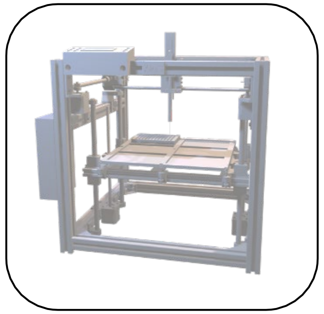
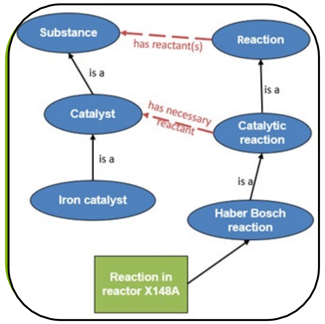
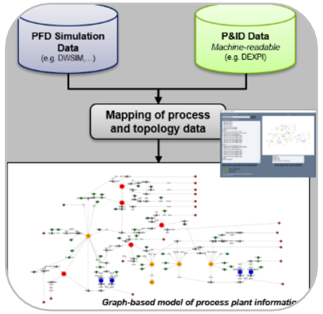
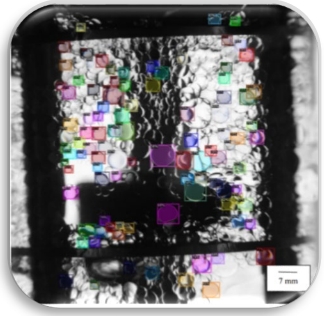
Data-Driven Laboratory - Lab Automation, Software Prototyping, and AI Modelling

R. Dinter[1], J. Oeing[1], L. Neuendorf[1], A. Behr[1], N. Kockmann[1]

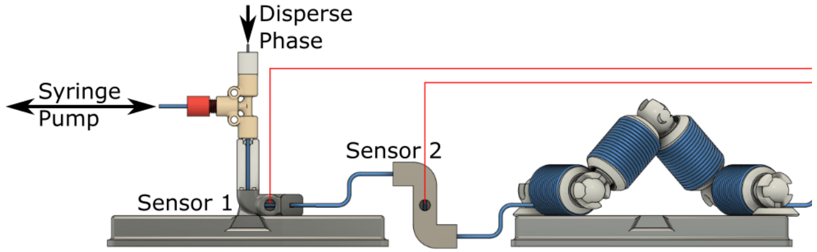
[1] *TU Dortmund University, Laboratory of Equipment Design, Dortmund/GER*

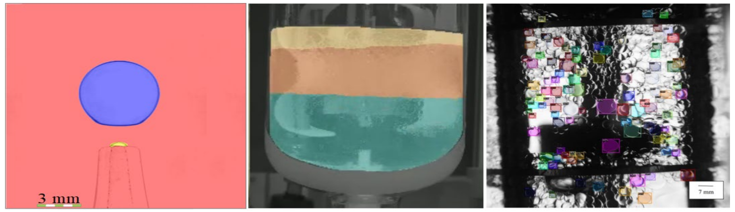



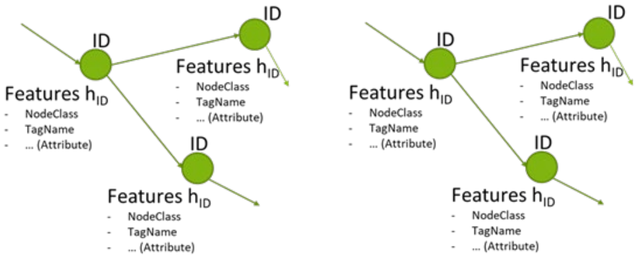



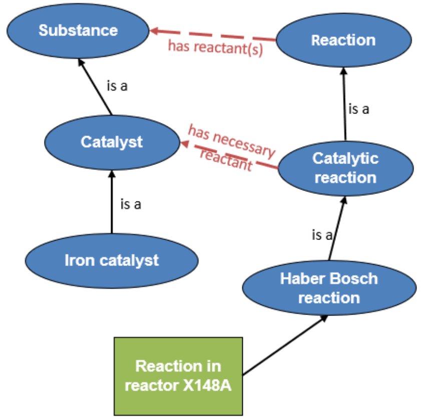



- Lab automation (Robin Dinter)
  - open-source hard-/software for an efficient fully integrated system
  - automated reaction screening strategy
- AI modelling (Laura Neuendorf)
  - AI-based optical sensors to automate evaluation of lab applications
- Process engineering (Jonas Oeing)
  - machine & deep learning
  - graph-based engineering design
- Research data management (Alex Behr)
  - ontology development & knowledge graphs



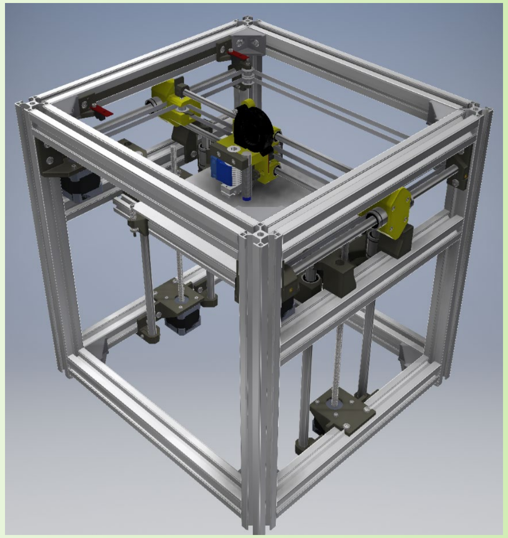


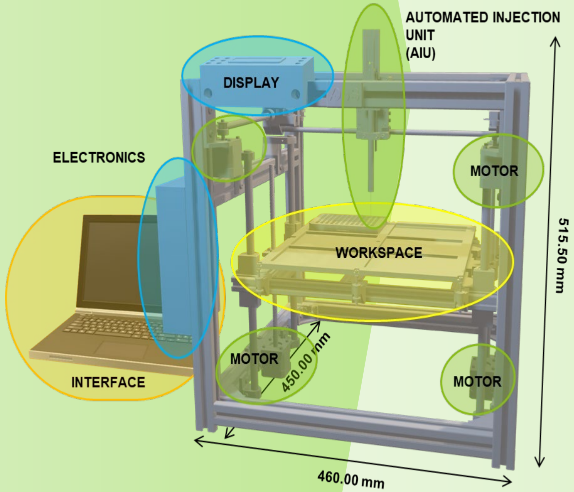


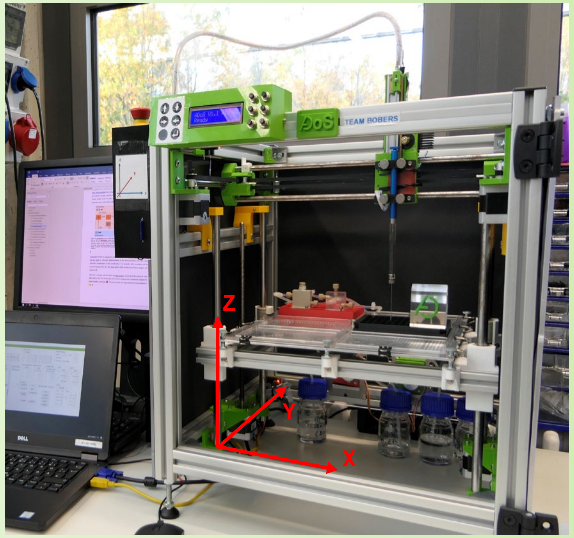


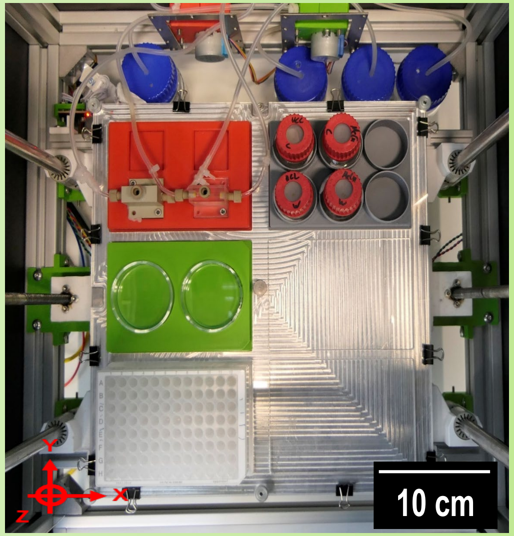
evolution of an automated dosage system

- open-source 3D-printer[1] as base
  - easy to build and low-cost
  - documentation available
  - controlled with microcontrollers ARDUINO
- mechanical design of the ADoS[2]
  - printer head substituted by GC syringe (50 µL volume)
  - modularized workspace in well plate format
  - customized interface for code generation python







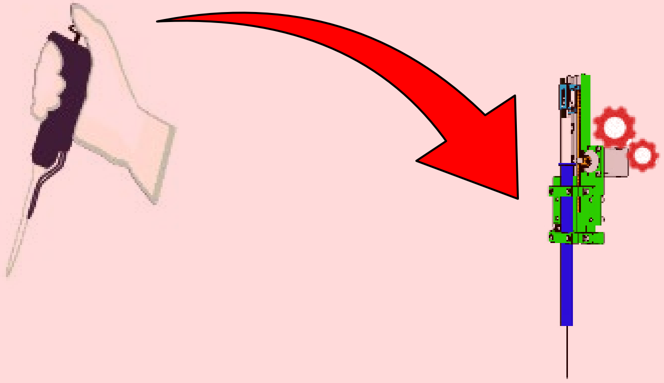


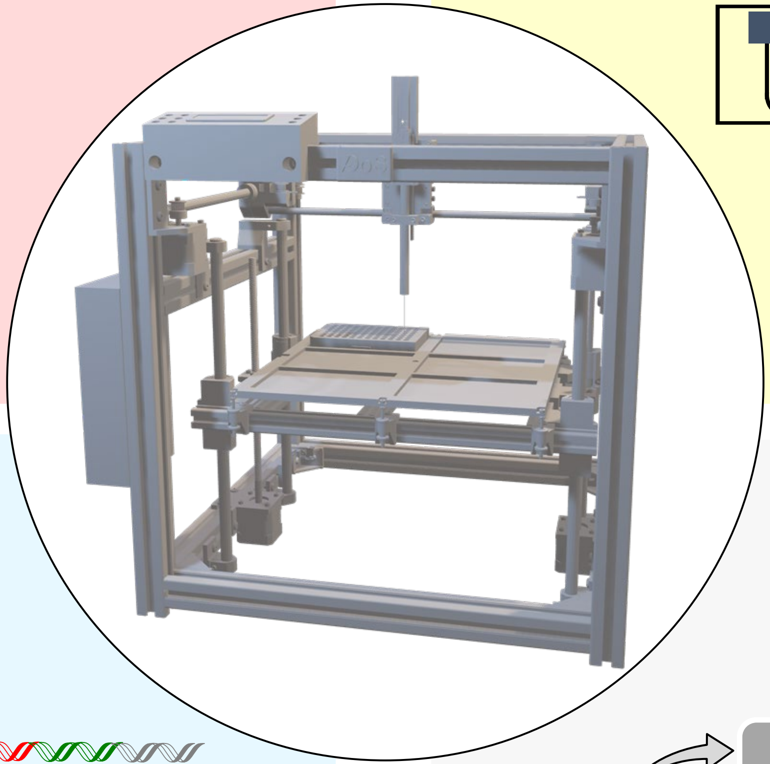
adaptation on specific lab applications by self-designed modules

[1] SCOTT_3D: https://www.thingiverse.com/thing:2254103 (accessed 06.09.2021)
[2] J. Bobers et al., ACS Comb. Sci., **2020**

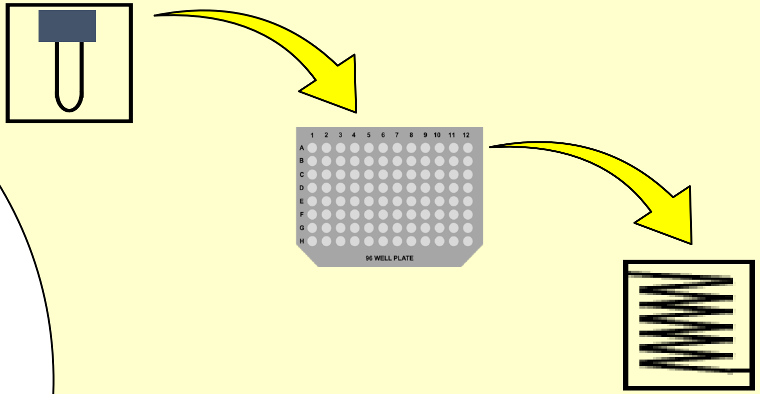
towards an efficient fully integrated system

- from manual to automated pipetting

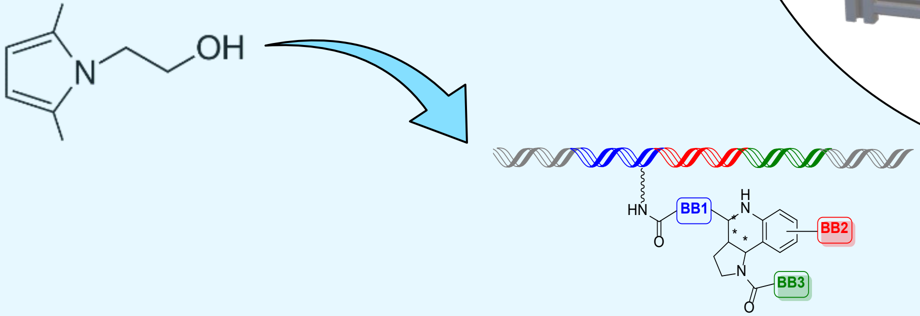




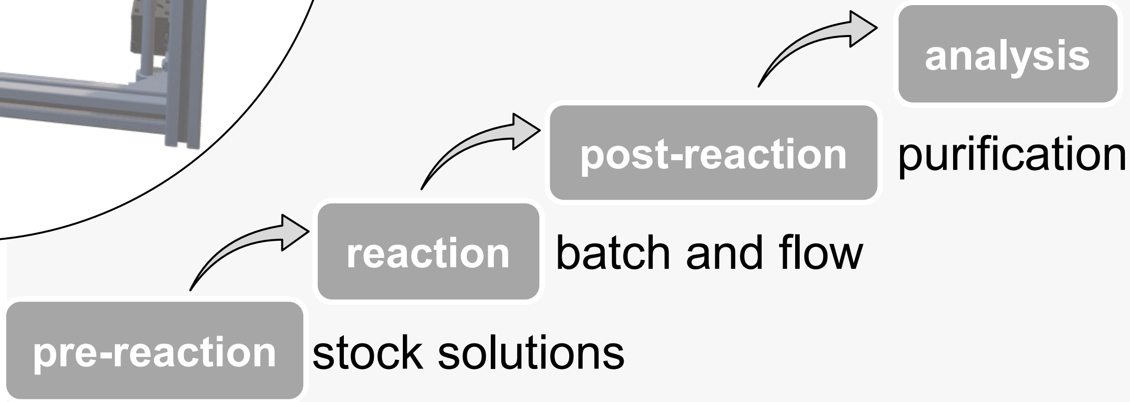
- from batch to well plate to continuous flow



- from model reaction to DNA-substrates for DEL

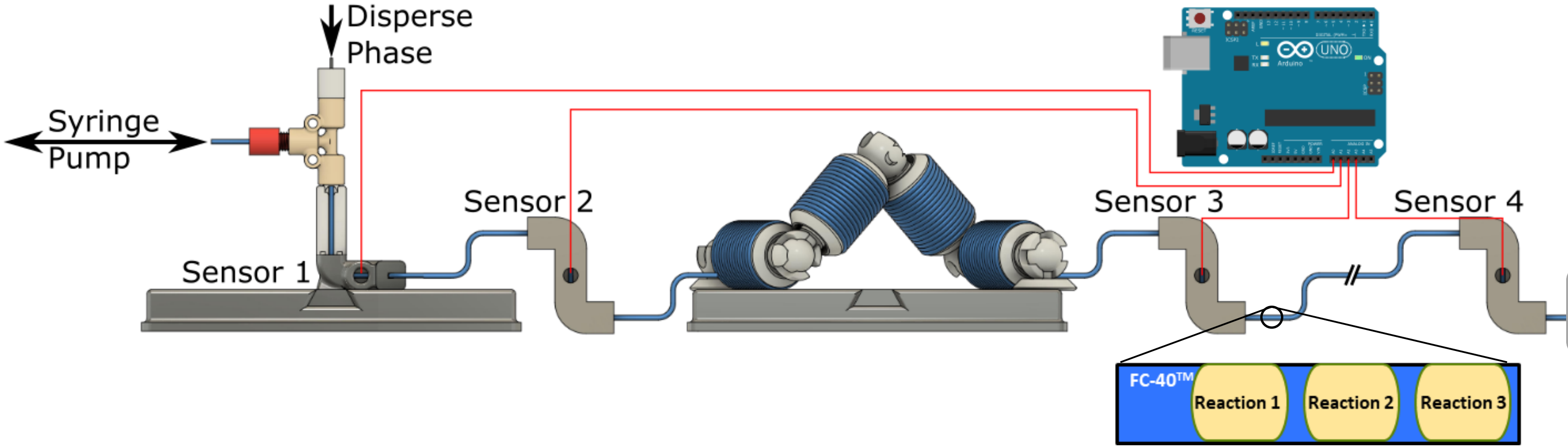


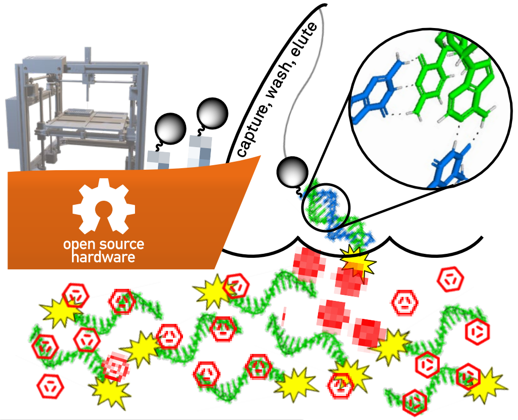
- increase level of automation



screening strategy for translation to DEL[1,2]

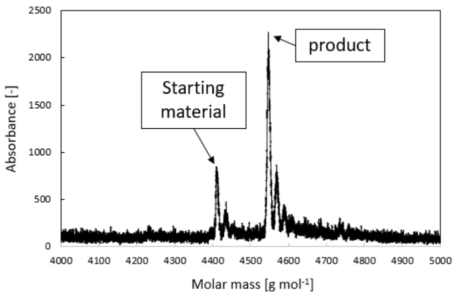
pre-reaction stock solutions





post-reaction purification

analysis MALDI-MS
HPLC



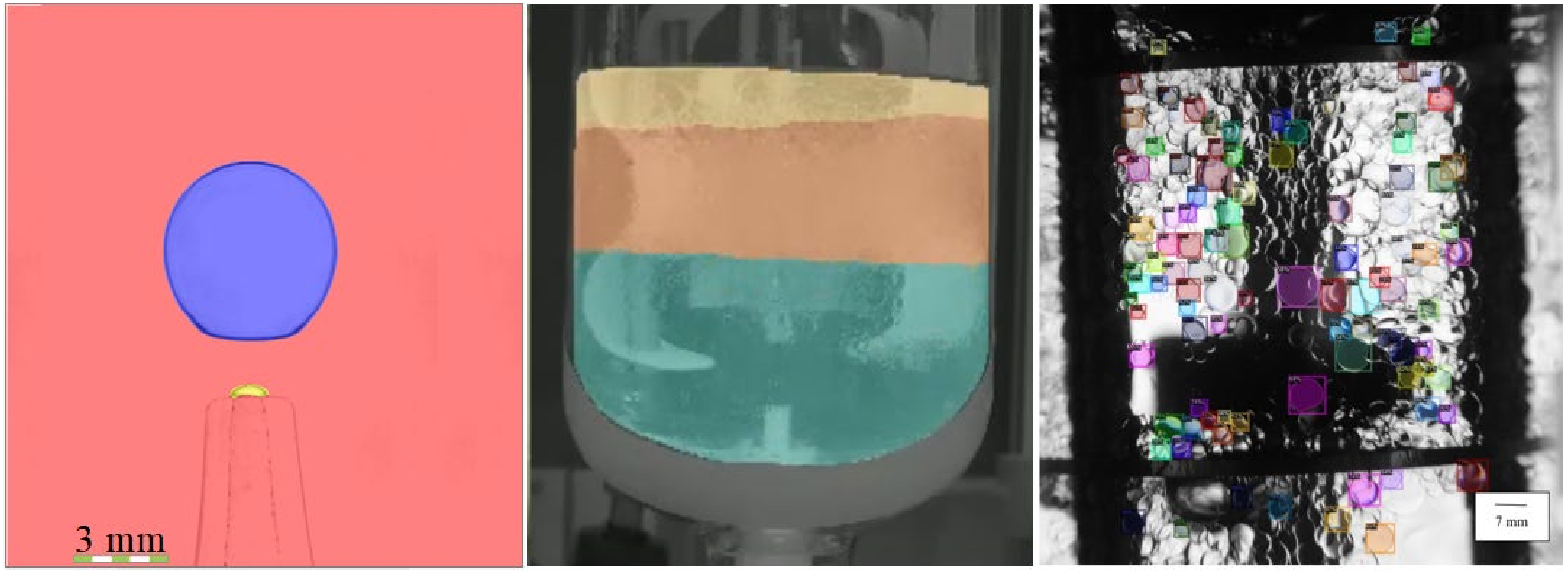
- disperse phase
  - reaction mixture
- continuous phase
  - inert liquid FC40
- screening parameters
  - reagents
  - reaction conditions
- high-throughput automated platform

[1] J. Bobers et al., Org. Process Res. Dev., **2020**

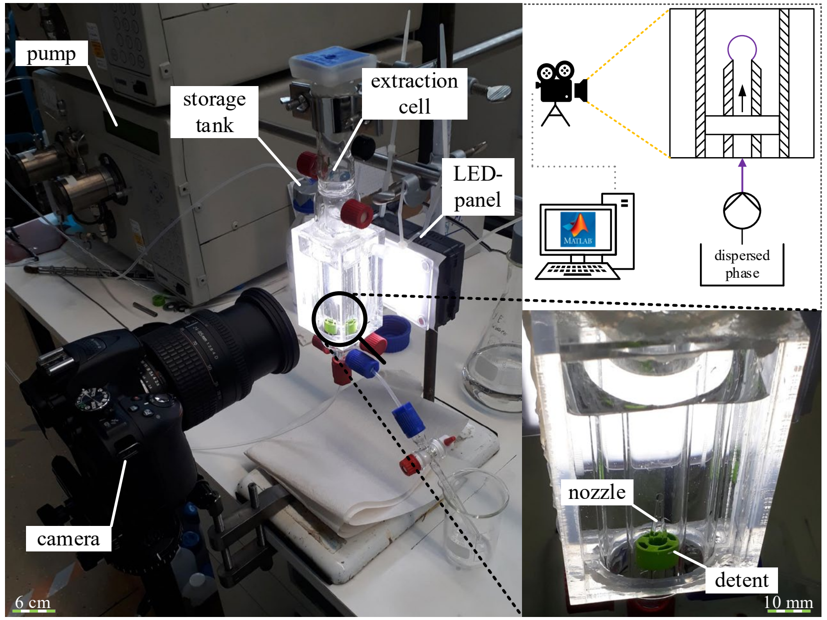
[2] K. Götte et al., ACS Omega, **2022** (submitted)

AI modelling

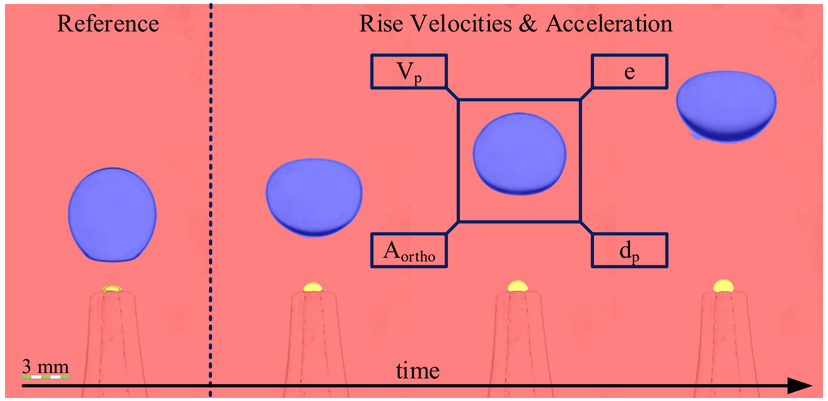
- development of AI-based optical sensors to automate evaluation of lab applications
  - single droplet tracking for parameter estimation
  - liquid-liquid coalescence tracking
  - solvent extraction supervision

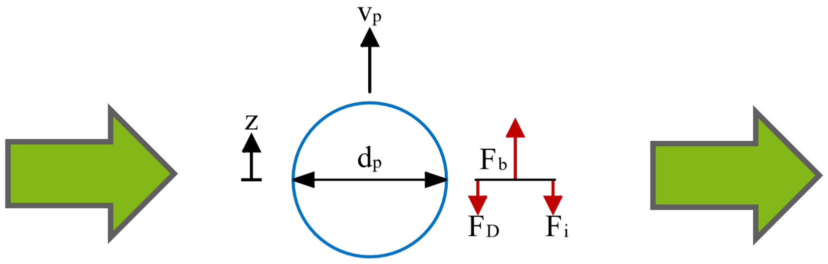


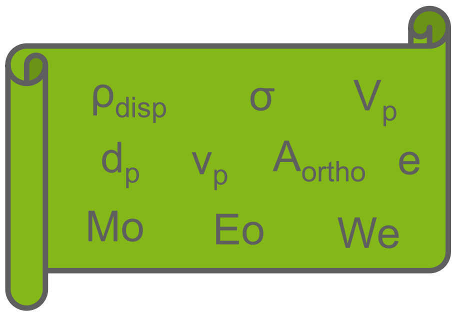
Single Droplet detection



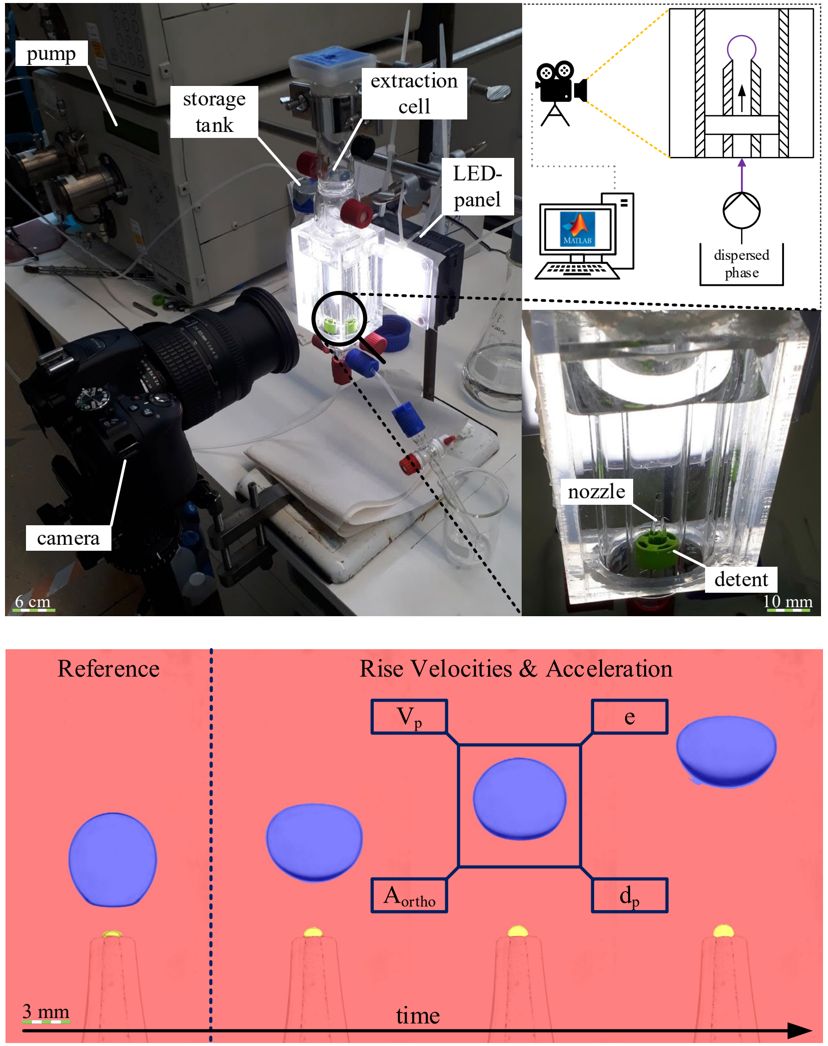
- cheap, easy to use online-densitometer, -tensiometer
- by iteratively solving a force balance of the uprising droplet, various substance parameters can be derived



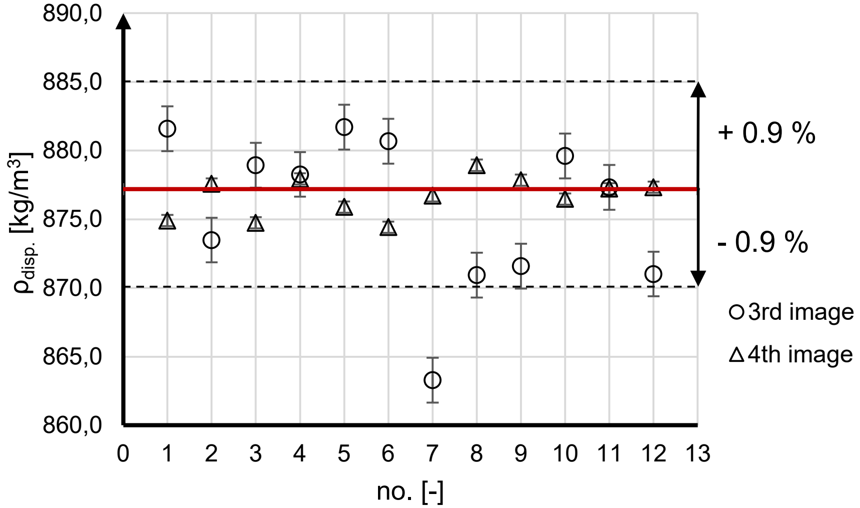




Single Droplet detection



- density measured for n-butylacetate droplets in water
- good agreement with literature values[1]

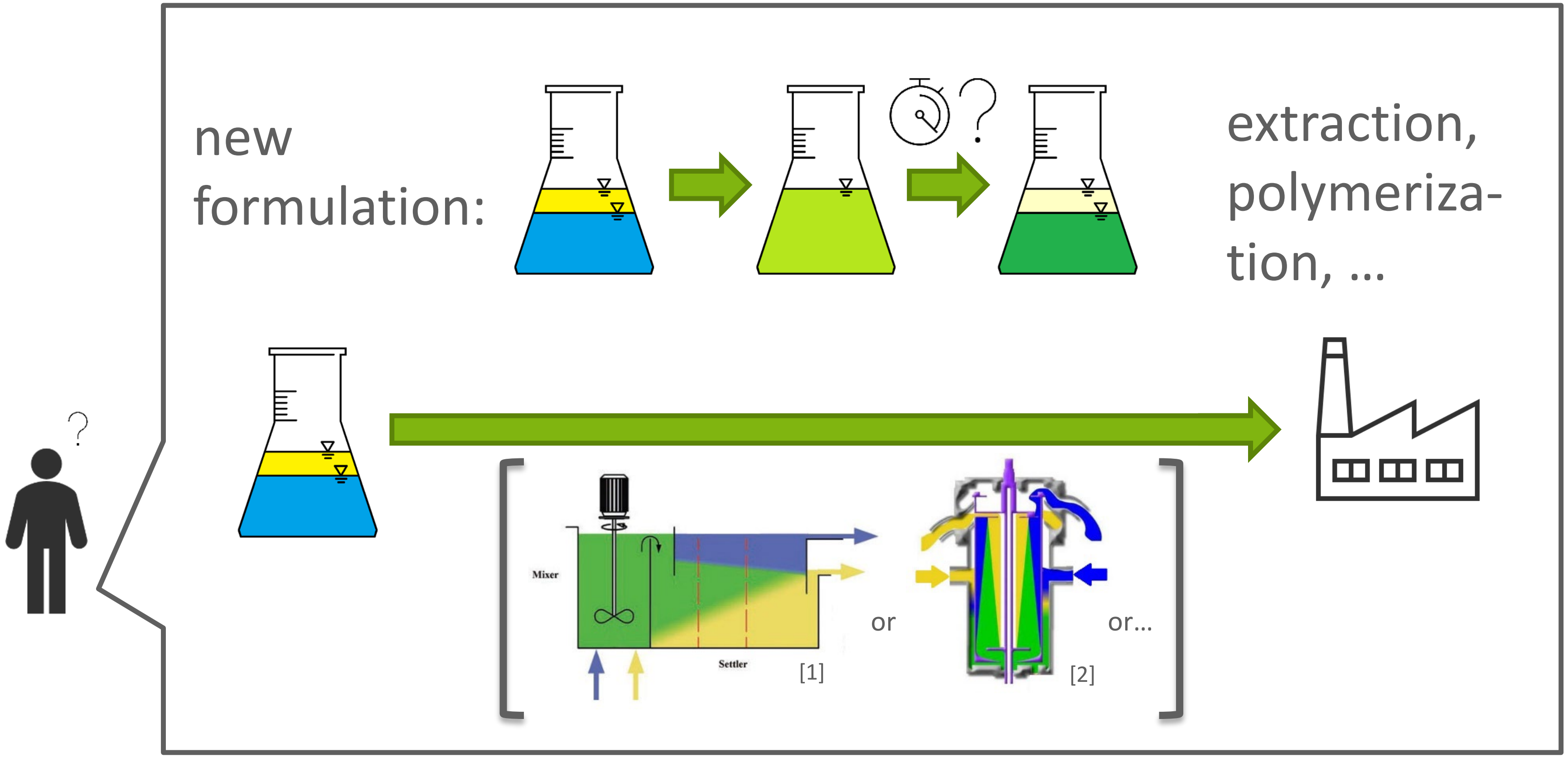


[1] Bäumler, K. et al. Drop rise velocities and fluid dynamic behavior in standard test systems for liquid/liquid extraction—experimental and numerical investigations. Chemical Engineering Science, 2011, 66, 426–439.

Liquid-Liquid Coalescence Tracking

new formulation:

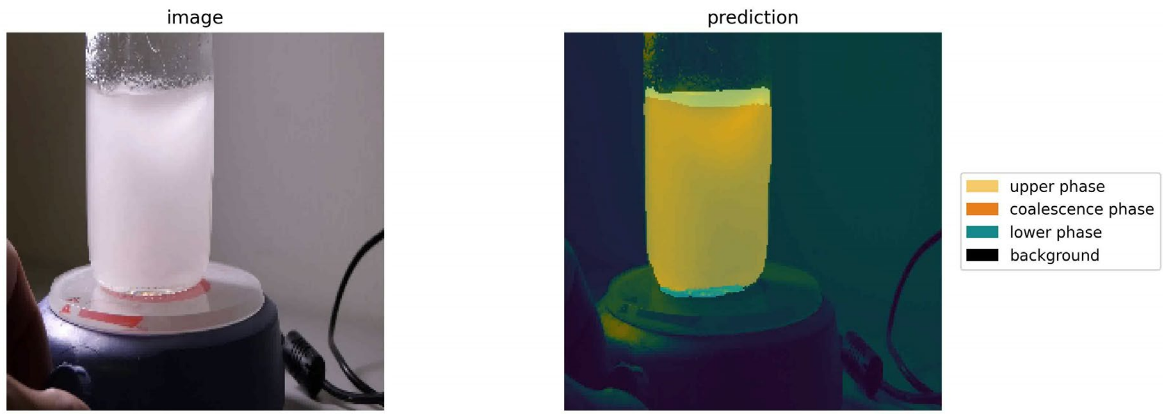
extraction, polymeriza-tion, …

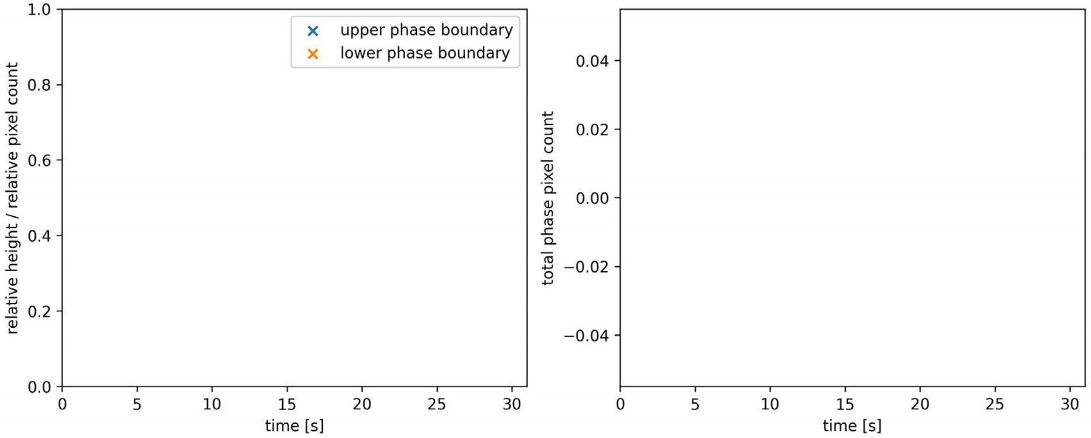


Liquid-Liquid Coalescence Tracking

- precise tool for coalescence observation
  - on clear and colored substance systems
  - flexible (vessel, internals, post-processing)
  - high temporal resolution
  - good reliability
- semi-autonomous
  - requires manual setting of post-processing parameters

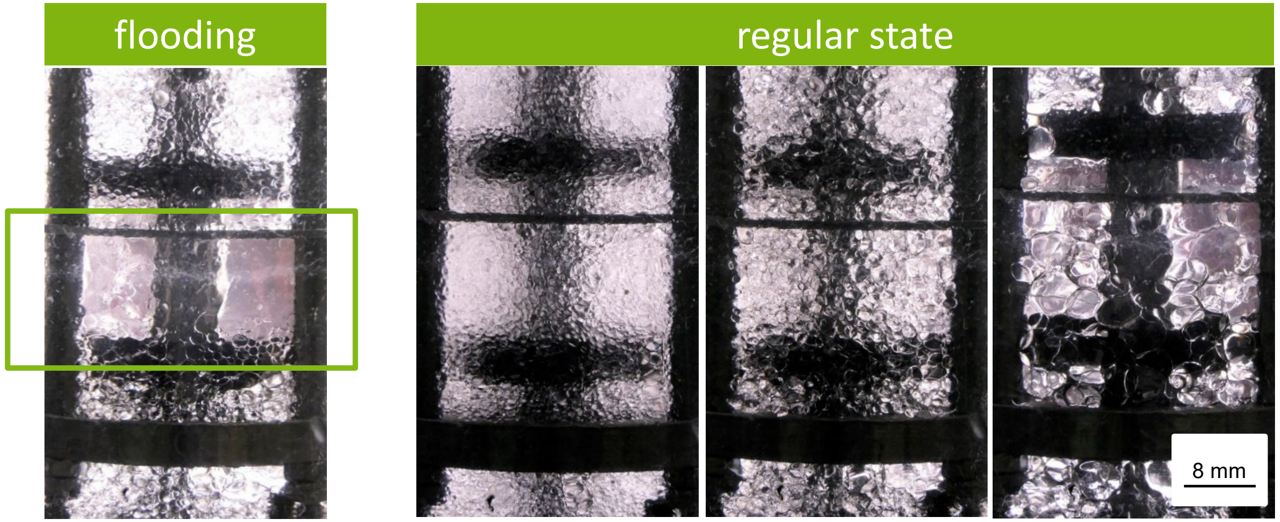
➢ assistance of routine work

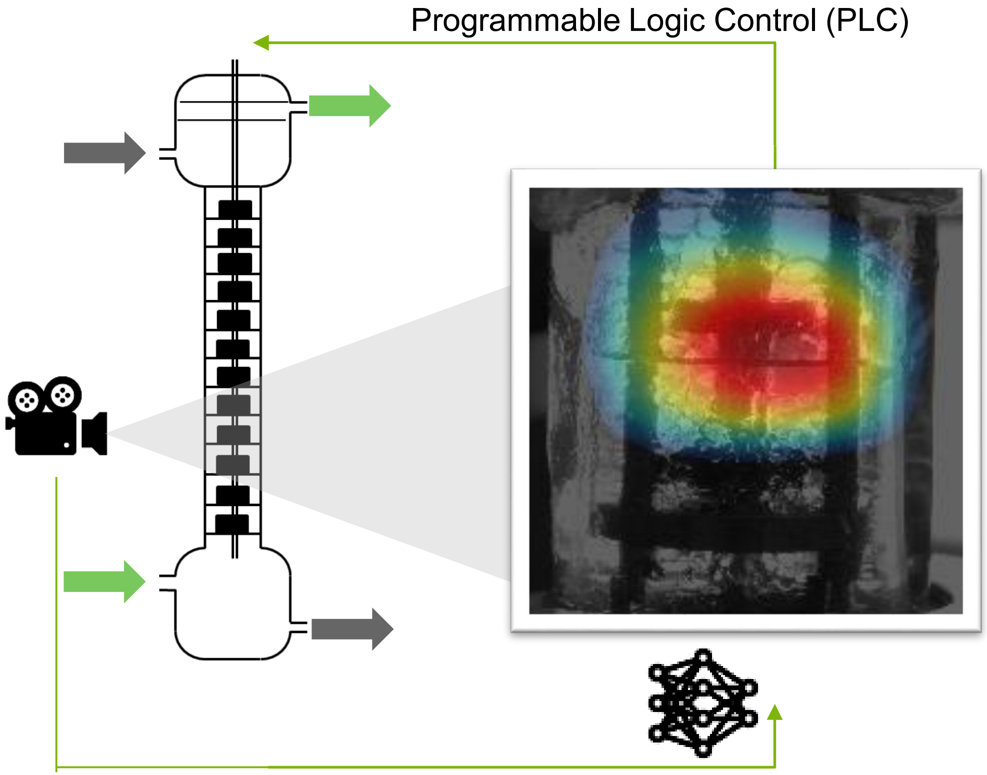




Solvent extraction supervision

- supervision of the hydrodynamics for optimizing operation[1]
    - prevention of unwanted „flooding“ state and determining droplet size

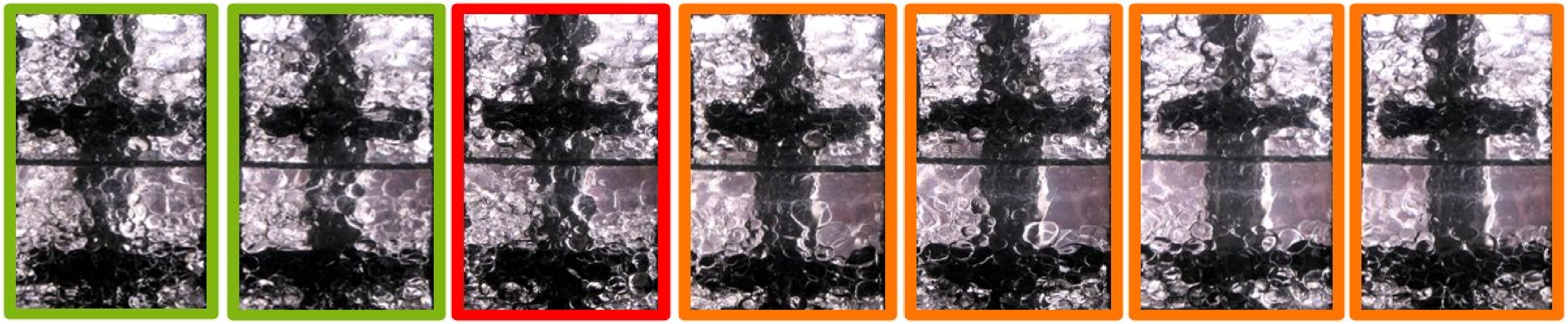




[1] Neuendorf, L.; Baygi, F.; Kolloch, P.; Kockmann, N.; „Implementation of a Control Strategy for Hydrodynamics of a Stirred Liquid–Liquid Extraction Column Based on Convolutional Neural Networks”, ACS Engineering Au, 2022, *DOI: 10.1021/acsengineeringau.2c00014*

Solvent extraction supervision

- supervision of the hydrodynamics for optimizing operation[1]
  - prevention of unwanted „flooding“ state in extraction column using Resnet18[2]
  - class activation map in reasonable region
  - live flooding detection, an image every 0.12 seconds, with early flooding detection as shown below in the image time

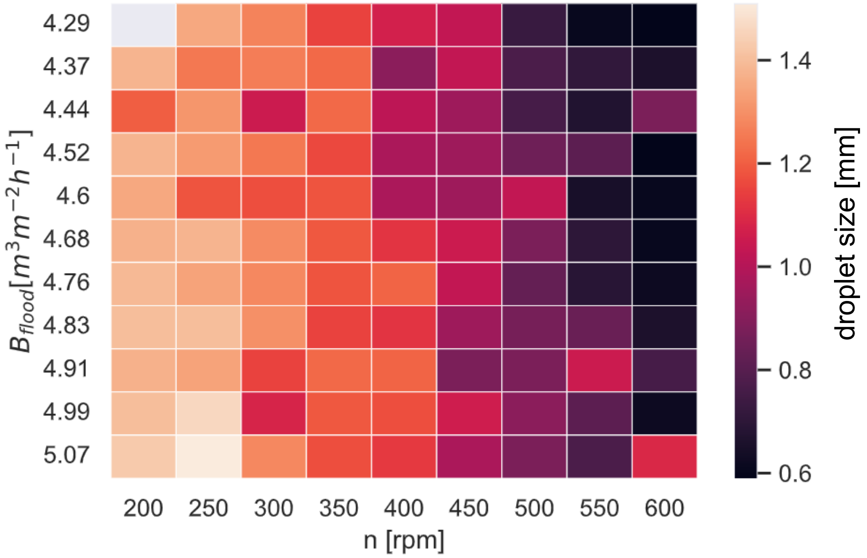


[1] Neuendorf, L.; Baygi, F.; Kolloch, P.; Kockmann, N.; „Implementation of a Control Strategy for Hydrodynamics of a Stirred Liquid–Liquid Extraction Column Based on Convolutional Neural Networks", ACS Engineering Au, 2022, *DOI: 10.1021/acsengineeringau.2c00014*
[2] He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition, IEEE Conference on CVPR , 2015

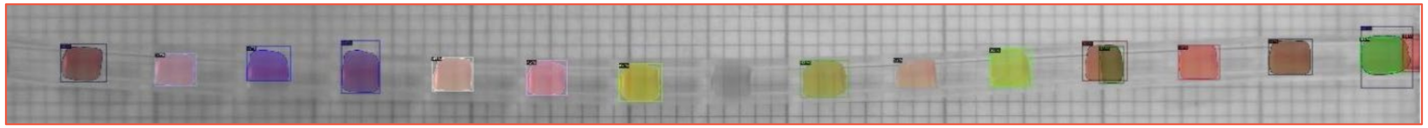
Solvent extraction supervision

- supervision of the hydrodynamics for optimizing operation
  - droplet size estimation[1] as a heatmap for different stirrer speeds n and Loads B

$$B = \frac{\dot{V}_c + \dot{V}_d}{A}$$



  - investigation of transferability to other systems such as tubes

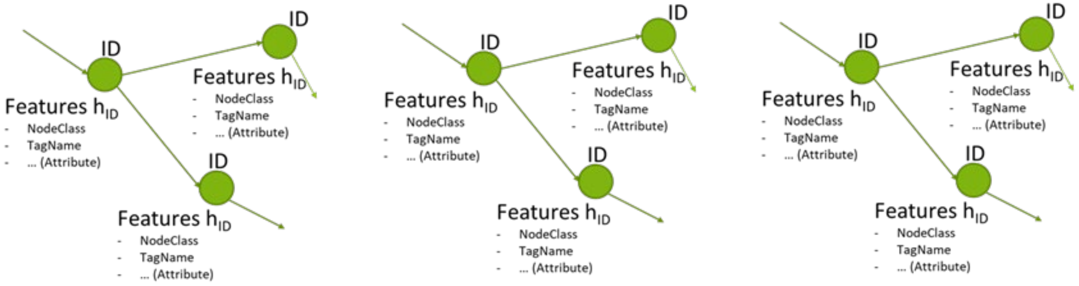


[1] He, K.; Gkioxari, G.; Dollar, P.; Girshick, R.; "Mask R-CNN", Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017, pp. 2961-2969

Artificial intelligence (AI) and engineering

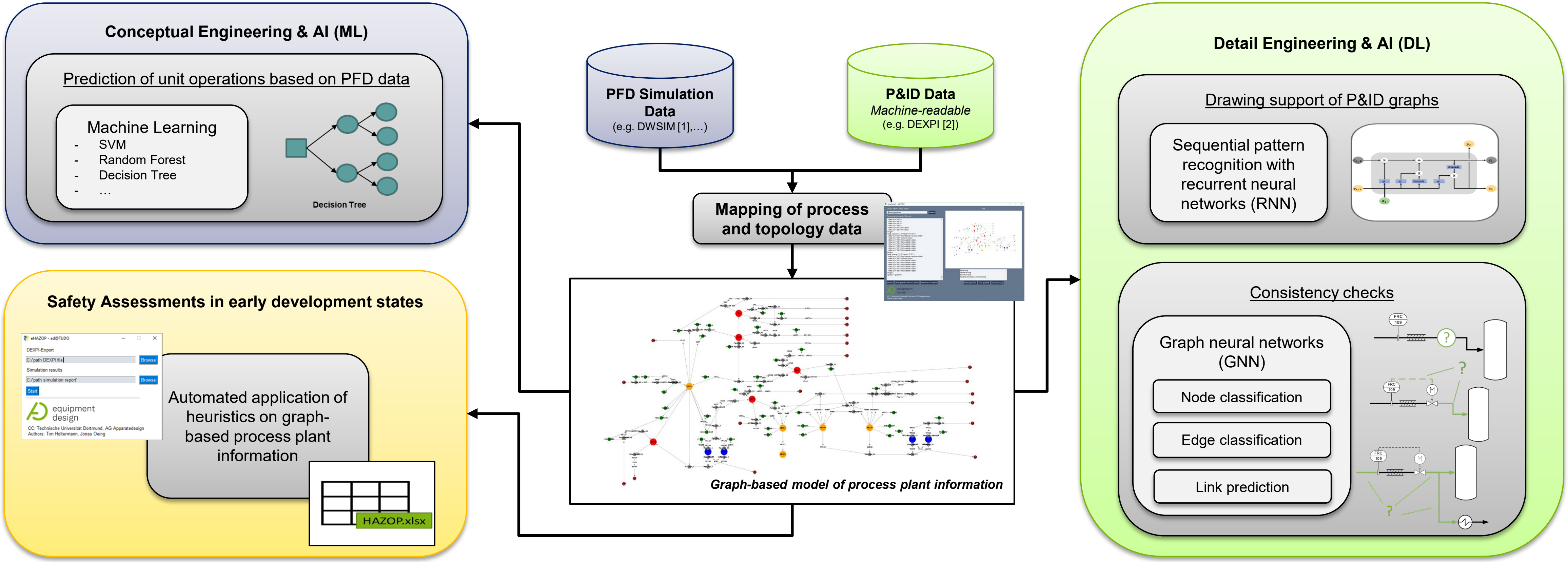
How engineering data / workflows should be designed to be accessible for artificial intelligence (AI) or deterministic algorithms?

- many different areas with a lot of different, often still analog (printed) data formats
  - process simulation flowsheets
  - piping and instrumentation diagrams
  - safety documentation
  - lists of equipment and piping
- high amount of correlating data but no machine-readable connection of the data
- engineering knowledge is available but cannot be learned by AI methods due to poor harmonization [1]
- graph-based data structures!



[1] Wiedau, Tolksdorf, Oeing, Kockmann, Chem. Ing. Tech., 2021

Advantages of graph-based process plant information



[1] DWSIM open-source process simulation, www.dwsim.org, 2022

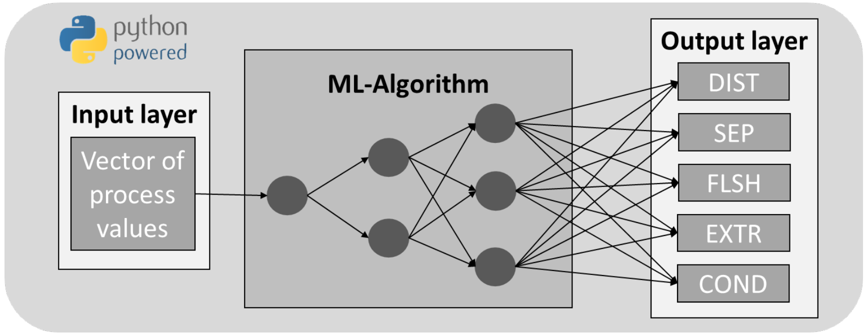
[2] Data exchange in process industry, www.dexpi.org, 2022

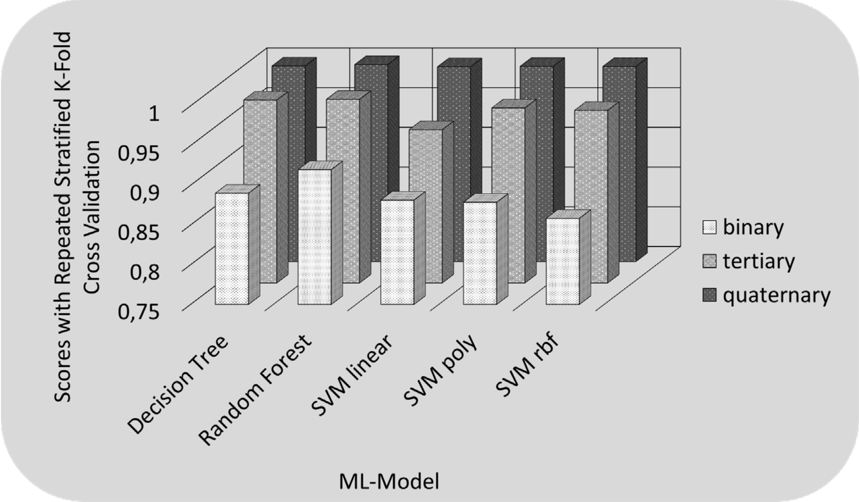
ML-prediction of unit operations

Machine Learning (ML) to suggest unit operations[1]

- ML-based prediction of separation units
  - inputs: process values (e.g. T, p, p^{LV}, H etc.)
  - output: separation units (e.g. distillation, flash, condensor etc.)
  - trained for binary, tertiary and quarternary substance mixtures

- results can acceleration future synthesis and simulation of processes
  - consistency checks
  - automated process simulation





[1] Oeing, Henke, Kockmann, Chem. Ing. Tech., 2021

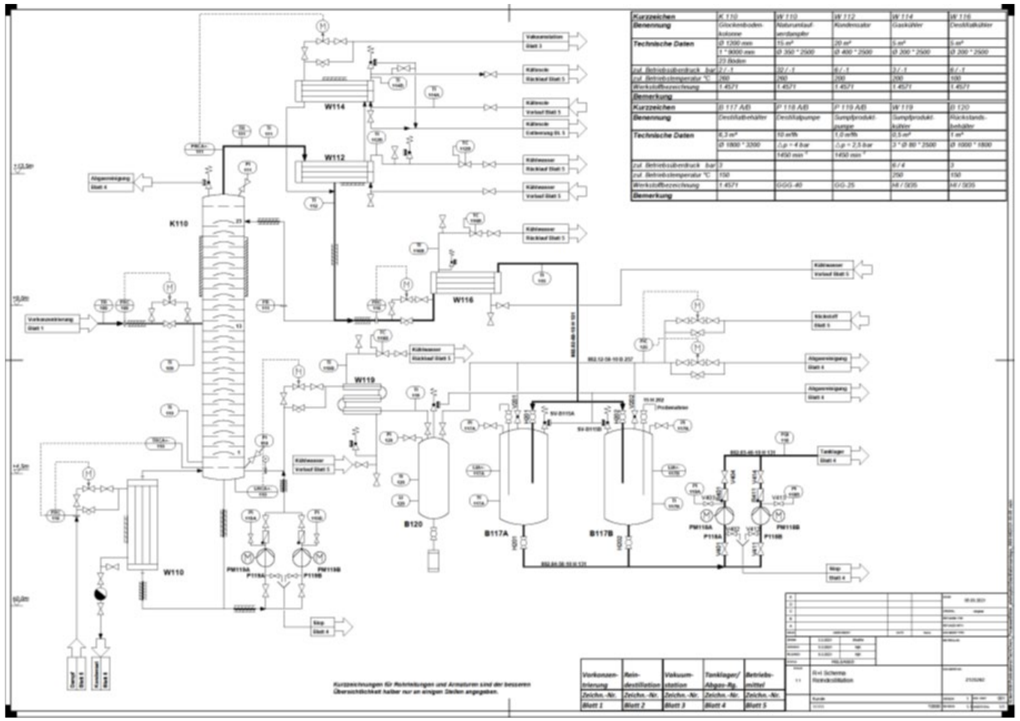
consistency checks of P&IDs

Piping and instrumentation diagram (P&ID)

- P&ID is the most important document of a plant
- creating and maintaining P&IDs is a very time-consuming task

P&ID:

- describes the plant topology in the process industry (equipment, piping, control loops, etc.)
- describes plant specifications (max./min. temperatures, max./min. pressures, materials etc.)
- uniform documentation of equipment and piping according to DIN EN ISO 10628
- uniform documentation of the process control technology according to EN 62424



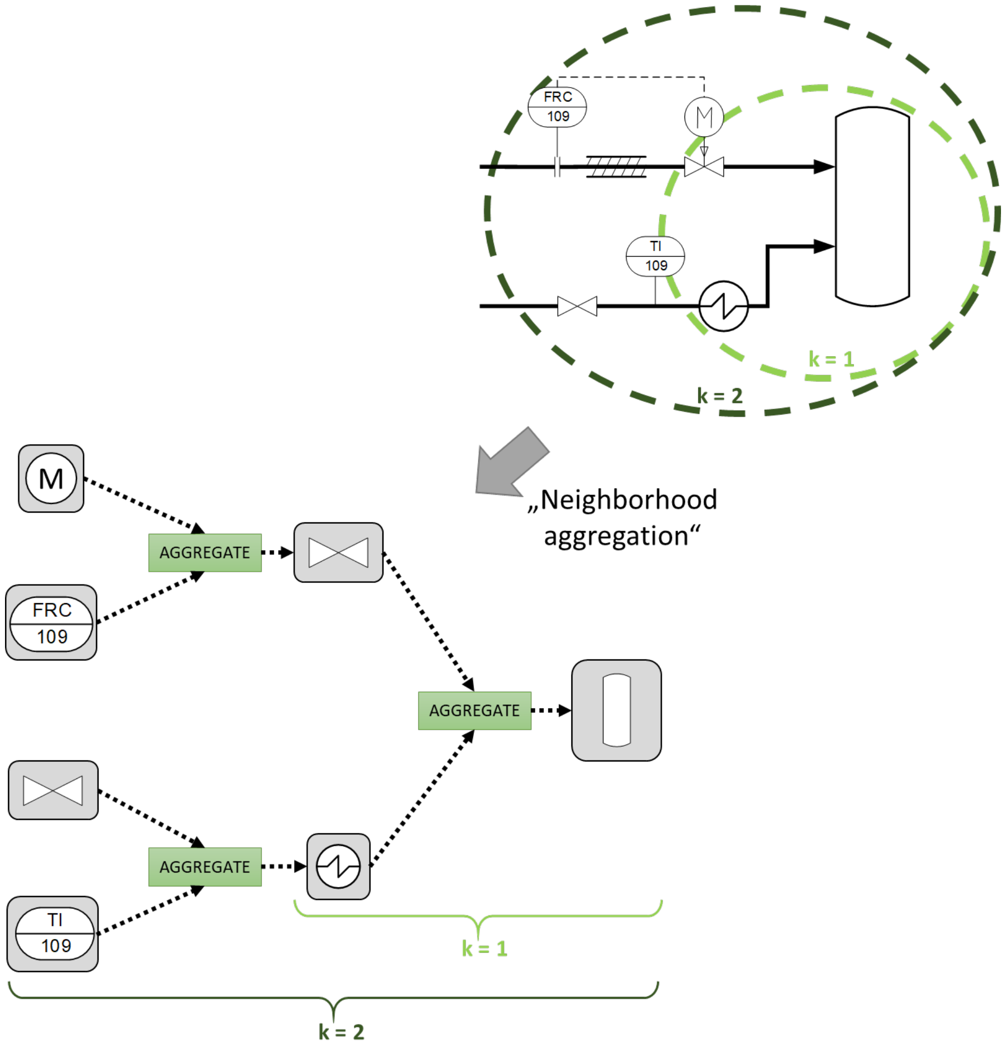
P&ID of a distillation plant

Graph neural network (GNN)

- k-layer GNN's generate embeddings h_u of node u based on their neighborhood structure[1]
- message passing in GNNs[2]

$$h_u^{(k+1)} = Update^{(k)}\left(h_u^{(k)}, Aggregate^{(k)}\left(\left\{h_v^{(k)}, \forall v \in N(u)\right\}\right)\right)$$

  - aggregation: collects neighborhood information
  - update: introduces a non-linearity into the output of a neuron

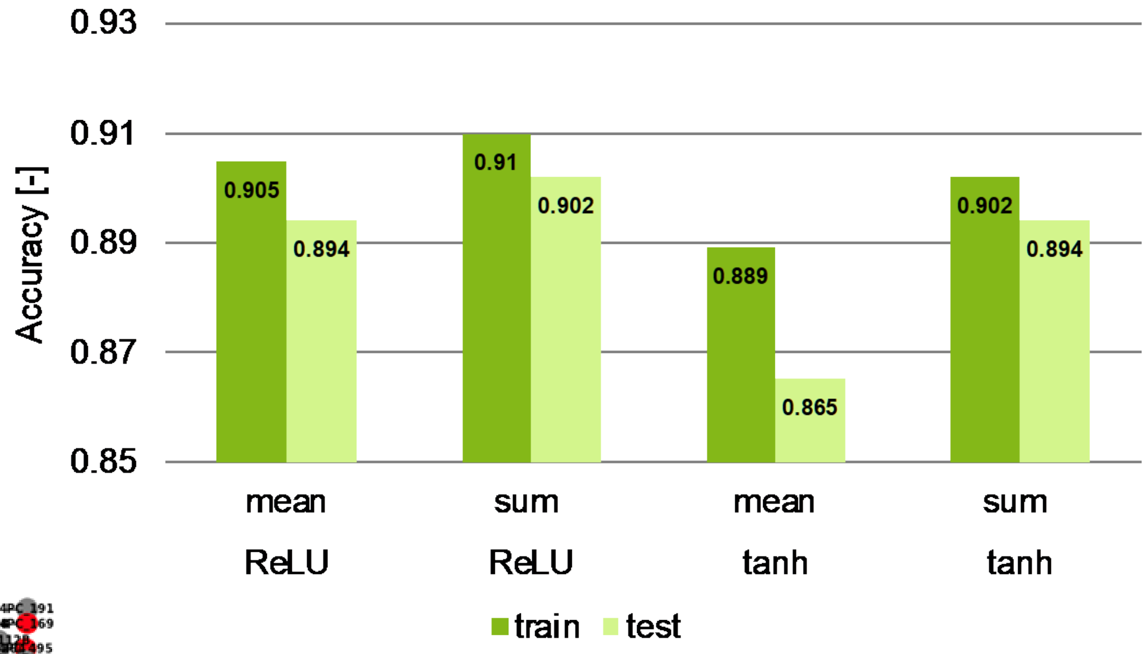


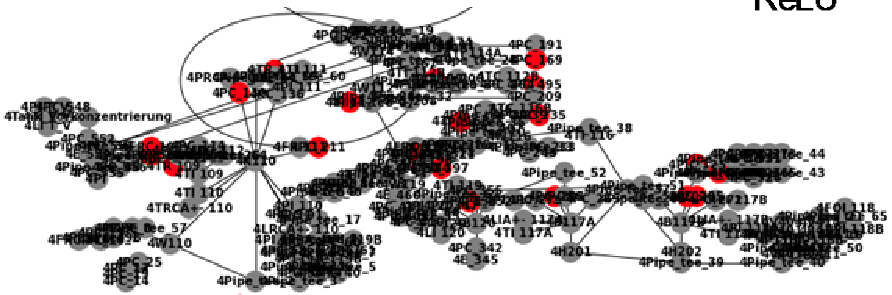
[1] Leskovec, J., Inductive Representation Learning on Large Graphs, 2017
[2] Hamilton, W., Graph Representation Learning, 2020

Results - GNN

- node classification via a GNN[1]
- 12 P&ID graphs
  - 2005 nodes, 2167 edges
  - train/test split: 0.7 / 0.3

- weightedSAGEconv[2] (2-layer)
  - aggregation: sum
  - activation: ReLU
  - hidden neurons[3]: 35
  - 91.0 % Training
  - 90.2 % Test





[1] Hamilton, W. et al., Inductive Representation Learning on Large Graphs, 2018 [2] Deep Graph Library, 2022. [3] Hagen, M. et al., Neural Network Design, 2014

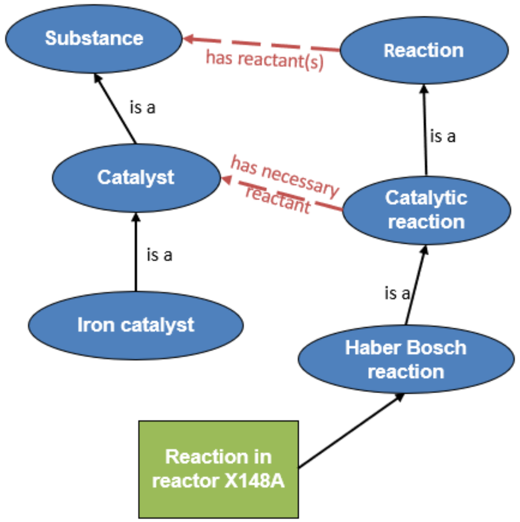
Research data management - NFDI

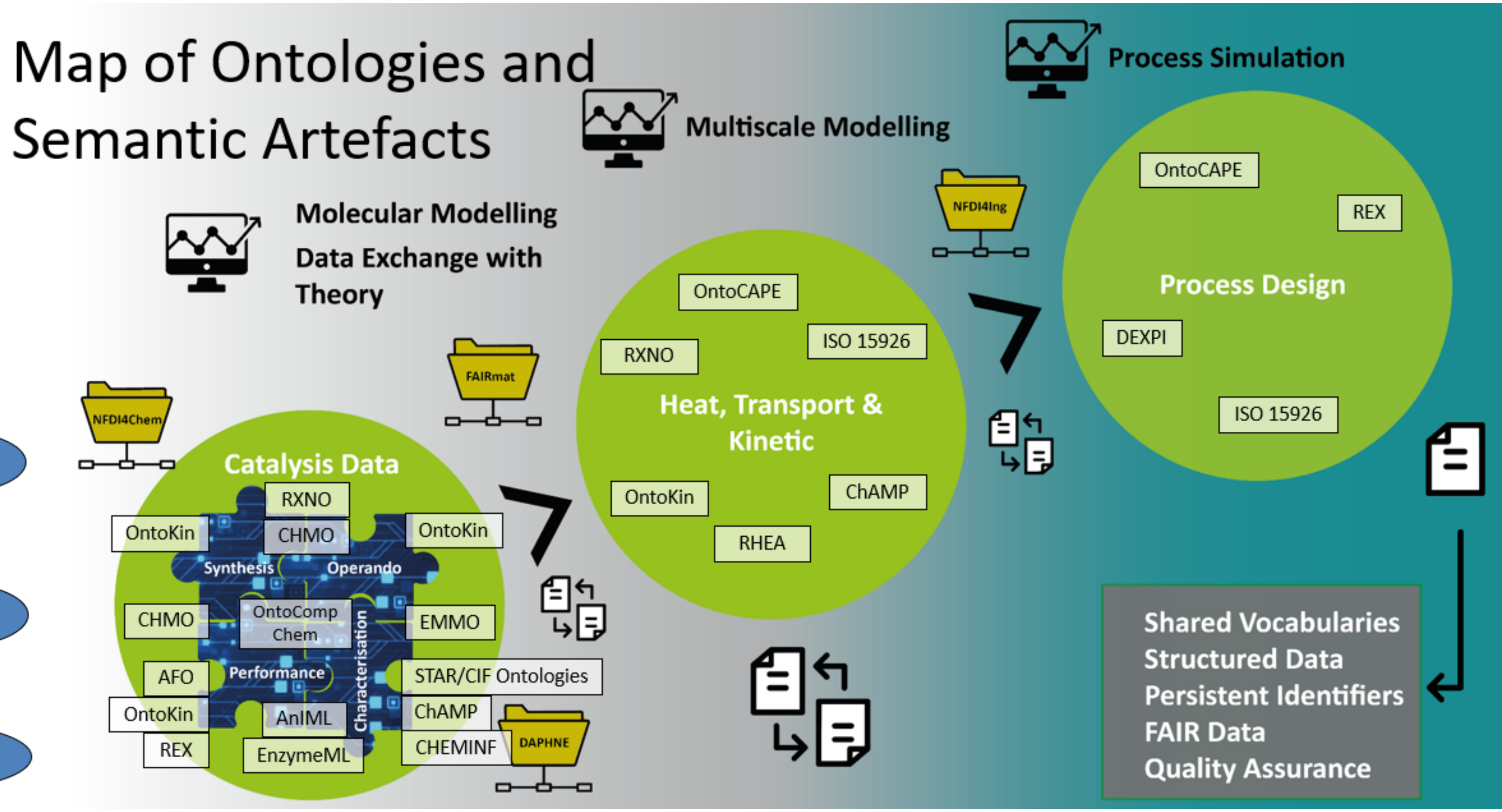
RDM in Catalysis and Process Engineering

NFDI4Cat

- ontology overview in NFDI4Cat [1]
- contact to NFDI4Chem and NFDI4Ing

- BCI-AD is engaged in
  - ontology development
  - metadata standards
  - related workflows



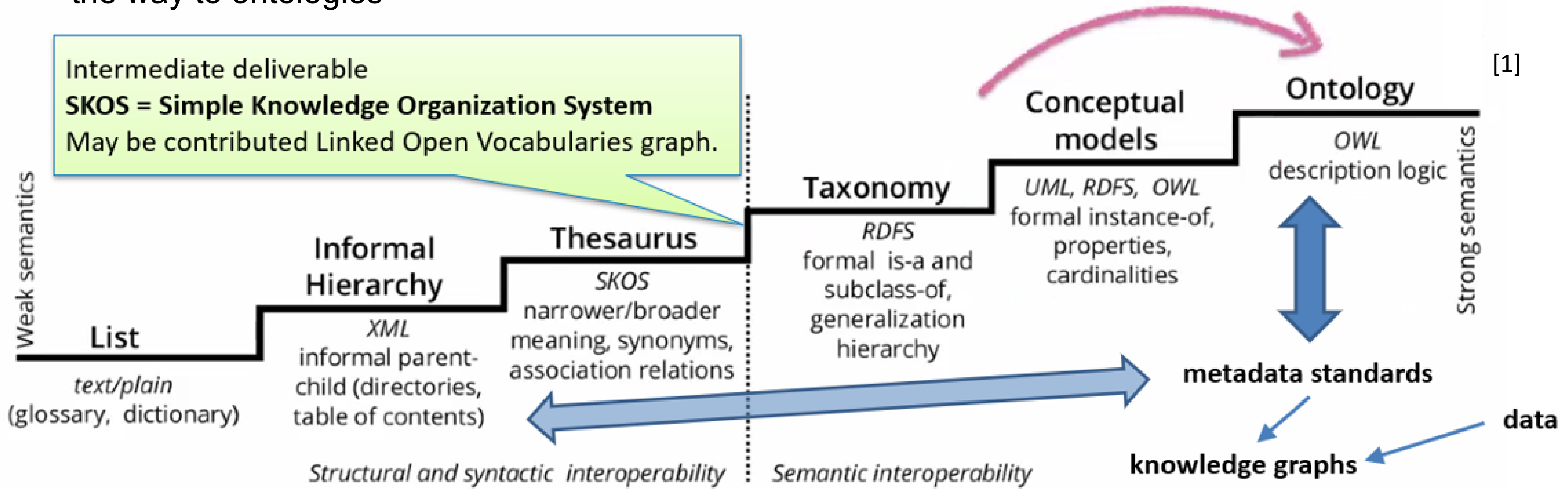


[1] See also: nfdi4cat.org/ontology-collection/

Ontologies in Catalysis and Process Engineering

NFDI4Cat

- the way to ontologies



[1]

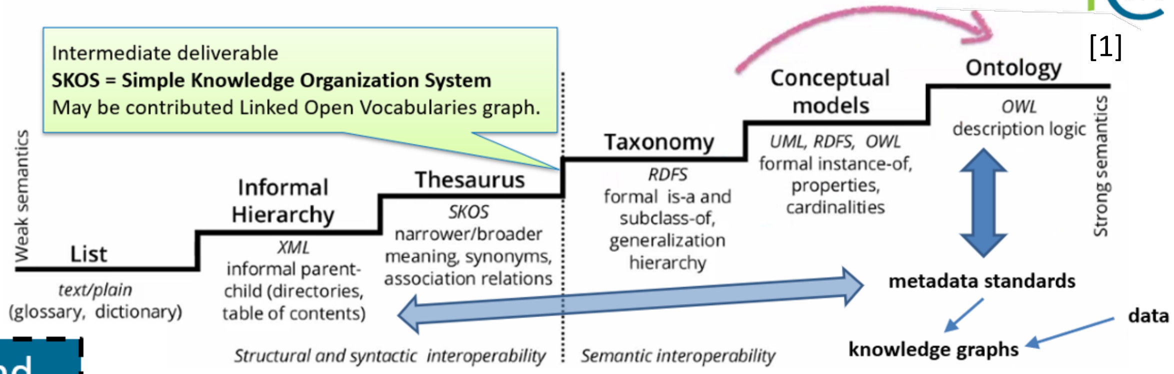
[1] adapted from https://th.fhi-berlin.mpg.de/meetings/fairdi2020/index.php?n=Meeting.PosterDetails&poster_id=18

Research data management - NFDI

Ontologies in Catalysis and Process Engineering

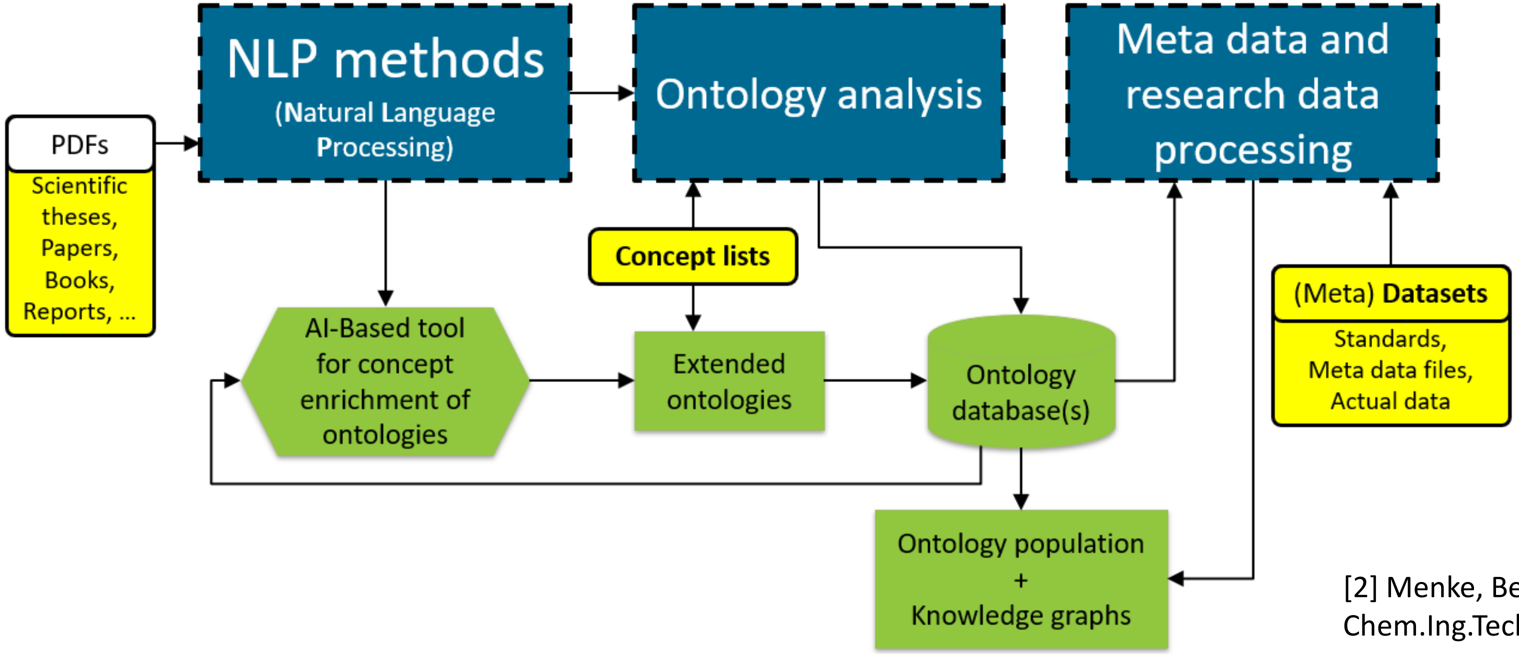
NFDI4Cat

- the way to ontologies



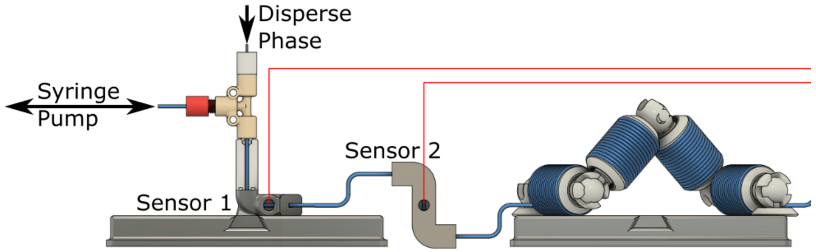
[1] adapted from https://th.fhi-berlin.mpg.de/meetings/fairdi2020/index.php?n=Meeting.PosterDetails&poster_id=18

- NFDI4Cat-TA1 metadata workflow
  - AI tools for text processing

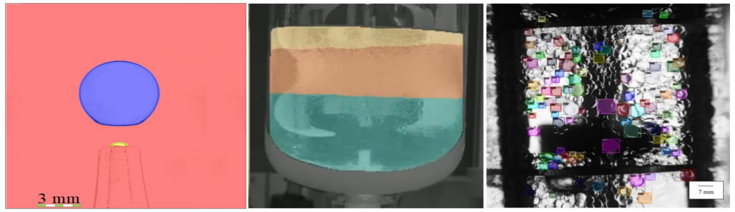


[2] Menke, Behr et al., Development of an Ontology for Biocatalysis, Chem.Ing.Techn., 2022, submitted

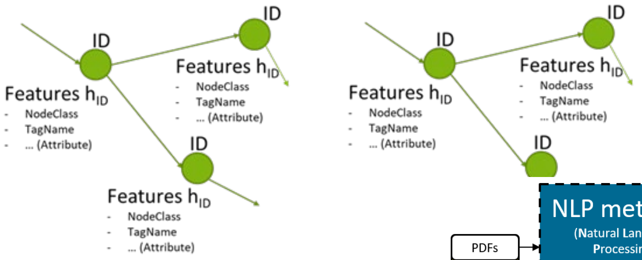
- Lab automation (Robin Dinter)



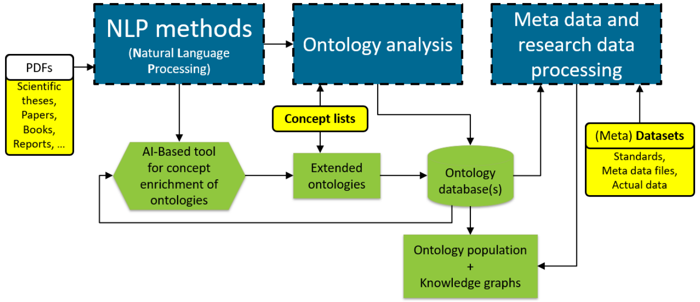
- AI modelling (Laura Neuendorf)



- Process engineering (Jonas Oeing)



- Research data management (Alex Behr)



Thank you for your attention!

www.ad.bci.tu-dortmund.de