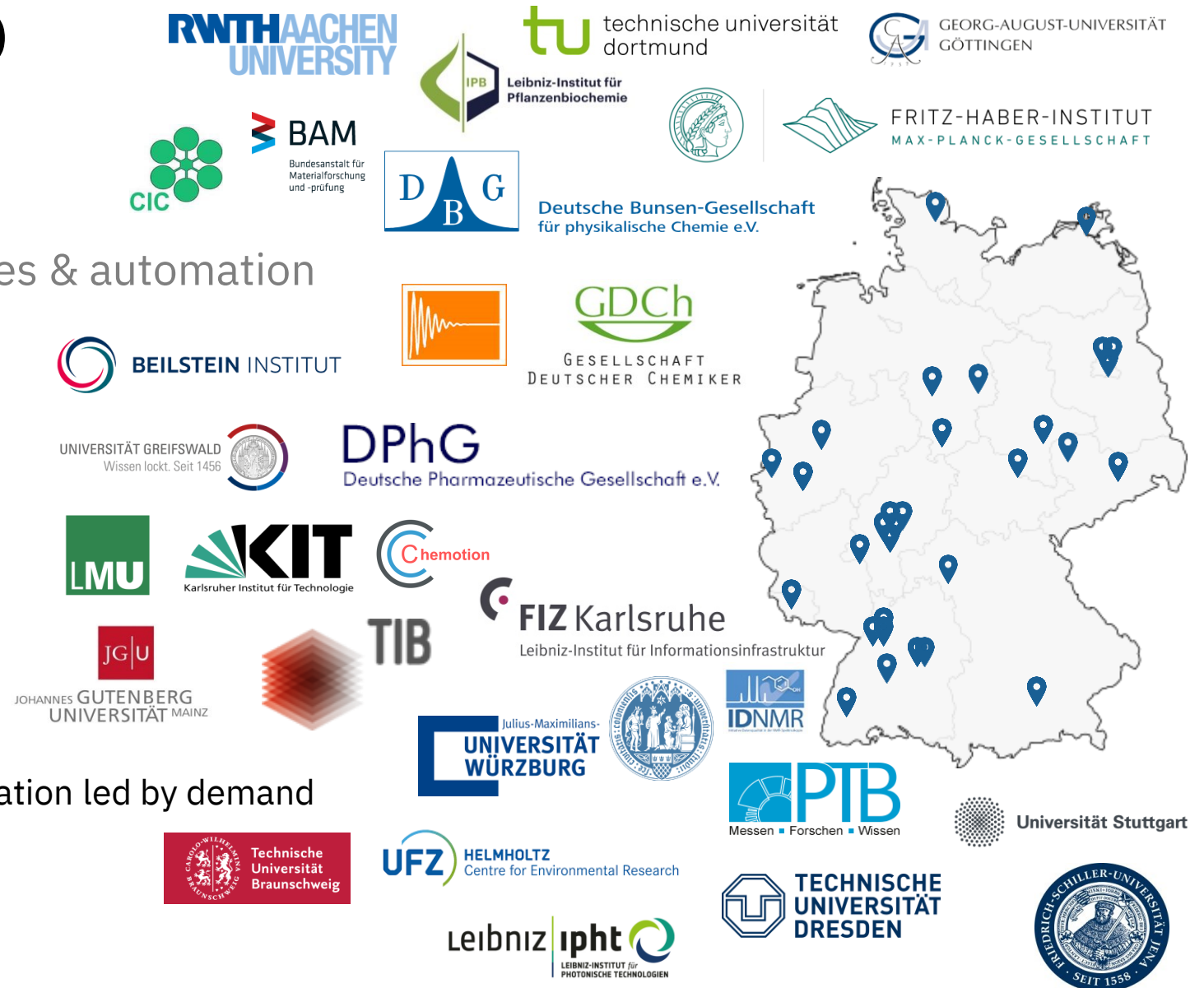


NFDI₄Chem

ENHANCE
YOUR
DATA.

State of the App

- 2012: We were lucky!
 - Time was ripe: RDM, resources & automation
 - Had critical mass (40 users)
- We strive to remain lucky:
 - In it for the long run
 - Building trust
 - university partners
 - collaborative
 - Constant improvements
 - Very feedback driven
 - “Customer satisfaction”: Innovation led by demand



State of the App

- 2012: We were lucky!
 - Time was ripe: RDM, resources & automation
 - Had critical mass (40 users)
- We strive to remain lucky:
 - In it for the long run
 - Building trust
 - university partners
 - collaborative
 - Constant improvements
 - Very feedback driven
 - “Customer satisfaction”: Innovation led by demand
- To “Spin-off” or “not-to-Spin-off”
 - Funding vs.
 - Trust and support of the Academia

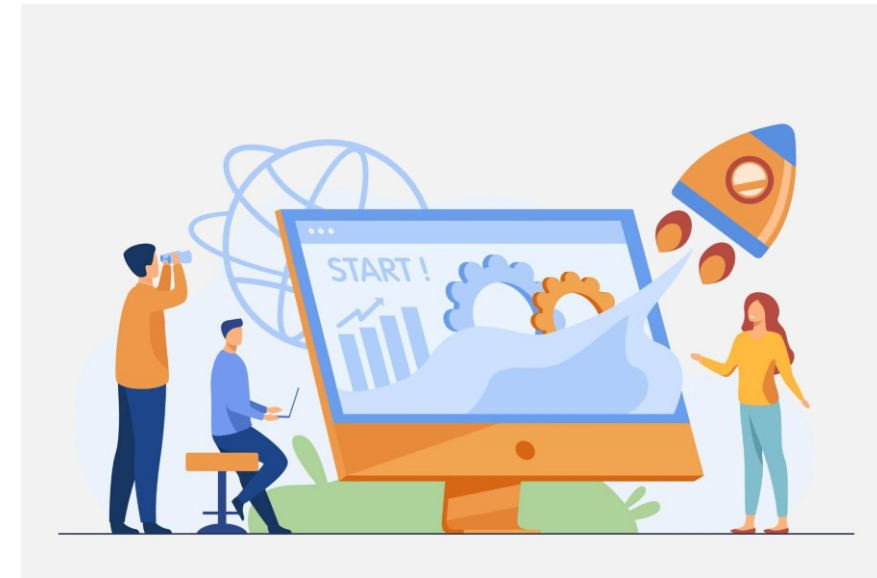


Image from: tradebrains.in/corporate-spin-offs

Thanks!

- To you
- To Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under the National Research Data Infrastructure – NFDI4Chem project – project number 441958208.
- To all our partners, collaborators, contributors and users ☺