

Wissenschaftliches Vorgehen: Der beste Weg zu guter Praxis

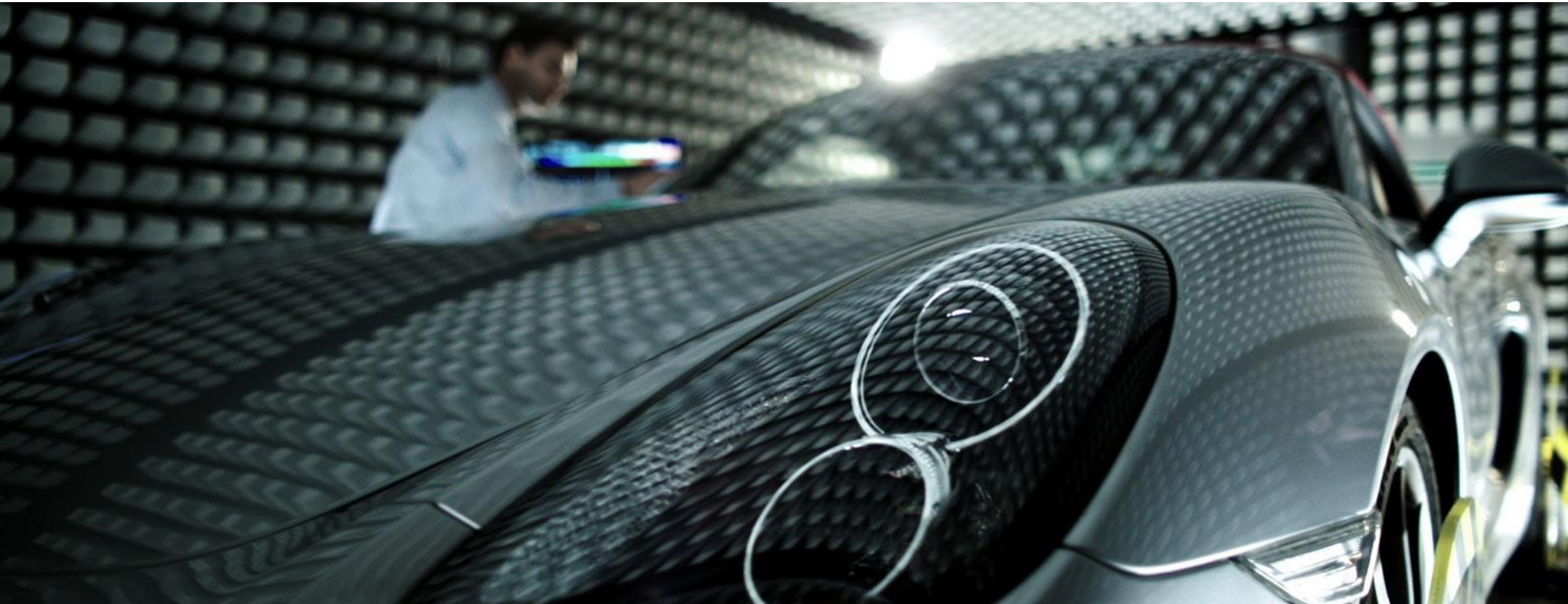
Thilo Stadelmann

BEST TEACHING – BEST PRACTICES:
VERMITTLUNG VON WISSENSCHAFTLICHKEIT

06. SEPTEMBER 2017, ZÜRICH, TONI-AREAL.

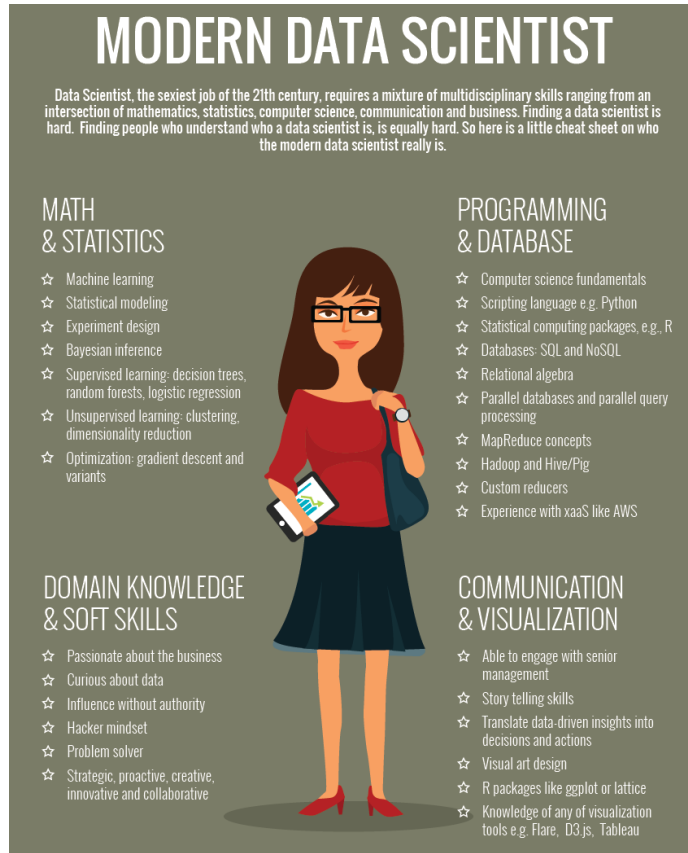


Wie entwickelt man ein Produkt?



Quelle: Produktentstehungsprozess – Vom Konzept zur Serie. <http://www.porscheengineering.com/peg/de/services/>

Wie entwickelt man ein Datenprodukt?



MODERN DATA SCIENTIST

Data Scientist, the sexiest job of the 21st century, requires a mixture of multidisciplinary skills ranging from an intersection of mathematics, statistics, computer science, communication and business. Finding a data scientist is hard. Finding people who understand who a data scientist is, is equally hard. So here is a little cheat sheet on who the modern data scientist really is.

MATH & STATISTICS

- ☆ Machine learning
- ☆ Statistical modeling
- ☆ Experiment design
- ☆ Bayesian inference
- ☆ Supervised learning: decision trees, random forests, logistic regression
- ☆ Unsupervised learning: clustering, dimensionality reduction
- ☆ Optimization: gradient descent and variants

PROGRAMMING & DATABASE

- ☆ Computer science fundamentals
- ☆ Scripting language e.g. Python
- ☆ Statistical computing packages, e.g. R
- ☆ Databases: SQL and NoSQL
- ☆ Relational algebra
- ☆ Parallel databases and parallel query processing
- ☆ MapReduce concepts
- ☆ Hadoop and Hive/Pig
- ☆ Custom reducers
- ☆ Experience with xaaS like AWS

DOMAIN KNOWLEDGE & SOFT SKILLS

- ☆ Passionate about the business
- ☆ Curious about data
- ☆ Influence without authority
- ☆ Hacker mindset
- ☆ Problem solver
- ☆ Strategic, proactive, creative, innovative and collaborative

COMMUNICATION & VISUALIZATION

- ☆ Able to engage with senior management
- ☆ Story telling skills
- ☆ Translate data-driven insights into decisions and actions
- ☆ Visual art design
- ☆ R packages like ggplot or lattice
- ☆ Knowledge of any of visualization tools e.g. Flare, D3.js, Tableau

Quelle: Marketing Distillery – the modern data scientist graphic.
http://www.marketingdistillery.com/wp-content/uploads/2014/11/mds_f.png

Wie entwickelt man ein Datenprodukt?

MODERN DATA SCIENTIST

Data Scientist, the sexiest job of the 21st century, requires a mixture of multidisciplinary skills ranging from an intersection of mathematics, statistics, computer science, communication and business. Finding a data scientist is hard. Finding people who understand who a data scientist is, is equally hard. So here is a little cheat sheet on who the modern data scientist really is.

MATH & STATISTICS

- ☆ Machine learning
- ☆ Statistical modeling
- ☆ Experiment design
- ☆ Bayesian inference
- ☆ Supervised learning: decision trees, random forests, logistic regression
- ☆ Unsupervised learning: clustering, dimensionality reduction
- ☆ Optimization: gradient descent and variants

PROGRAMMING & DATABASE

- ☆ Computer science fundamentals
- ☆ Scripting language e.g. Python
- ☆ Statistical computing packages, e.g., R
- ☆ Databases: SQL and NoSQL
- ☆ Relational algebra
- ☆ Parallel databases and parallel query processing
- ☆ MapReduce concepts
- ☆ Hadoop and Hive/Pig
- ☆ Custom reducers
- ☆ Experience with xaaS like AWS

DOMAIN KNOWLEDGE & SOFT SKILLS

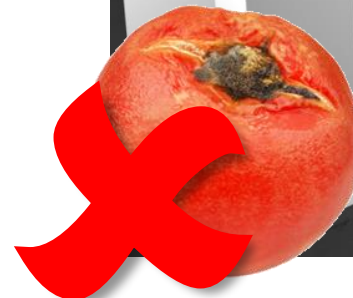
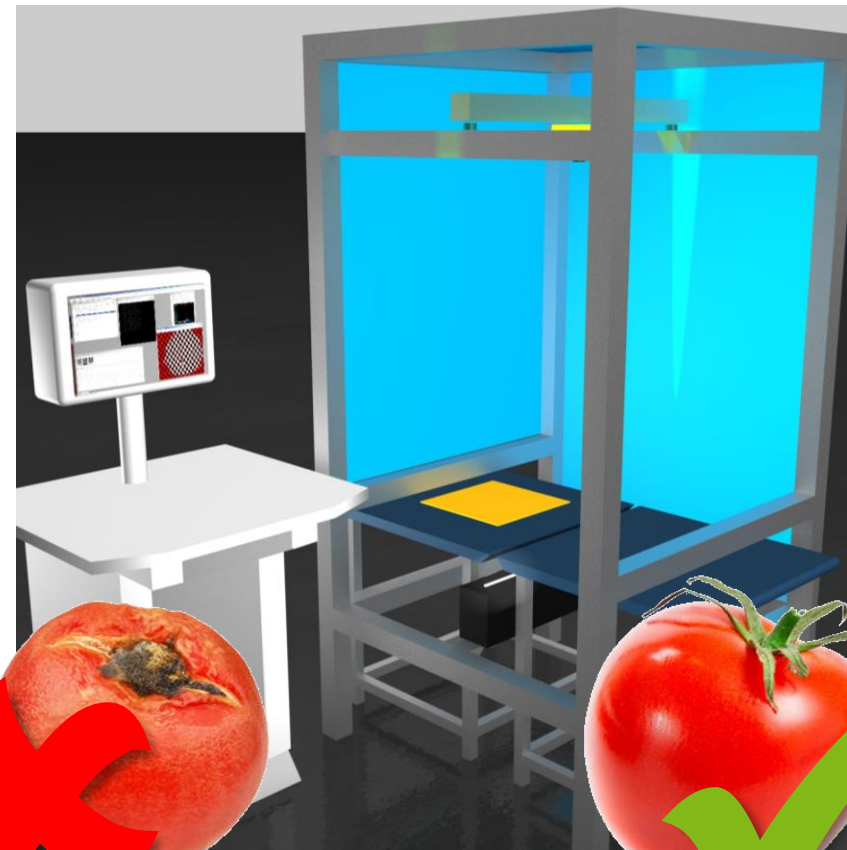
- ☆ Passionate about the business
- ☆ Curious about data
- ☆ Influence without authority
- ☆ Hacker mindset
- ☆ Problem solver
- ☆ Strategic, proactive, creative, innovative and collaborative

COMMUNICATION & VISUALIZATION

- ☆ Able to engage with senior management
- ☆ Story telling skills
- ☆ Translate data-driven insights into decisions and actions
- ☆ Visual art design
- ☆ R packages like ggplot or lattice
- ☆ Knowledge of any of visualization tools e.g. Flare, D3.js, Tableau

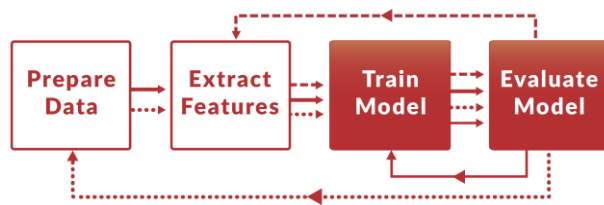
Quelle: Marketing Distillery – the modern data scientist graphic.
http://www.marketingdistillery.com/wp-content/uploads/2014/11/mds_f.png

Quelle: Industrielle Bildverarbeitung – ...ich sehe was, was du nicht siehst.
<http://www.fraenz-jaeger.de/index.php/de/technologien-de/industrielle-bildverarbeitung>



Schlüssel: Die Wissenschaftliche Methode

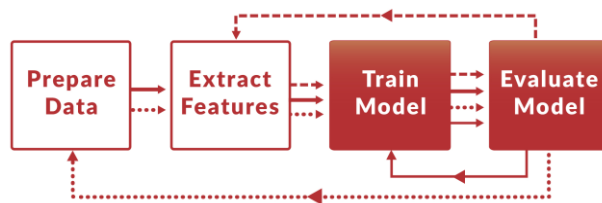
Maschinelles Lernen in der Praxis



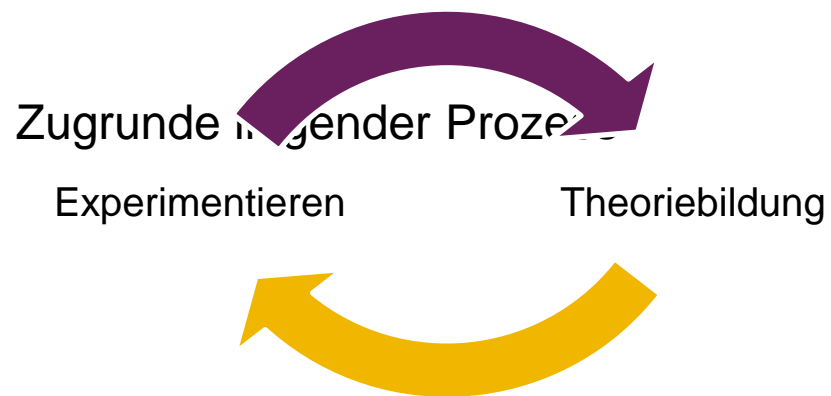
Quelle: data-intelligence.ai

Schlüssel: Die Wissenschaftliche Methode

Maschinelles Lernen in der Praxis

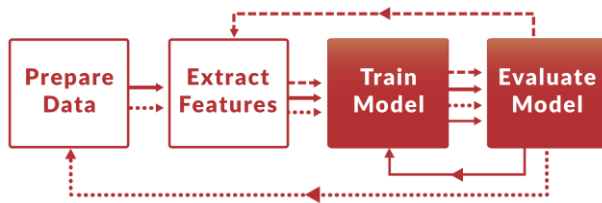


Quelle: data-intelligence.ai

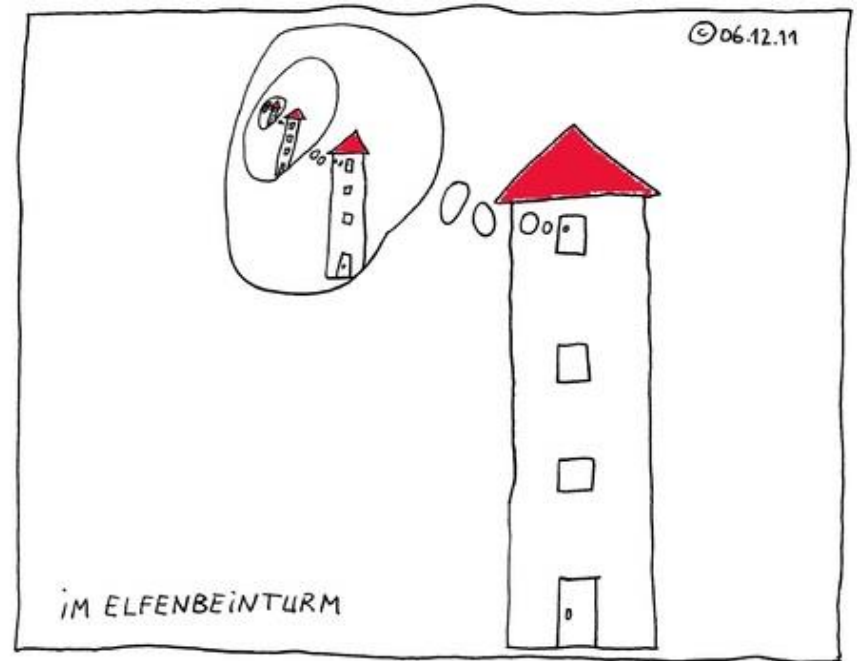
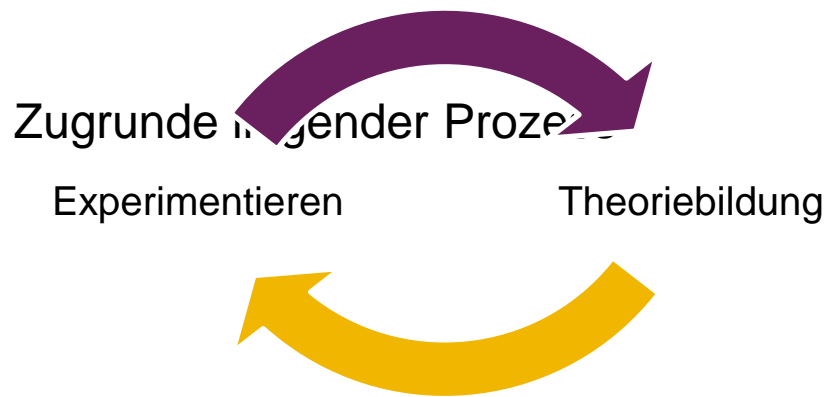


Schlüssel: Die Wissenschaftliche Methode

Maschinelles Lernen in der Praxis



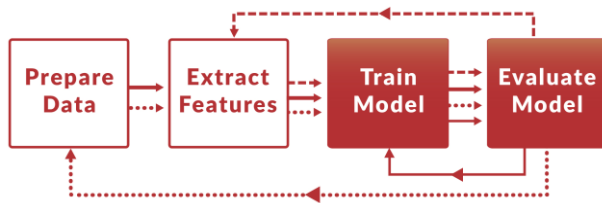
Quelle: data-intelligence.ai



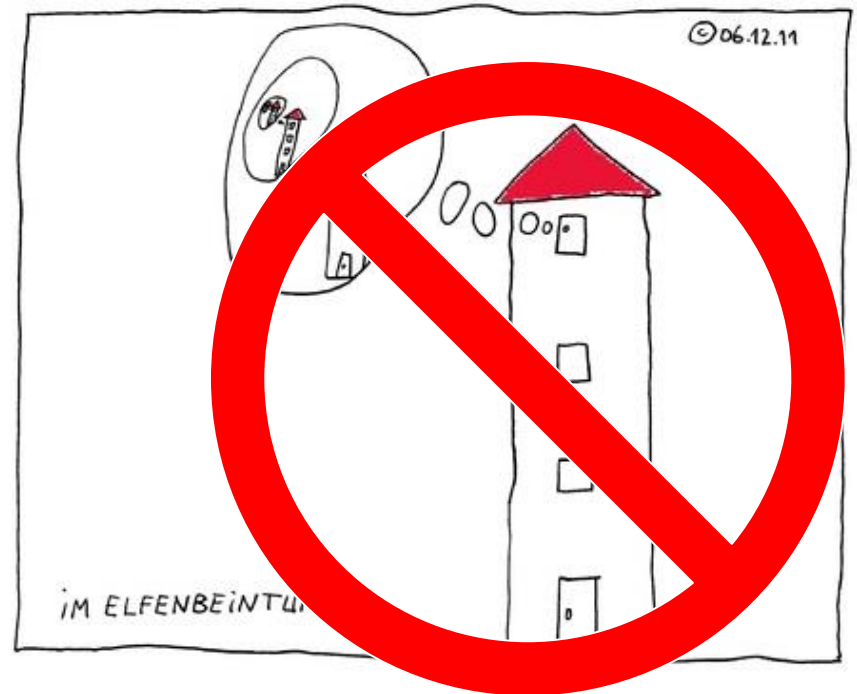
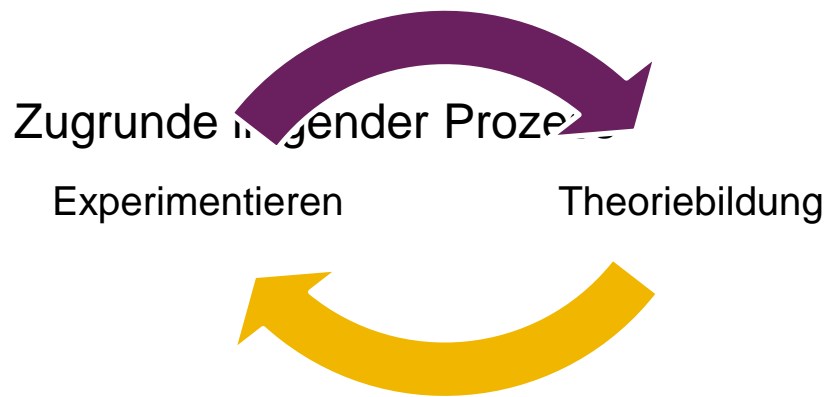
Quelle: Elfenbeinturm by Müller.
https://www.toonpool.com/cartoons/Elfenbeinturm_173927

Schlüssel: Die Wissenschaftliche Methode

Maschinelles Lernen in der Praxis



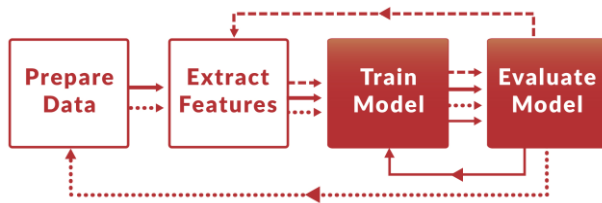
Quelle: data-intelligence.ai



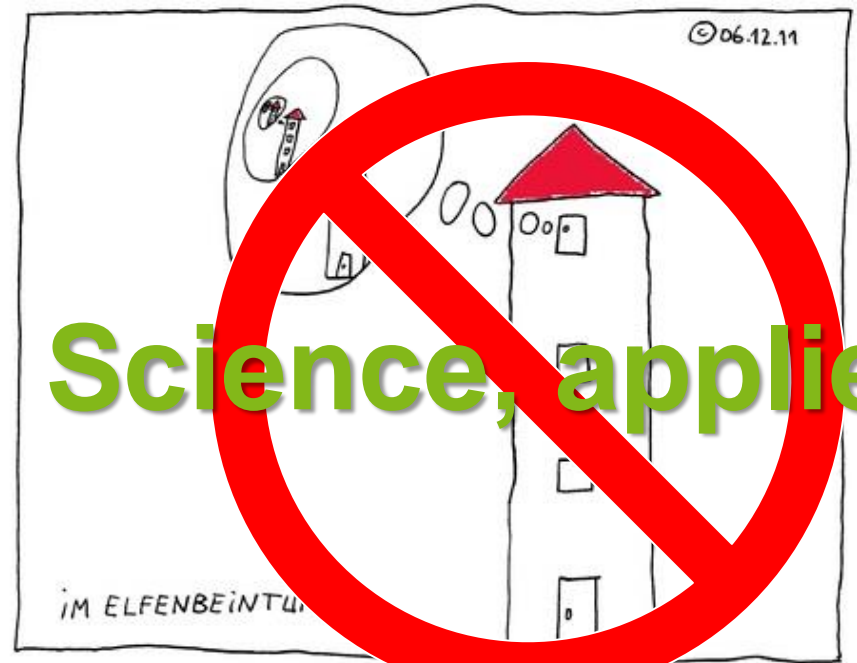
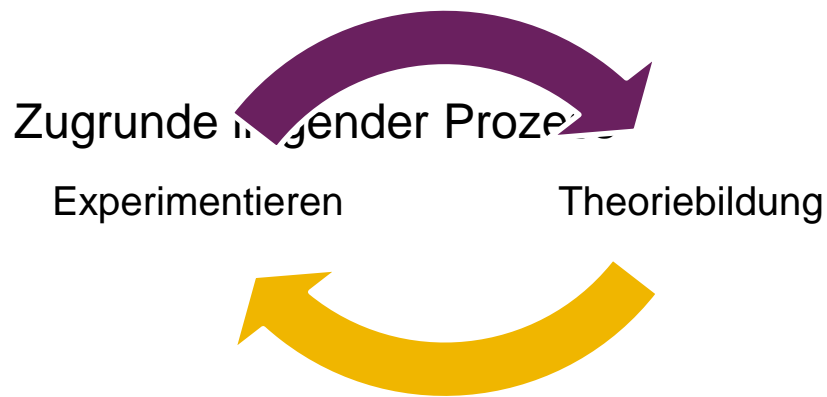
Quelle: Elfenbeinturm by Müller.
https://www.toonpool.com/cartoons/Elfenbeinturm_173927

Schlüssel: Die Wissenschaftliche Methode

Maschinelles Lernen in der Praxis



Quelle: data-intelligence.ai



Quelle: Elfenbeinturm by Müller.
https://www.toonpool.com/cartoons/Elfenbeinturm_173927

Schlüsselkompetenzen (nicht nur) für Datenwissenschaftler

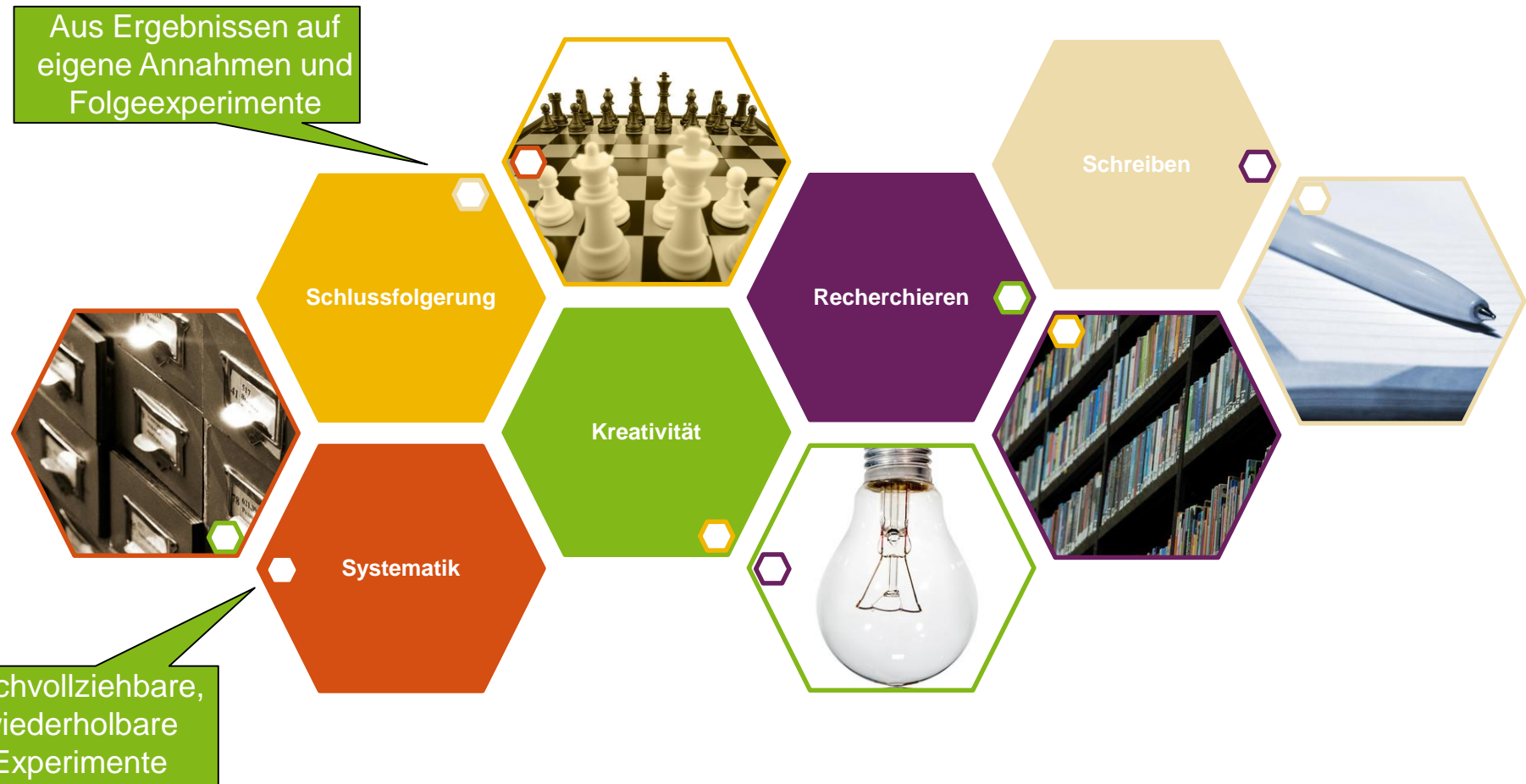


Schlüsselkompetenzen (nicht nur) für Datenwissenschaftler



Nachvollziehbare,
wiederholbare
Experimente

Schlüsselkompetenzen (nicht nur) für Datenwissenschaftler



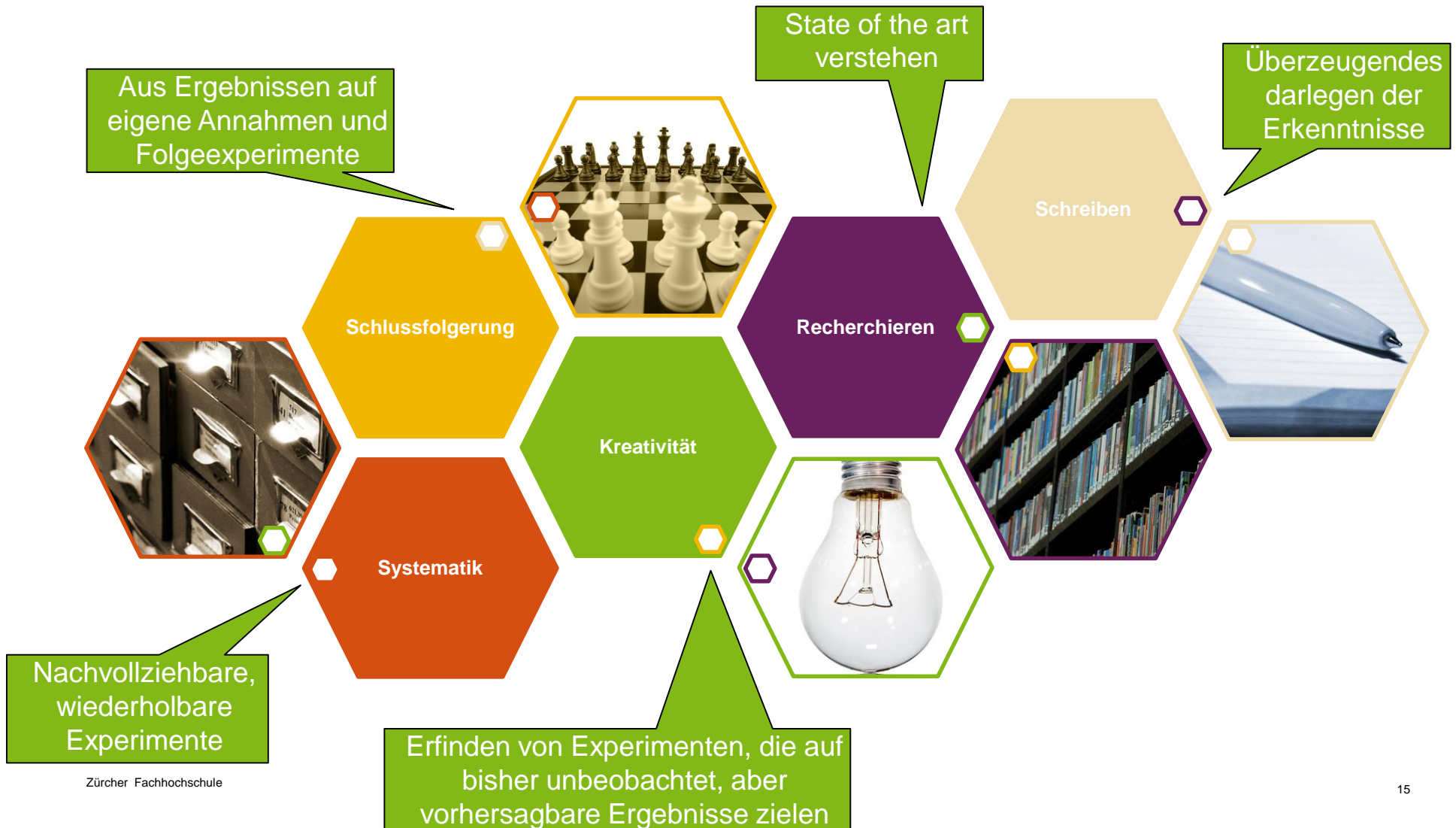
Schlüsselkompetenzen (nicht nur) für Datenwissenschaftler



Schlüsselkompetenzen (nicht nur) für Datenwissenschaftler



Schlüsselkompetenzen (nicht nur) für Datenwissenschaftler



Fazit

Das Arbeiten anhand der wissenschaftlichen Methode erlaubt die Entwicklung relevanter und praxistauglicher Data Products

Applied Science ist also das Anwenden der wissenschaftlichen Methode zur Lösung eines neuen, konkreten Use-Cases

zhaw School of Engineering
Speaker identification and clustering using convolutional neural networks
 Yanick Lukic, Carlo Vogt, Oliver Durr, and Thilo Stadelmann
 ZHAW Databab, Zurich University of Applied Sciences, Winterthur, Switzerland

Problem Statement

- Speaker recognition performance by machine vs. by humans [2]
- Extensive clustering performance across of background noise than identifier performance
- Improves voice speaker recognition performance on both tasks by using learned features instead of hand-crafted
- Identifies speakers with high accuracy
- Less time consuming with smaller number of speakers

Approach

Feature learning (identification training) from raw features by using:

- 128 feature outputs arranged in 16x16x16 volume
- 16 neural net layers (preprocessing, feature learning)
- 100000 iterations by speaker

Applications (clustering test):

- Fast speaker identification (CNN as inferencer)
- Very large number of speakers
- Large number of speakers
- Large number of speakers
- Large number of speakers
- Large number of speakers
- Large number of speakers
- Large number of speakers
- Large number of speakers
- Large number of speakers

Experimental Setup

- Classical baseline [20] (DNN)
- Architecture of CNN from [21]
- Evaluation by speaker
- Classification rate table
- Implementation in python
- 1.1s/c.s.f. for speaker
- Pre-trained CNN for clustering
- Clustering based on speaker embedding

Results

	1000	10000	100000
1.5 error	0.80	0.78	0.78
1.0 error	0.10	0.08	0.08
0.5 error	0.00	0.00	0.00

References

This work has been supported by the Bachelor's programme 6, the Institute of Applied Information Technology 107, and the Institute of Data Analysis and Process Design (DAP) 2018/2019. Thank you!

2017 IEEE INTERNATIONAL WORKSHOP ON MACHINE-LEARNING FOR SIGNAL PROCESSING, SEPT. 25- 28, TOKYO, JAPAN

LEARNING EMBEDDINGS FOR SPEAKER CLUSTERING BASED ON VOICE EQUALITY

Yanick X. Lukic, Carlo Vogt, Oliver Durr, and Thilo Stadelmann
 Zurich University of Applied Sciences, Winterthur, Switzerland

ABSTRACT

Recent work has shown that convolutional neural networks (CNNs) trained in a supervised fashion for speaker identification are able to extract features from spectrograms which can be used for speaker clustering. These features are represented by the activations of a certain hidden layer and are called embeddings. However, previous approaches require plenty of additional speaker data to learn the embedding, and although the clustering results are often on par with more traditional approaches using MFCC features etc., room for improvement stems from the fact that these embeddings are trained with a surrogate task that is rather far away from segregating unknown voices - namely, identifying few specific speakers. We address both problems by training a CNN to extract embeddings that are similar for equal speakers (regardless of their specific identity) using weakly labeled data. We demonstrate our approach on the well-known TIMT dataset that has often been used for speaker clustering experiments in the past. We extend the clustering performance of all previous approaches, but require just 100 instead of 500 annotated speakers to learn an embedding suited for clustering.

Index Terms— Speaker Clustering, Speaker Recognition, Convolutional Neural Network, Speaker Embedding

1. INTRODUCTION

Speaker clustering handles the “who spoke when” challenge in a given audio recording without knowing how many and which speakers are present in the audio signal. It is called speaker duration when the task of segmenting the audio stream into speaker-specific responses is handled simultaneously [1]. The problem of speaker clustering is omnipresent in digital audio archiving, e.g., recordings of lectures, conferences or debates [2]. For their quantitative indexing, automatic extraction of figures like the number of speakers or talk time per person is important. This further facilitates automatic transcripts using existing speaker recognition procedures, based on the accurate automatic assignment of speech utterances to groups that each represent a (previously unknown) speaker.

The lack of knowledge of the number and identity of speakers leads to a much more complex problem compared to

2. LEARNING SPEAKER DISSEMBLITY

2.1. Related work

The design of CNNs makes it possible to recognize patterns in minimally preprocessed digital images or other data with

zhaw School of Engineering
Fully Convolutional Neural Networks for Newspaper Article Segmentation
 Benjamin Meier, Thilo Stadelmann, Jan Stampf, Marek Arnold, and Mark Ciallebach
 ZHAW Databab, Zurich University of Applied Sciences, Winterthur, Switzerland

Problem Statement

- In order to open up large volumes of classical print media (e.g. newspapers) to content-based information retrieval, digitized print media pages have to be segmented into their semantically corrected components, i.e. articles.
- The idea is to use machine learning methods to create a model that is able to learn how to segment newspaper pages into articles.
- Therefore no static rules have to be defined.
- The plain newspaper page / image is given and also a OCR preprocessed page on which text, lines and images are marked.

Approach

Semantic Segmentation based on Fully Convolutional Networks (FCNN) [1] Using a grayscale image of the newspaper page as input and also a OCR preprocessed version of the page

- The network classifies each pixel as background or foreground
- The classification map is used to extract polygons which are finally the articles
- there are many labeled pages required to train the neural network
- our dataset contains only about 3'000 newspaper pages and only 400 pages are fully labeled

Implementation

- Architecture based on FCNN
- Evaluation by F1 score
- error score (best value, higher values are worse) and
- Completeness score (fraction of exactly detected articles; lowest value, lowest value)
- implementation in python using Jupyter notebooks
- training from scratch about 3h
- Competition with a patch based / sliding window CNN baseline
- Increased quality by 48.2% (F1 score)
- Increased quality by a factor of 2.6 (Compl. score)
- Classification time improved by a factor of 35

References

This work has been supported by the Master's programme 7 (MSE), the Institute of Applied Information Technology 107 and the Institute of Data Analysis and Process Design (DAP) 2018/2019. Thank you!



APPENDIX

What is Data Science?

Enables Data Products
 → Applied Science
 → Interdisciplinary

Data Science := "Unique blend of skills from analytics, engineering & communication aiming at generating value from the data itself [...]"
 (ZHAW Datalab*)

