Research Data Management with LinkAhead

Shared RDM Event Series "Research Data Management in Austria"

Timm Fitschen

IndiScale GmbH

9 Jan 2024

LinkAhead
Thinking Data Management Ahead

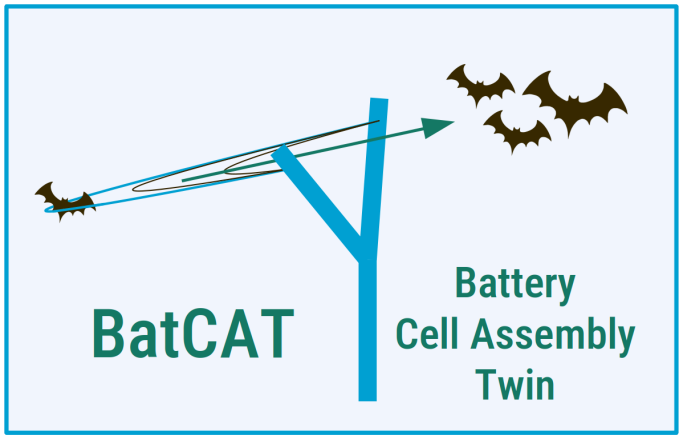
IndiScale
Data Services

BatCAT

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101137725.

Funded by the European Union

Battery Cell Assembly Twin (BatCAT)



Timeframe: 1 Jan 2024 – 30 Jun 2027.

Goal

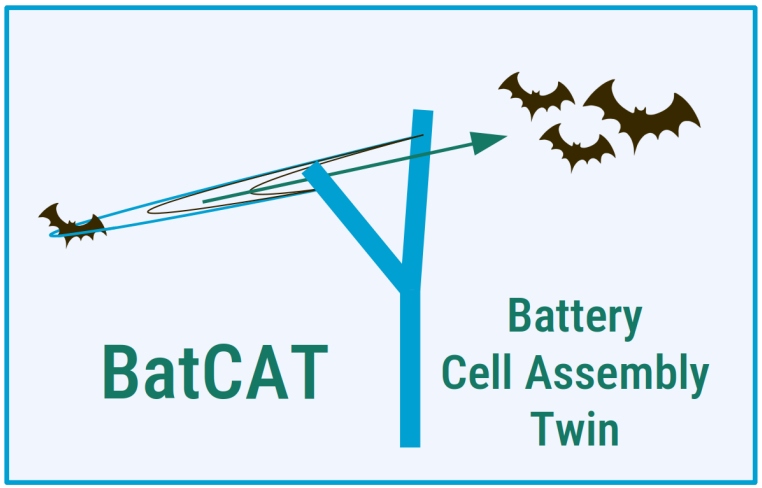
Digital twin for battery manufacturing *towards new greener and more sustainable manufacturing processes*.

K. Edström et al., "Inventing the sustainable batteries of the future" (BATTERY 2030+ Roadmap), 2022.

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101137725.

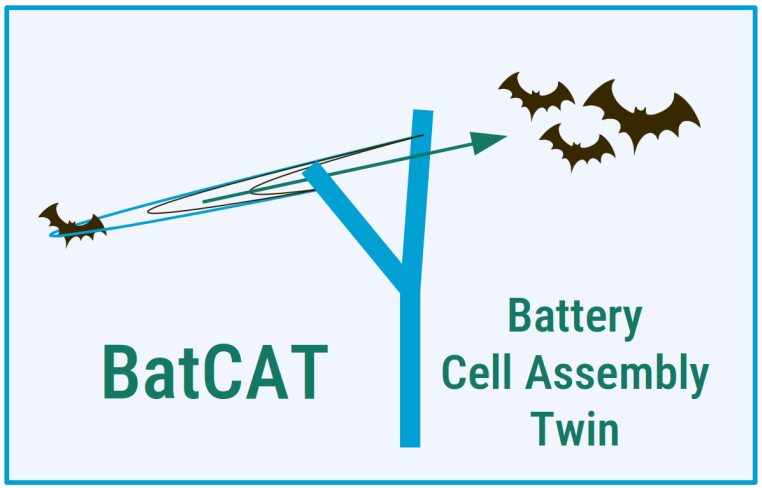
Battery Cell Assembly Twin (BatCAT)



Partners

- *Norwegian University of Life Sciences (Norway)*
- Fraunhofer Institute for Industrial Mathematics ITWM (Germany)
- Kemijski Institut (Slovenia)
- Politecnico di Torino (Italy)
- Rheinland-Pfalzische Technische Universitat (Germany)
- IFP Energies Nouvelles (France)
- *Universitaet Klagenfurt (Austria)*
- Danmarks Tekniske Universitet (Denmark)
- Fundacion Universidad Loyola Andalucia (Spain)
- Hochschule Kaiserslautern (Germany)
- Simula Research Laboratory AS (Norway)
- BI-REX - Big Data Innovation Research Excellence (Italy)
- Vanevo GmbH (Germany)
- Luxembourg Institute of Science and Technology (Luxembourg)
- United Kingdom Research and Innovation (United Kingdom)
- Centre for Process Innovation Limited LBG (United Kingdom)
- Goldbeck Consulting Limited (United Kingdom)
- *IndiScale GmbH (Germany)*

Battery Cell Assembly Twin (BatCAT)

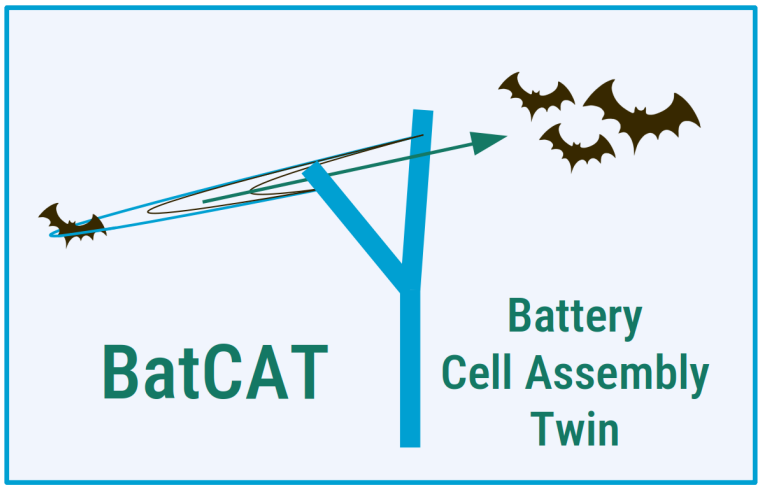


IndiScale
Data Services

- IndiScale: Data management services using LinkAhead.
- Develop LinkAhead as *federated hierarchical knowledgebase* [Horsch 2022]
- Create iteroperable data space accross national and institutional boundaries.

M. T. Horsch et al., in Proc. DAMDID 2021, pp. 166–177, doi:10.1007/978-3-031-12285-9_10, 2022.

Battery Cell Assembly Twin (BatCAT)



LinkAhead
Thinking Data Management Ahead

- Integrate data from experiments, simulations, analyses.
- Feed machine learning algorithms and AI training.
- Share data between partners and *secure sensitive data, intellectual property rights, regulatory and contractual obligations.* [Horsch 2022]

Research Data Management System

✓ Supports researchers in data management tasks

LinkAhead: A Research Data Management System

Research Data Management System

- ✓ Supports researchers in data management tasks
- ✓ according to applicable rules and policies

- § FAIR Guiding Principles
- § Research data policies
- § Standards, best-practices
- § Organisational, legal, individual rules

LinkAhead: A Research Data Management System

Research Data Management System

- ✓ Supports researchers in data management tasks
- ✓ according to applicable rules and policies
- ✓ taking into account the conditions

- ? Capacities
- ? Data structure, file formates und software
- ? Established workflows
- ? Legacy data resources

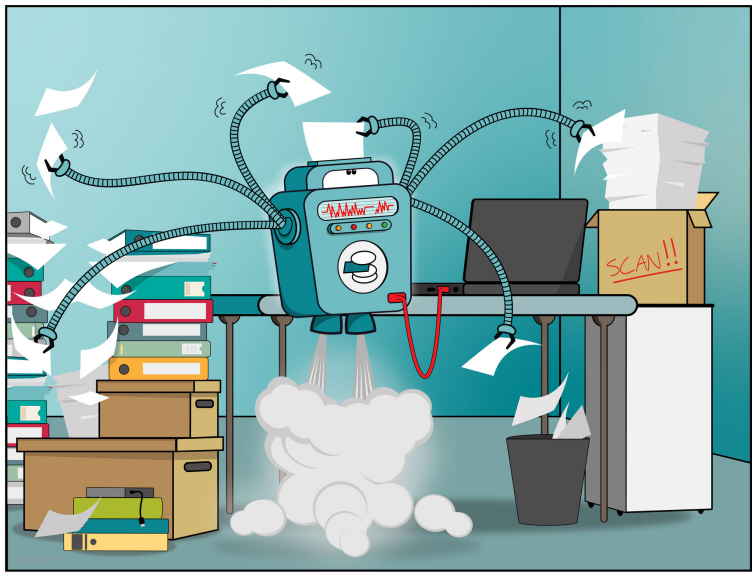
LinkAhead: Vision

- ✓ Flexible: Adaptale to different conditions
- ✓ Agile: Step-by-step change, improvement and growth
- ✓ Unobtrusive: researchers determine the workflow, the tools

LinkAhead: Vision

LinkAhead
Thinking Data Management Ahead

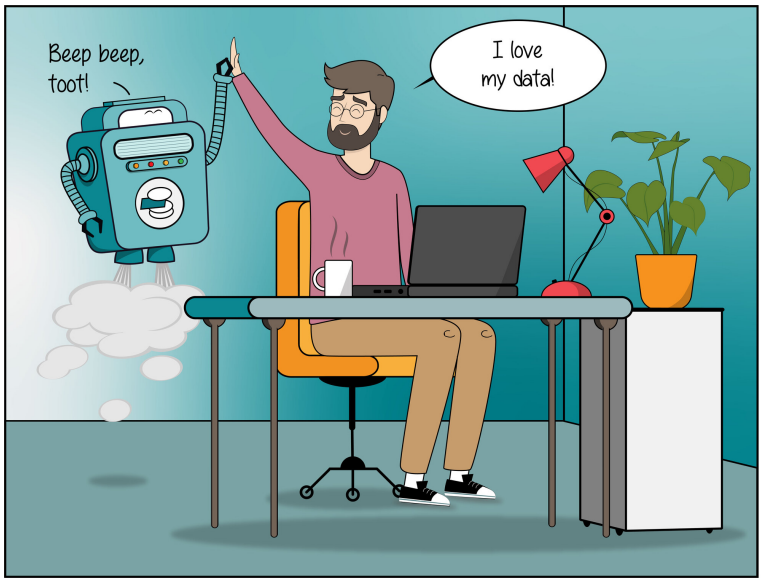
- ✓ Flexible: Adaptale to different conditions
- ✓ Agile: Step-by-step change, improvement and growth
- ✓ Unobtrusive: researchers determine the workflow, the tools



LinkAhead: Vision

LinkAhead
Thinking Data Management Ahead

- ✓ Flexible: Adaptale to different conditions
- ✓ Agile: Step-by-step change, improvement and growth
- ✓ Unobtrusive: researchers determine the workflow, the tools



LinkAhead: Looking Back

- ✓ LinkAhead Development since 2011 at MPI-DS (Göttingen) (Former name: CaosDB)
- ✓ Used in production since 2016
- ✓ Published as Open Source Software 2018
- ✓ Increasing prevalence since 2020
- ✓ Sample management, literature database, management of raw data, automation of data processing, laboratory automation, indexing and annotation of data lakes and repositories, ...

Caosdb
an open scientific database

LinkAhead: Looking Back

- ✓ LinkAhead Development since 2011 at MPI-DS (Göttingen) (Former name: CaosDB)
- ✓ Used in production since 2016
- ✓ Published as Open Source Software 2018
- ✓ Increasing prevalence since 2020
- ✓ Sample management, literature database, management of raw data, automation of data processing, laboratory automation, indexing and annotation of data lakes and repositories, ...

Caosdb
an open scientific database

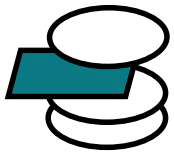
GitLab

LinkAhead: Looking Back

- ✓ LinkAhead Development since 2011 at MPI-DS (Göttingen) (Former name: CaosDB)
- ✓ Used in production since 2016
- ✓ Published as Open Source Software 2018
- ✓ Increasing prevalence since 2020
- ✓ Sample management, literature database, management of raw data, automation of data processing, laboratory automation, indexing and annotation of data lakes and repositories, ...

Caosdb
an open scientific database

GitLab

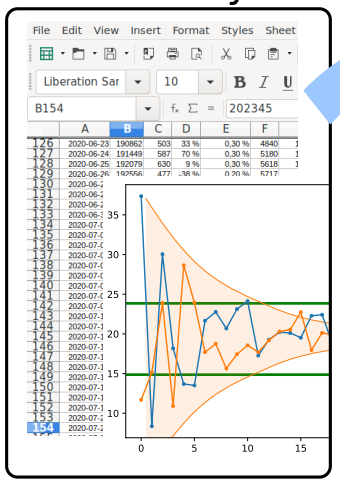


LinkAhead

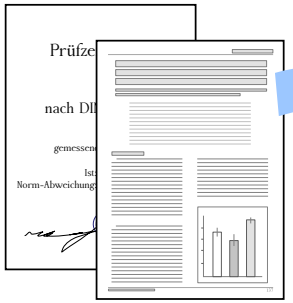
IndiScale
Data Services

LinkAhead in the Daily Research Routine

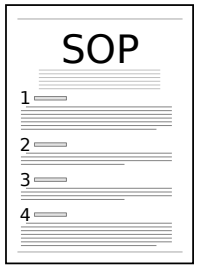
Processing,
Data Analysis



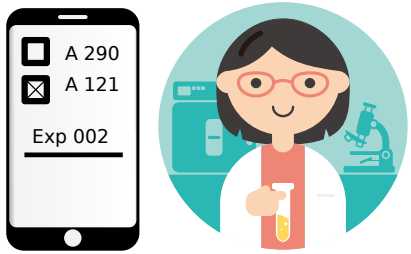
Publications



Ideas, Protocols,
Internal Documents,



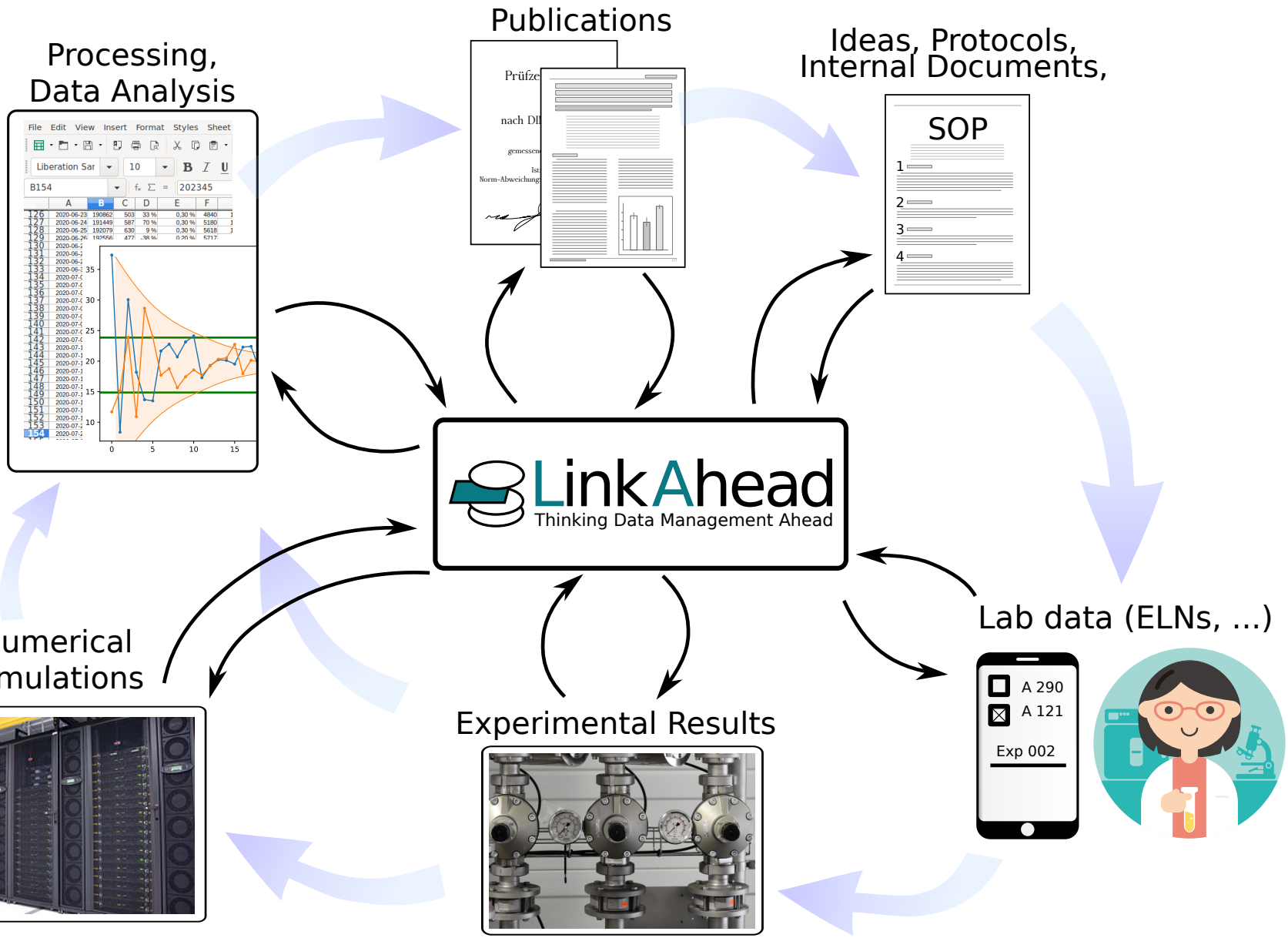
Lab data (ELNs, ...)



Experimental Results

Numerical
Simulations

LinkAhead in the Daily Research Routine



Agile Data Management

- ☛ Research happens before standardization.
- ☛ Agile: Start small, grow iteratively, adapt.

LinkAhead's Innovative Data Model

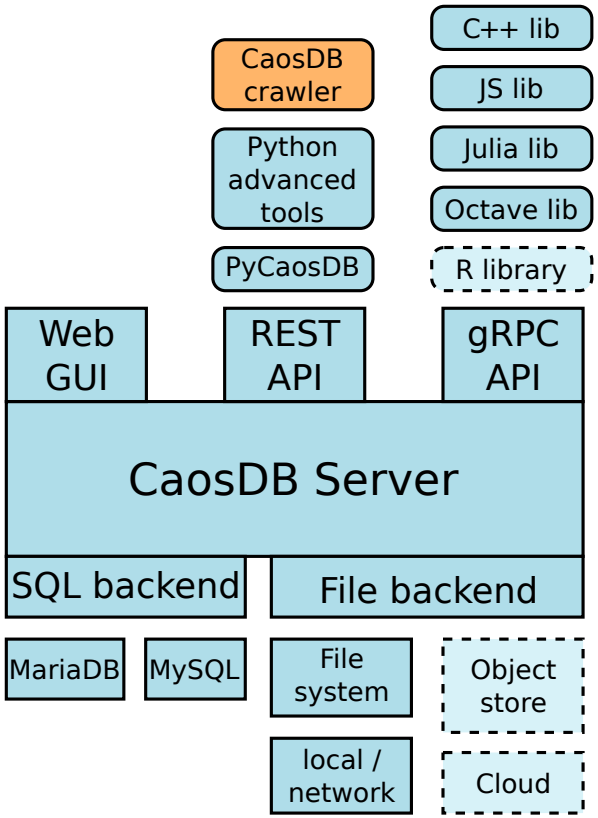
- ✓ No migrations necessary.
- ✓ Integrate legacy data AS IT IS.
- ✓ Future changes don't have to bother you now.

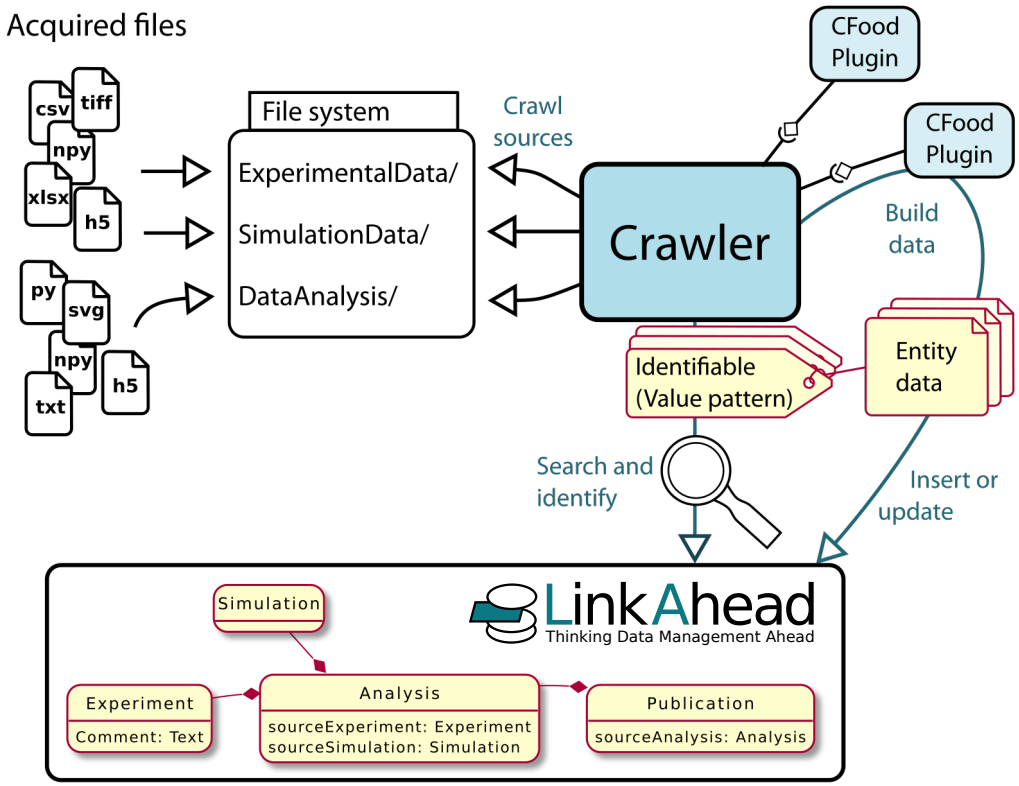
Features

LinkAhead

- ✓ Semantic data model und searching capabilities
- ✓ Schema/ontology defined by the users
- ✓ Web-based user interface
- ✓ Interaction with mit external file storages
- ✓ Server-side scripting
- ✓ Integrated ETL system
- ✓ User management with LDAP, NIS, and more
- ✓ Role-based access control
- ✓ Software libraries for many programming languages

Architecture: Overview





Integration of Data Analysis Software

LinkAhead

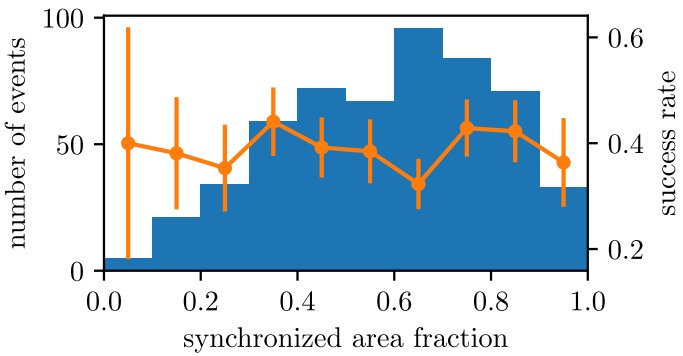
jupyter Introduction to the CaosDB-Python Client (autosaved)

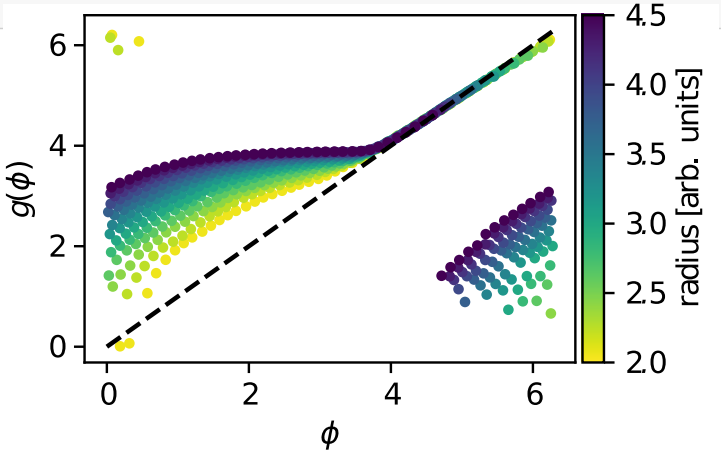
File Edit View Insert Cell Kernel Widgets Help Not Trusted

Run Markdown

Data Analysis with Jupyter

In [ ]:
```
import caosdb as db
data = db.execute_query("SELECT quality_factor FROM RECORD Analysis with quality_factor" )
table = to_table(data)
plt.plot(table.x, table.y)
```

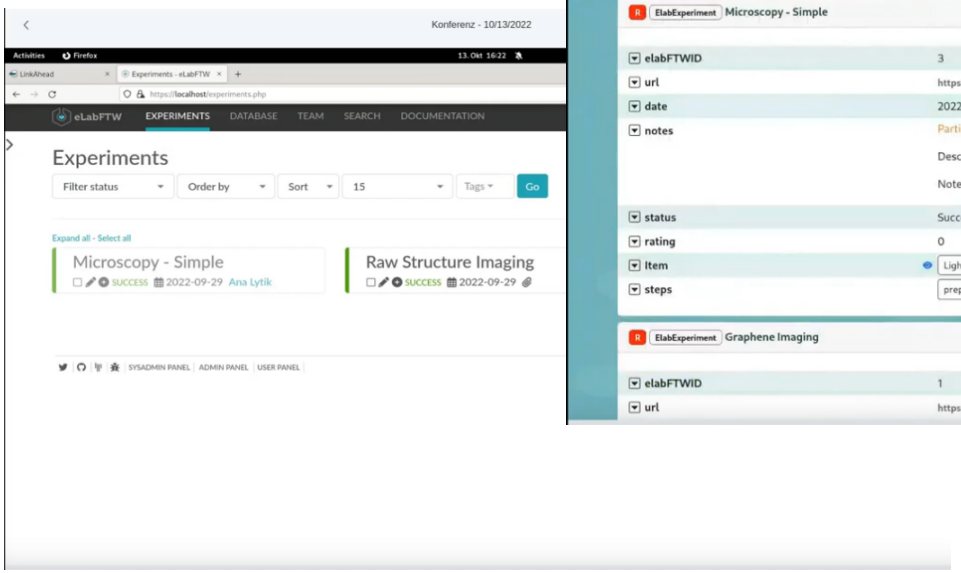


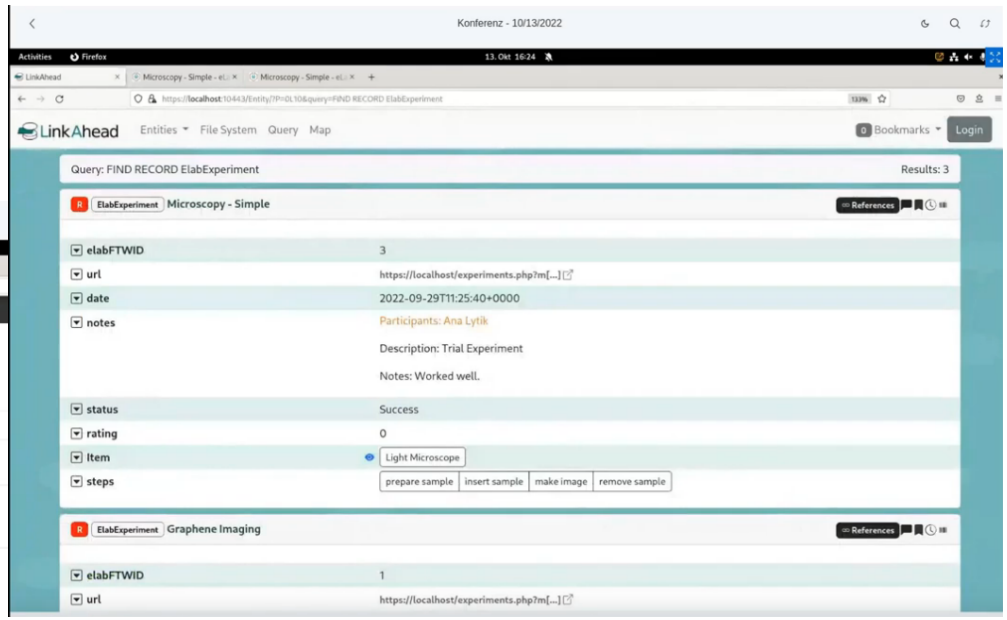


Integration of ELN

LinkAhead

eLabFTW





https://youtu.be/mJ4MSuryjEE

Demo

Live-Demo

https://demo.indiscale.com

Deployment and Licensing

- ✓ Deployment on-premise
- ✓ Cloud service
- ✓ All components are open source software (AGPLv3)
- ✓ Documentation is freely available online

Deployment and Licensing

- ✓ Deployment on-premise
- ✓ Cloud service
- ✓ All components are open source software (AGPLv3)
- ✓ Documentation is freely available online

Advertisement

IndiScale
Data Services

Consulting and service for deployment, customisation and development

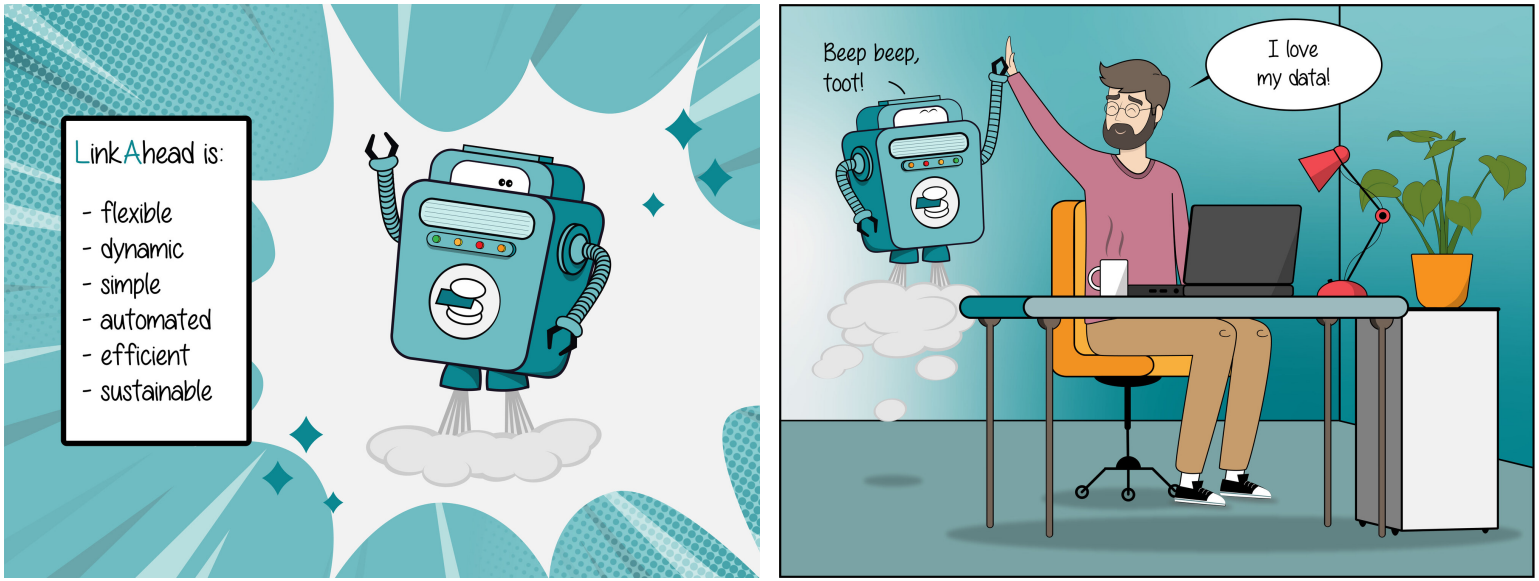
Current Activities

- ELN Consortium: https://github.com/TheELNConsortium
- Omero Integration: https://www.openmicroscopy.org/omero/
- FAIR Digital Objects: https://fairdo.org/
- Horizon Europe Consortium BatCAT
  https://www.nmbu.no/en/research/projects/batcat

Further Resources

- ✓ Web site: https://getlinkahead.com
- ✓ Online demo: https://demo.indiscale.com
- ✓ Documentation: https://docs.indiscale.com
- ✓ Software repository: https://gitlab.com/linkahead
- ✓ Fitschen et al. 2019 in ***Data***: https://doi.org/10.3390/data4020083
- ✓ Hornung et al. (accepted) in **ing.grid**:
  Agile Research Data Management with Open Source: LinkAhead.
  https://doi.org/10.48694/inggrid.3866

Summary



- ✓ Focus on researchers and their activities
- ✓ Flexible, agile and unobtrusive
- ✓ Integrates into established workflows
- ✓ Open source software