HOHENSTEIN

Application of Statistical Concepts to Laundry Business

Dr. Igor Kogut, Hohenstein

Outline

- Why the utilization of statistical concepts is especially useful for laundry business?
- Evaluation of historical data and strategical data acquisition
  - How I can process already stored (historical) data?
  - Which strategy I can use to obtain valuable data in my laundry?
  - Which Software Environment is necessary for realization of “data science” projects in your laundry
- Opportunities and challenges of the application of statistical concepts
- Examples from research & industrial applications

belongs to
Data Science

How useful is „Data science“ for laundries and their suppliers?

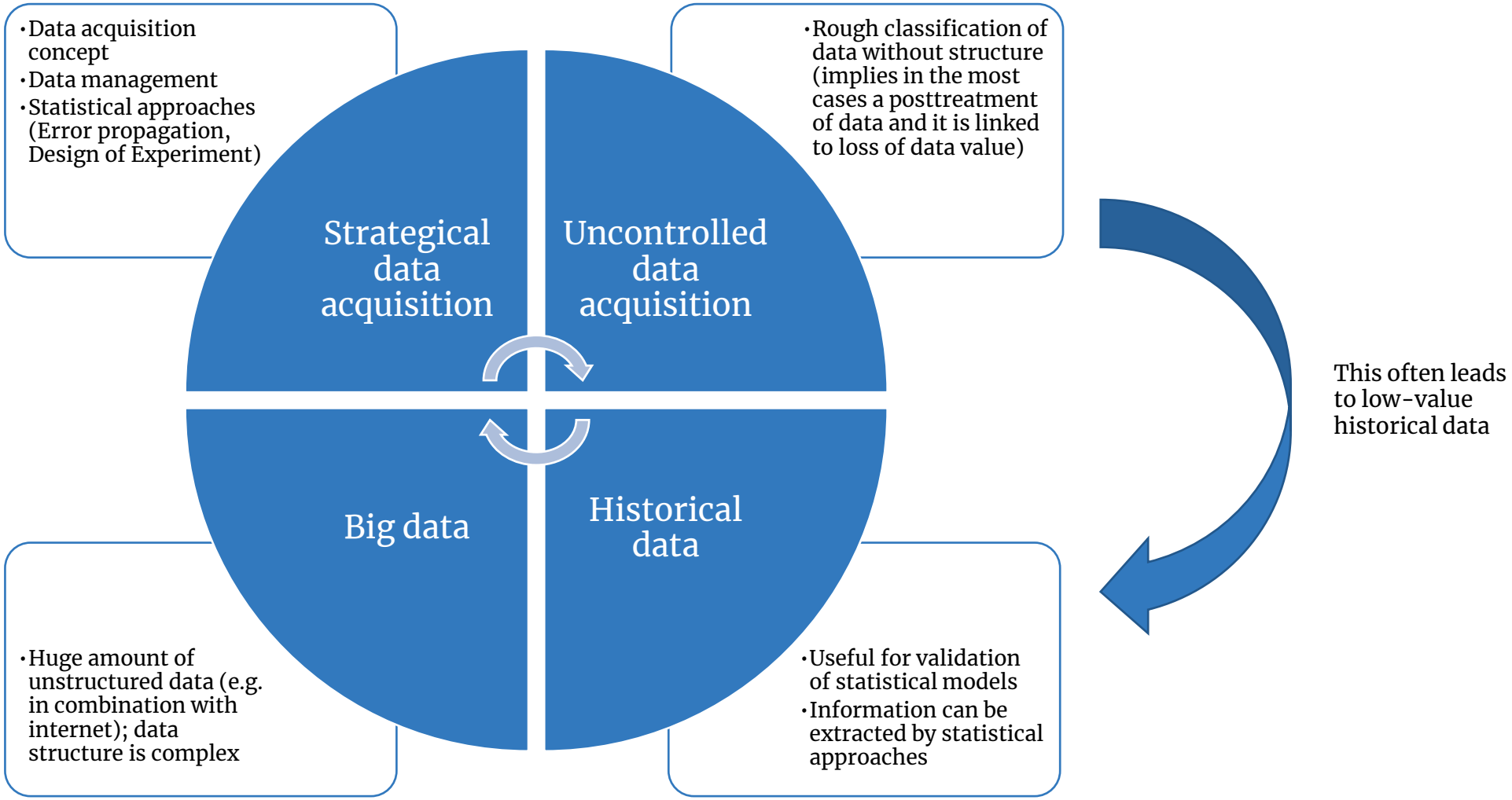
- Laundry business is determined by tailor-made technological equipment (i.e. laundry machines) as well as tailor-made usage of cleaning chemistry

  → fast adjusting/optimization of all process steps are important

- Development of novel cleaning processes for special customers/application cases e.g. personal protection equipment of firefighters (PPE)

- The processes are not too complex and “data science projects” can be easily realized with a relatively small „budget“

Sources of data in industrial environment



Strategical data acquisition

Set the " data science" concept

- For which purpose do you collect the data (process screening, process optimization)?
- Which data quality you want to achieve (e.g. accuracy, redundancy)?
- Costs for data acquisition (time, money)

Strategy & pre-evaluation

- Set strategy

(e.g. web scrapping, RShiny)

- Set technology

(Tableau, Python, SQL etc.)

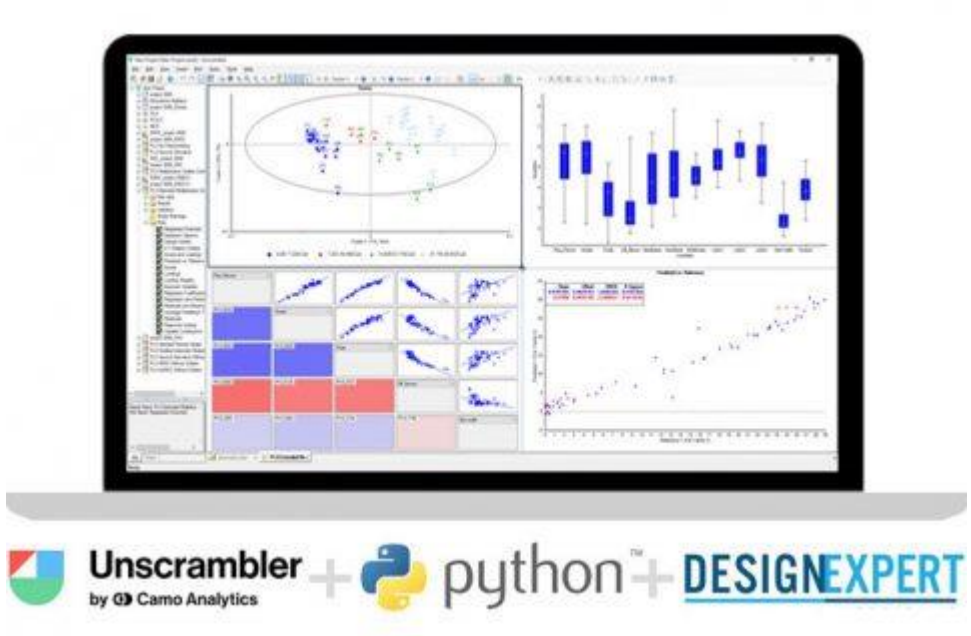
- Check/simulate the statistical parameters of data acquisition models (e.g. statistical power)

Transformation, evaluation and validation

- Downsample, clean, simplify, transform data
- Define test and train datasets for machine learning
- Set models and evaluate the results
- Validate the models e.g. for predictive analytics

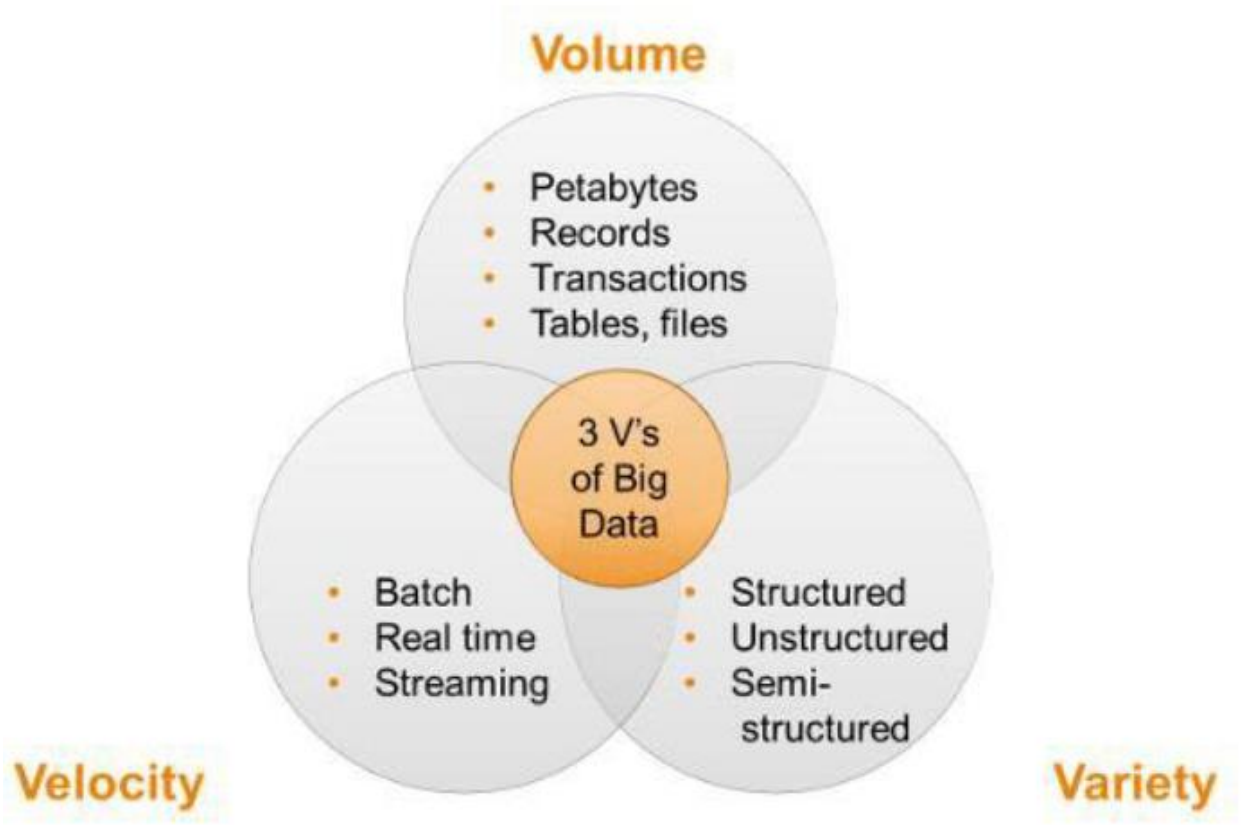
“Historical” data analysis

- For this kind of data, only “output” data exist (we don’t know which inputs have a significant impact)
- Low-quality of data (e.g. missing values)
- Can be used as “Test Data Set” for machine learning projects (e.g. linear, logistic regression etc.)
- Can be analyzed with clustering algorithms: which groups of “outputs” have similarities? (e.g. multivariate data analysis with Unscrambler, R, Python etc.)

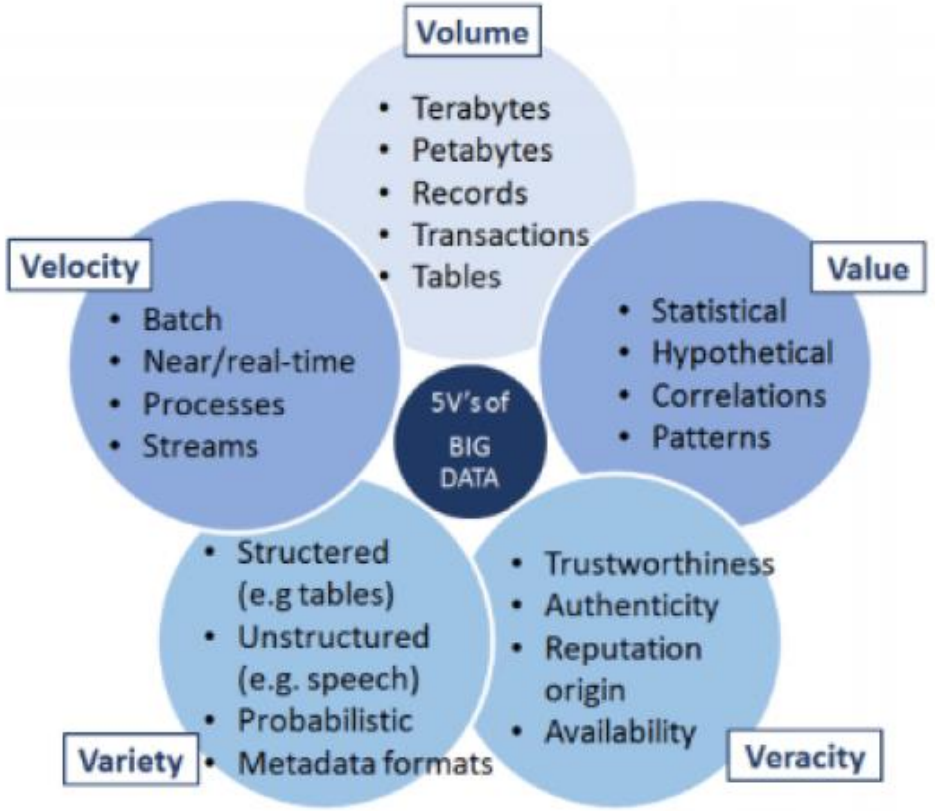


www.camo.com

Big data: 3 V's vs. 5 V's definitions



researchgate.net



researchgate.net

Strategies & Technologies for Big Data analysis (as example only in R)

- High-complexity projects which include all data of company or supply chain (e.g. laundry suppliers → laundries → health care sector)
- Data can be additionally collected from internet (e.g. web scrapping https://www.dataquest.io/blog/web-scraping-in-r-rvest/

  → in internet a lot of valuable information is published for laundries e.g. lists of disinfection chemicals (https://www.desinfektionsmittelliste.de/)

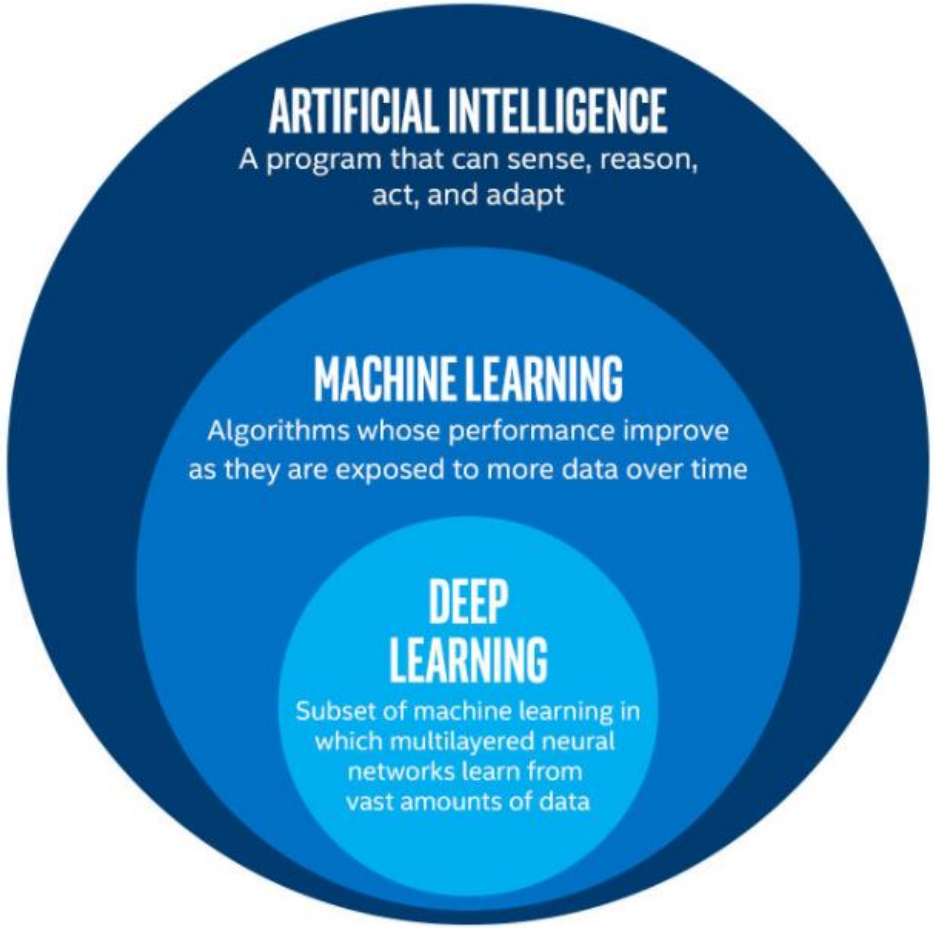
- Different strategies can be used for analyzing Big Data (e.g. downsample to database → pull sample to Dev Machine → Build Model https://rviews.rstudio.com/2019/07/17/3-big-data-strategies-for-r/)

- Technologies as Apache Hadoop, Spark/Sparklyr, H2O may be considered for big data science projects (https://dzone.com/articles/rsparkling-gt-the-best-of-r-h2o-spark)

Artificial intelligence vs. Machine Learning vs. Deep Learning

Machine Learning (ML) projects make sense for laundry business:

- Supervised ML: Correlation of output (continuous, discrete) with inputs variables to statistical predictive models (e.g. predicting chemical dosage in laundering process for reasons of resource efficiency)
- Unsupervised learning: only output variables exist and can be clustered to groups to find out if there are any correlations (e.g. analysis of customer groups of laundries i.e. hospitals, elderly residence etc.)
- Reinforcement learning is more interesting for robotics (i.e. machine laundry equipment) and will be not discussed here in detail



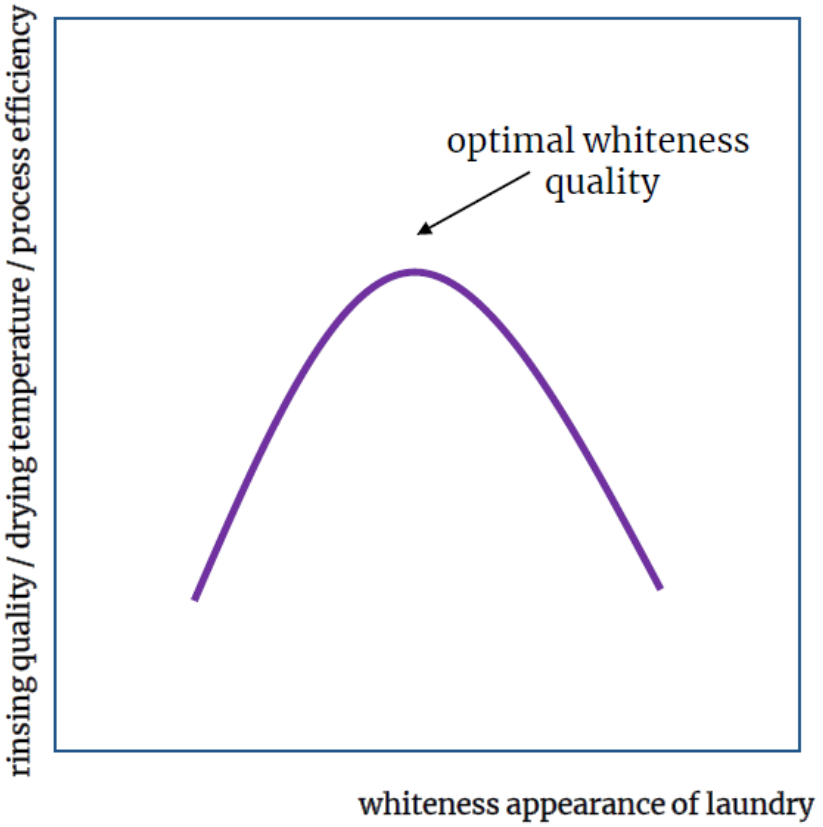
https://medium.com/

Opportunities of "data science" projects in laundries

- Small-scale projects need to be considered at first, as a risk of collecting low-value data (e.g. missing data, lack in error consideration) is high
- Effective reducing of resource consumption (e.g. energy consumption)

  → E.g., Screening & Optimization via Design Experiment 12

  (costs of software and education of leading laundry engineers are about 10000 €; the savings through optimization measures are much higher)
- Preparing high-quality data for digital transformation of laundries as a long-term investment



Challenges of “data science” projects in laundries

- People in laundry business need to be educated for quality management in accordance with goals of digital transformation process
- Clarification who owns the laundry data (i.e. machine industry or supplier of washing chemistry or the laundry itself)
- Preparing the hardware infrastructure (e.g. computer storage space, cloud solution etc.)
- The amortization time cannot be easily defined, as there is not much experience with “data science projects” in this industry area

Examples from scientific and industrial application areas

- 1) Screening of influence parameters related to release of microplastics
- 2) Development of soft sensors for monitoring of rinsing quality in laundry processes
- 3) Clustering of data for sustainability evaluation of laundry performance

Screening of influence parameters related to release of microplastics

Process parameters*	Unit	Level -1	Level 0	Level 1
Temperature	°C	40	60	80
Duration of washing process	min	30	60	90
Detergent concentration	g/l	0	2	4
Mechanics (number of metallic spheres)		0	50	100
Washing liquor to cotton fabric ratio		1:7	1:14	1:21

* Data were acquired in a lab washing machine

via Software „Design Expert“

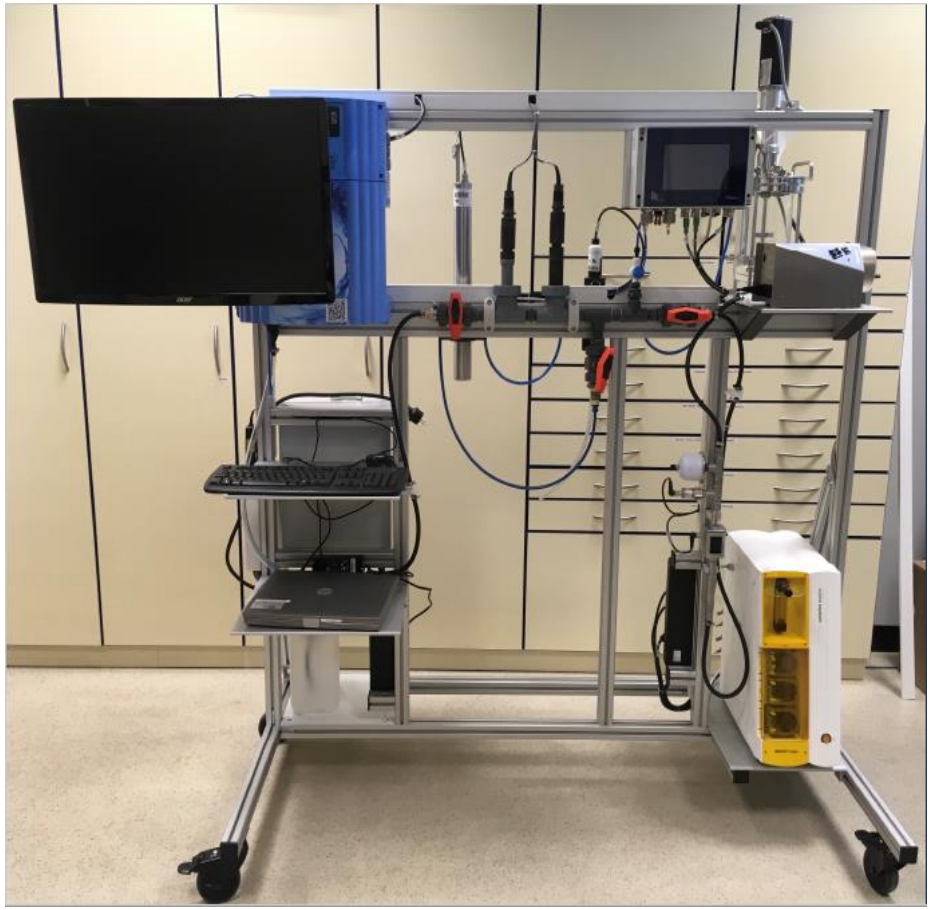
Jasmin Haap, Ph.D. thesis, Hohenstein, **2020**.

Screening: Mechanics and duration of washing have the strongest influence on the release of microplastics under defined lab washing conditions

Development of soft sensors for monitoring of rinsing quality in laundry processes

Systematic approach for development of predictive statistical models:

- 1) Collecting data of washing, rinsing and wastewater filtration processes by offline and online analytics
- 2) Screening of the most important data via Design of Experiment (i.e. Definitive Screening Design)
- 3) Evaluating statistical models via Kriging
- 4) Storage of data in PostGreSQL database
- 5) Transformation, Evaluation of data via Python
- 6) Set up machine learning algorithms in Python



I. Kogut, D. Gaida, E. Brohammer,
AiF project IMPRESS 20792N, 2019-2022

Clustering of data for sustainability evaluation of laundry performance

- In recent projects, sustainability related (mainly historical) data were collected and analyzed:

- By means of novel benchmarking methods

„Methodik des branchenunabhängigen ressourcen- und energiebezogenen Benchmarksystems, C. Mechel“, Ph.D. thesis, Hohenstein, 2017.

- And in the future, by innovative clustering methods & machine learning e.g.

B. Ferreira et al. “Monitoring sustainable development by means of earth observation data and machine learning: a review”, Environmental Sciences Europe volume 32, Article number: 120 (2020)

Outlook

In the future, "data science projects" and novel statistical approaches will gain more importance in laundry as well as textile industry and significantly influence the following areas:

- 1) Data-driven effective process optimization
- 2) Realizing "Industry of Things" projects
- 3) Analysis of big data
- 4) Effective Machine Learning Algorithms for enabling "smart laundries"
- 5) Accelerating digital transformation

Digital Transformation

HOHENSTEIN