From thermodynamics to processes: AI facing small data and big models
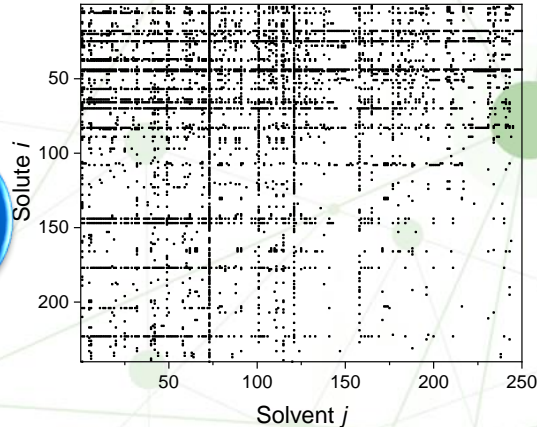
Michael Bortz

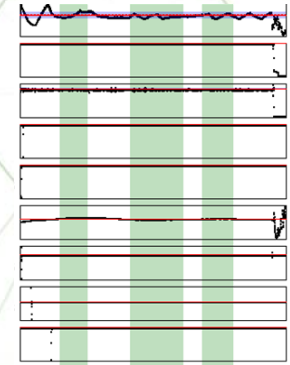**27.06.2023**

Small data

$\gamma^\infty (T = 298.15 \pm 1\ K)$ from DDB:
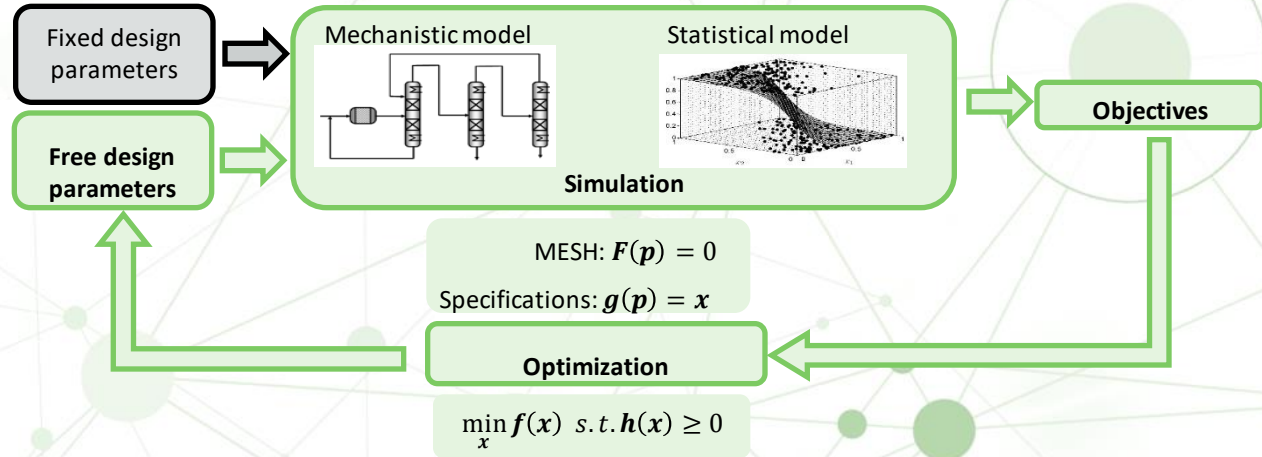240x250=60.000 entries; 4094=6,8% measured

Miniplant data over several hours

Big data

Value

Variability
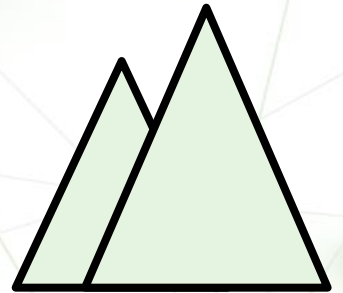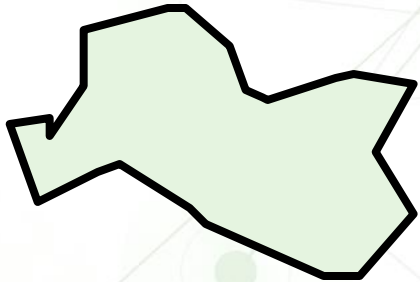
Velocity

Veracity

Volume

- Expensive experiments (high value, small volume)
- Restricted variability in process operation
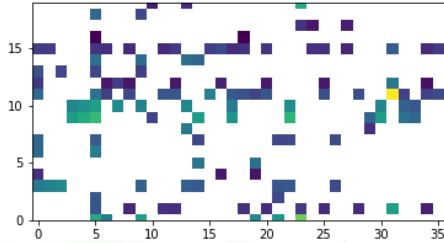- Restricted access, data are considered as confidential

Complex models

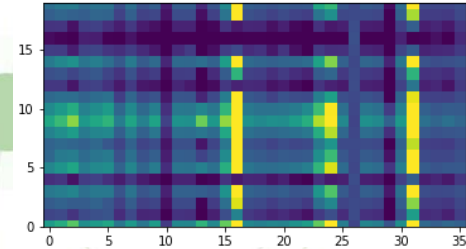- Mixture properties
- Topology
- Unit operations

Fixed design parameters		
	Mechanistic model	Statistical model
Free design parameters		
	Simulation	

$$\text{Objectives}$$

MESH: $\boldsymbol{F(p)} = 0$

Specifications: $\boldsymbol{g(p)} = \boldsymbol{x}$

Optimization

$$\min_{\boldsymbol{x}} \boldsymbol{f(x)} \ \ s.t. \boldsymbol{h(x)} \geq 0$$

- Feasible domain for $\boldsymbol{x}$ is generally unknown (convergence, constraints)
- Not much structural information

Contributions from KEEN

Matrix/tensor completion

- Integration of prior model knowledge
- Benchmarking on exemplary industrial test cases by physical models
- Experiments to create new data
- Software prototype

- Integration of user-defined units
- Real-time surrogates
- Explicit instead of implicit models
- Software prototypes

Simulation of process data