

Smart Product Data in der High-Tech-Industrie



Prof. Dr. Alexander Holland

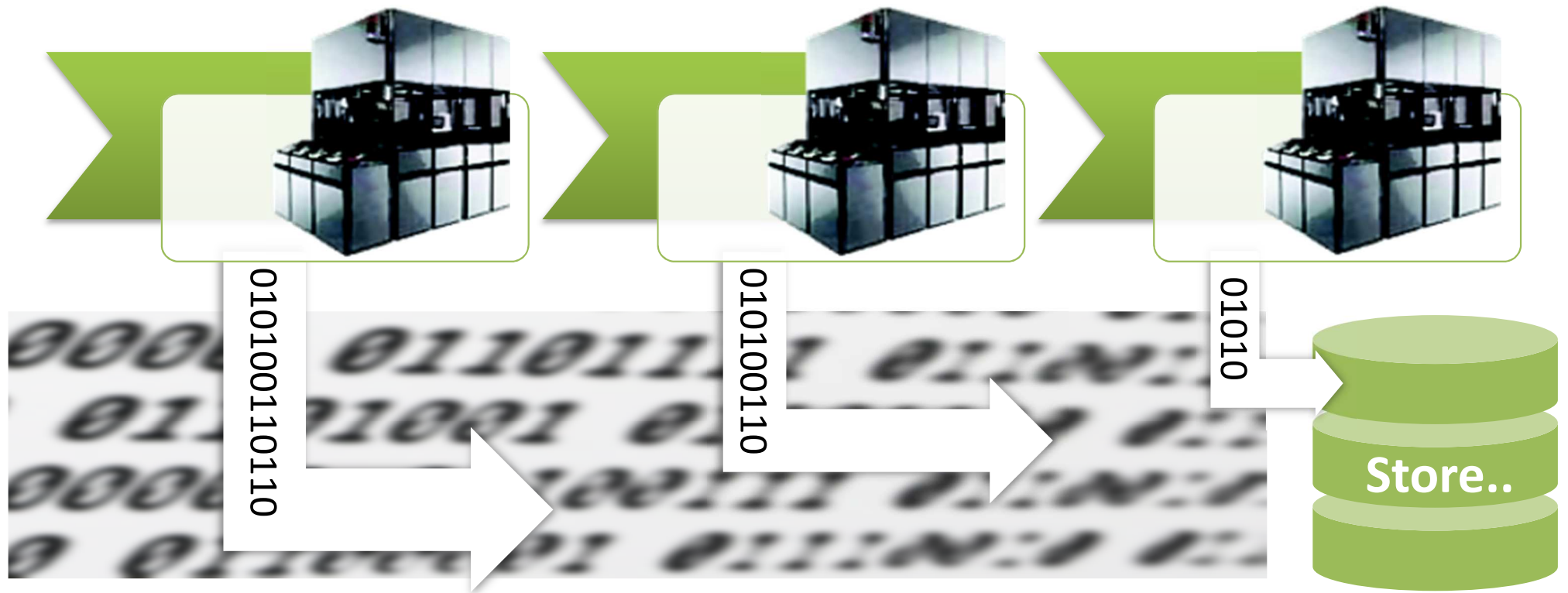
*FOM Hochschulzentrum Dortmund, Fachbereich
Wirtschaftsinformatik*

Christian Weber & Johannes Zenkert

*Universität Siegen, Institut für Wissensbasierte Systeme und
Wissensmanagement*

Datenzuwachs in Produktionsumgebungen

Stetig wachsende Datenmengen..
Aber wie handhaben?



Big Data Volume: the steady growth of data

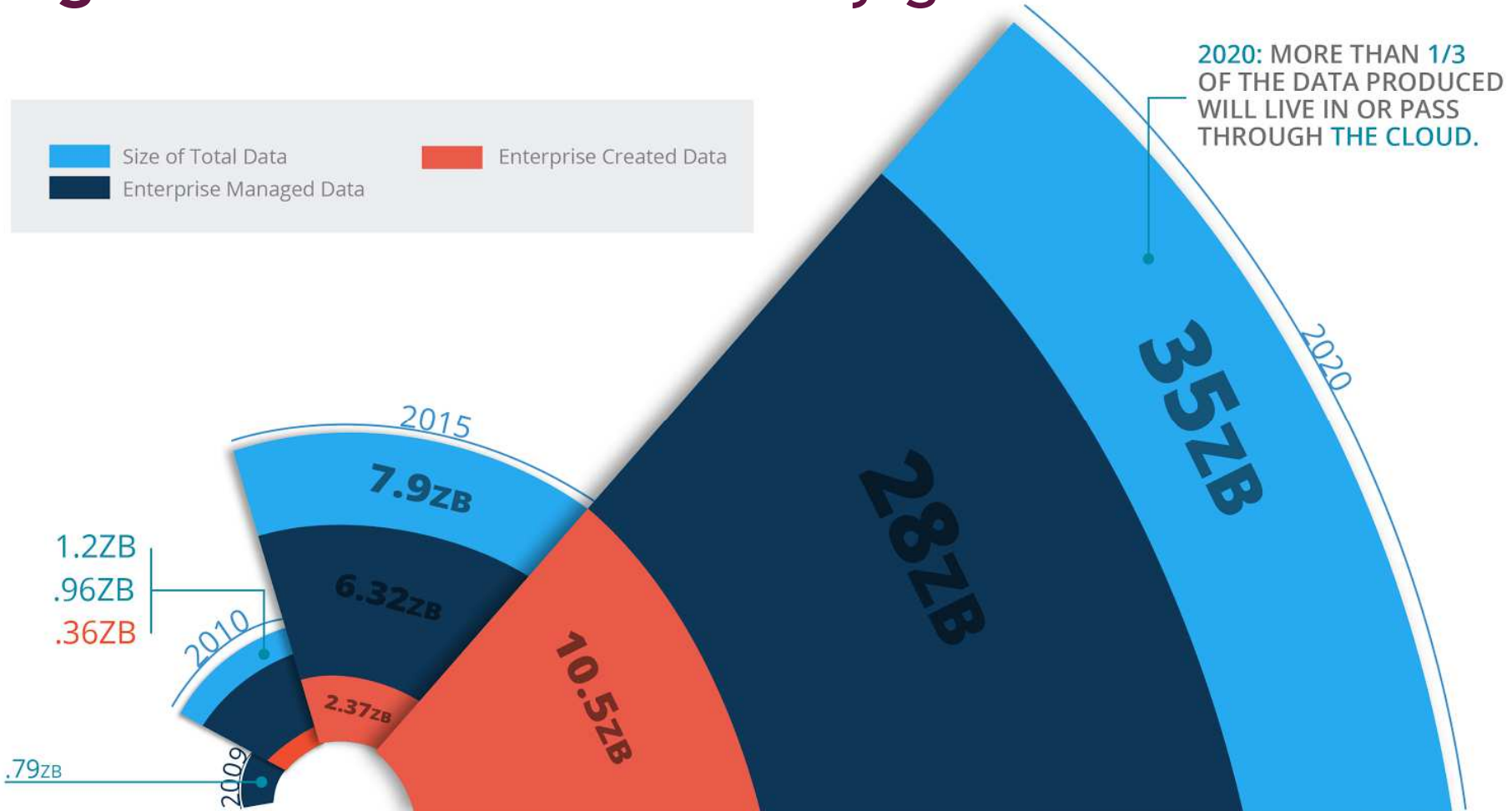


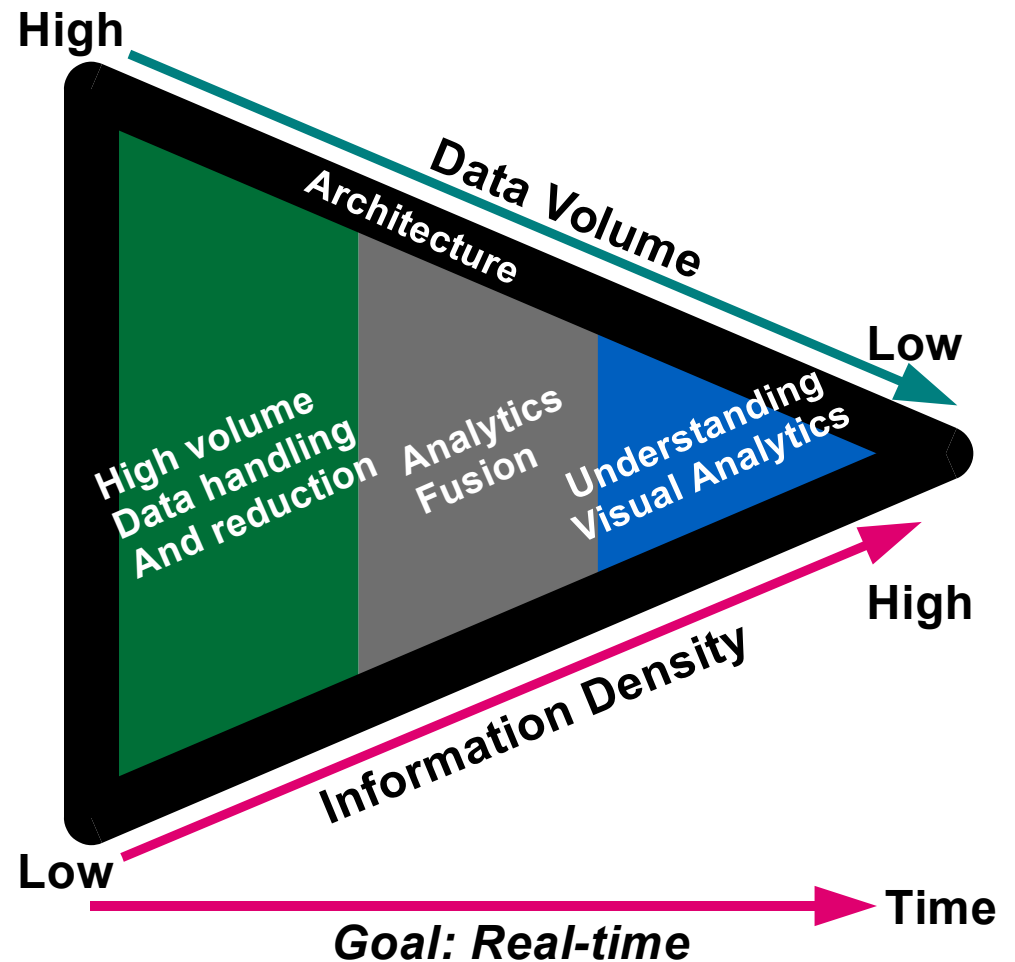
Image from here: <https://www.impactradius.com/blog/7-vs-big-data/>

Original from here: http://www.csc.com/insights/flxwd/78931-big_data_universe_beginning_to_explode

Data Intensive Computing

Phasen der Datenverarbeitung

- **Reduktion**
 - Auf Basis von Analysen
- **Prozessieren und Fusionieren**
 - Potentiell mit “large scale distributed computing”
- **Data Understanding**
 - Auswerten und Visualisieren

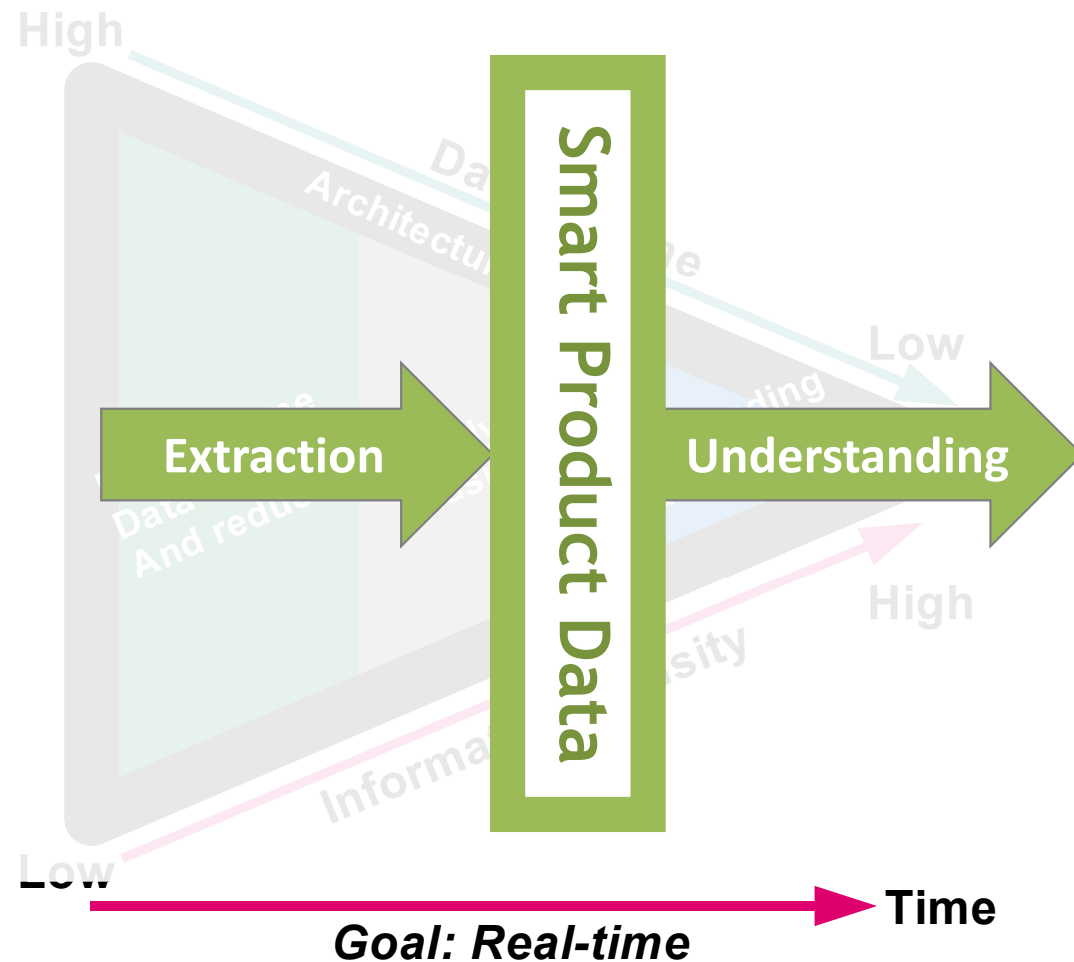


Gorton 2008, Software Architecture Challenges for Data Intensive Computing

Data Intensive Computing

Phasen der Datenverarbeitung

- **Reduktion**
 - Auf Basis von Analysen
- **Prozessieren und Fusionieren**
 - Potentiell mit “large scale distributed computing”
- **Data Understanding**
 - Auswerten und Visualisieren



Gorton 2008, Software Architecture Challenges for Data Intensive Computing

Big Data Landscape 2016 (Version 3.0)

Viele Komponenten am Markt..

Passende Lösungen
für KMUs möglich?



1: <http://mattturck.com/2016/02/01/big-data-landscape/>

Was ist die Ausgangssituation?

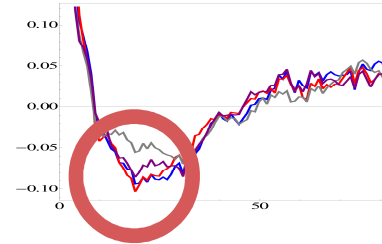


B	B	B	Fehler
		...	1
	...		0
...			1

Auch in Zeiten von Big Data maßgeschneiderte
Lösungen für KMUs?



WISSENSTRANSFER AUS DER FORSCHUNG IN DIE PRAXIS

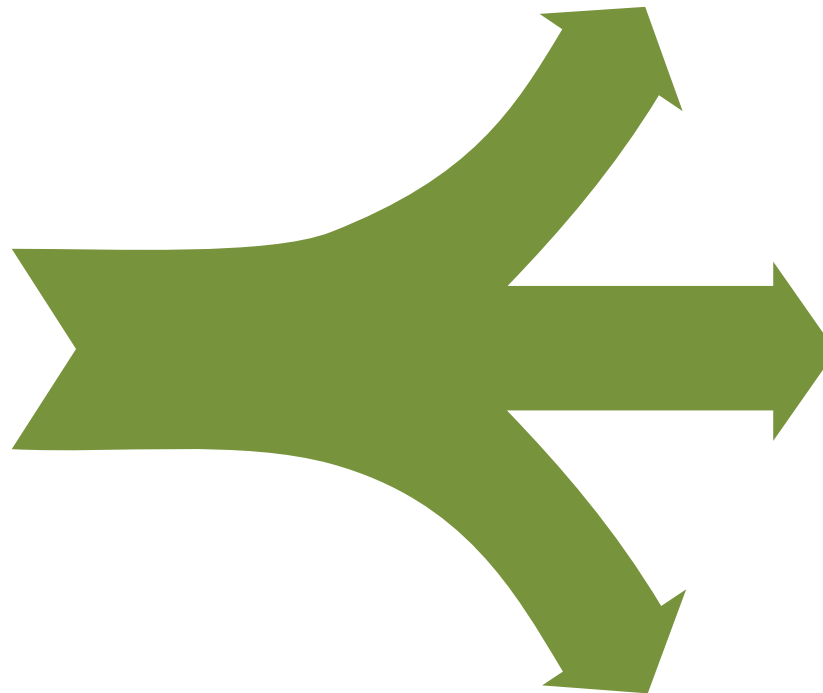


Effekte identifizieren

Ein Effekt -
zeitlich **begrenzt**.

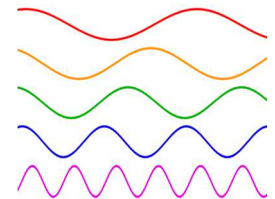
3 Sichten

Was entnehmen wir
unseren Daten?



Schwankungen identifizieren

Ein Effekt -
wiederkehrend.



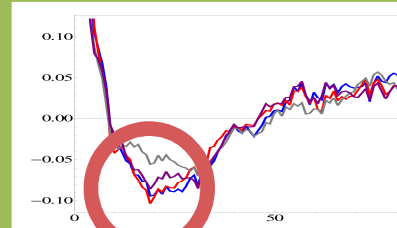
Effekte visualisieren

Effekte, Ergebnisse,
Zusammenhänge -
sichtbar machen.



3 Sichten

Was entnehmen wir unseren Daten?



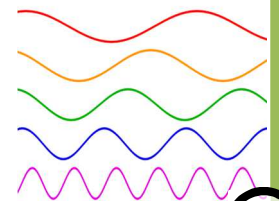
Effekte identifizieren

Ein Effekt -
zeitlich **begrenzt**.

1

Schwankungen identifizieren

Ein Effekt -
wiederkehrend.

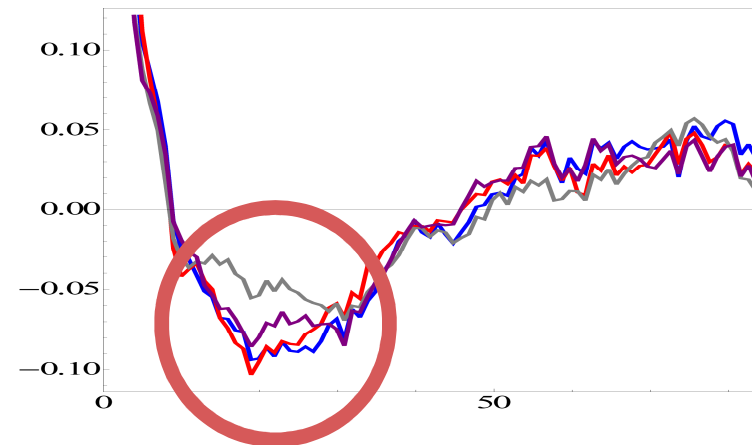


2

Effekte visualisieren

Effekte, Ergebnisse,
Zusammenhänge -
sichtbar machen.





Ein Effekt - zeitlich begrenzt

① EFFEKTE IDENTIFIZIEREN UND UNTERSUCHEN MIT KÜNSTLICHER INTELLIGENZ

Problemstellung

Produktionsevents
werden von **großen
Datenmengen**
beschrieben (pot.
“Big Data” Mengen).

Wie lassen sich Events
universell und flexibel auf
wenige aussagekräftige
Parameter herunter
brechen?

Lösung

**Flexible Datenextraktion
und Speicherung.**

Einsatz einer
“**Feature Selection**”
zur langfristigen
“**Feature Extraction**”.

Prüfen der Signifikanz mit
einem **intelligenten
Klassifikator.**

Die Drei Säulen der Analyse

Konzipieren



**User Interface mit
Job-Tickets
für Auswertungen**

Extrahieren



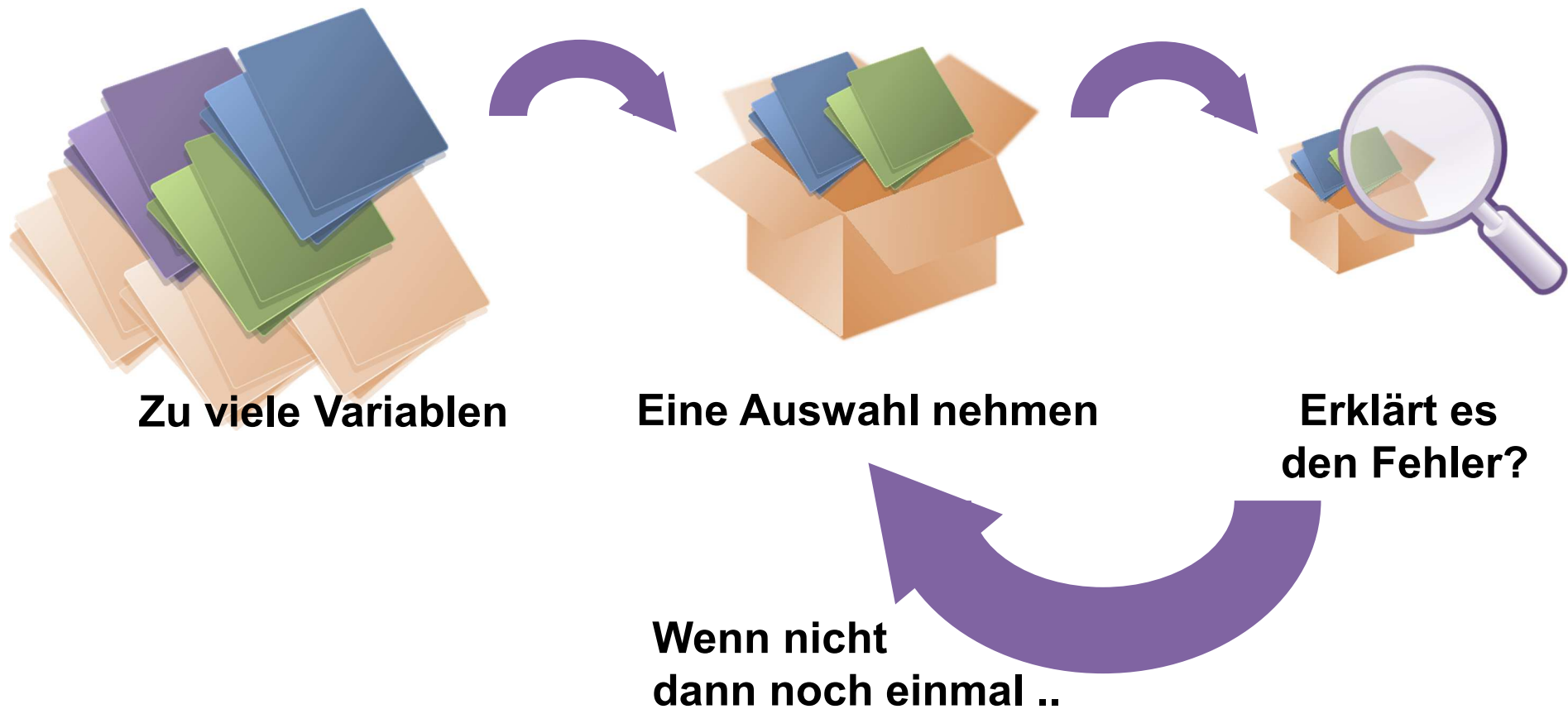
**Data Warehouse
oder Cloud für
aufbereitete Daten**

Auswerten



**Feature Selection
und Evaluierung
mit Reporting**

Daten reduzieren - Feature Selection



Daten reduzieren - Fehler identifizieren



**Erklärt es
den Fehler?**

Wenn ja ..

B	B	Fehler
		1
		0
		1

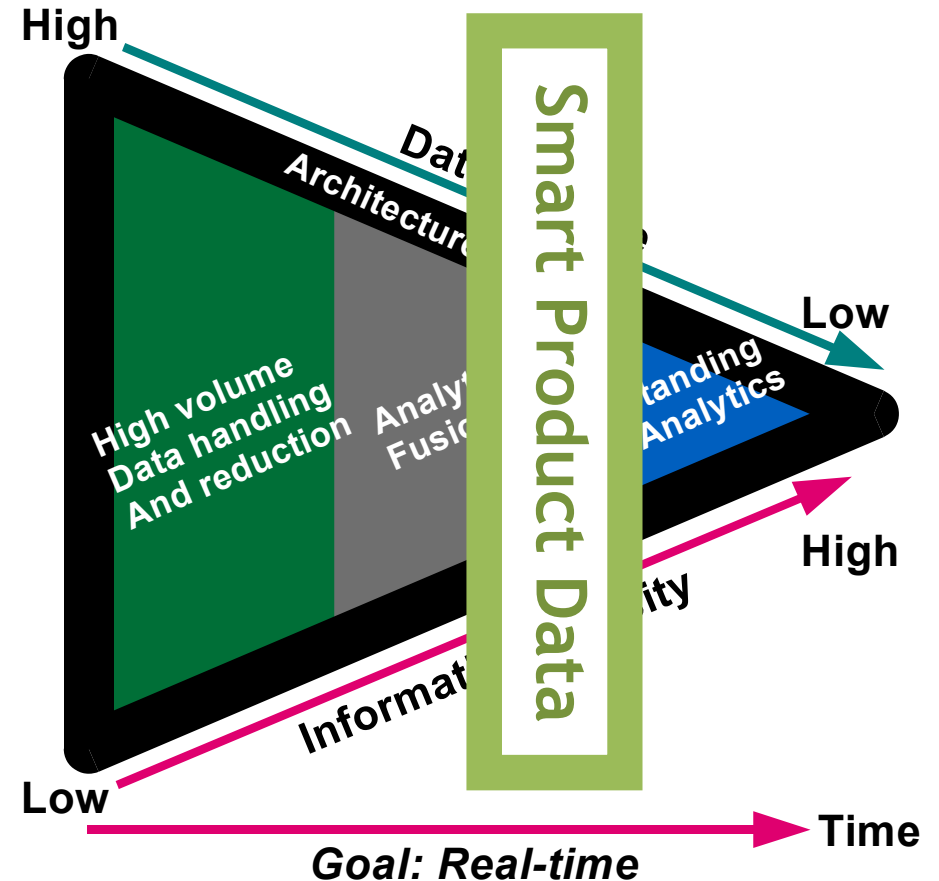
**Fehler
identifizieren /
klassifizieren..**

Daten reduzieren - Fehler identifizieren



**Erklärt es
den Fehler?**

Wenn ja ..





Feature Selection in der High-Tech Industrie

EINE BEISPIELLÖSUNG AUS DER PRAXIS



	SetId	Result Column Name	Rows	Columns
Show result	1212	YIELD	7410	152

Parameter Overview List:

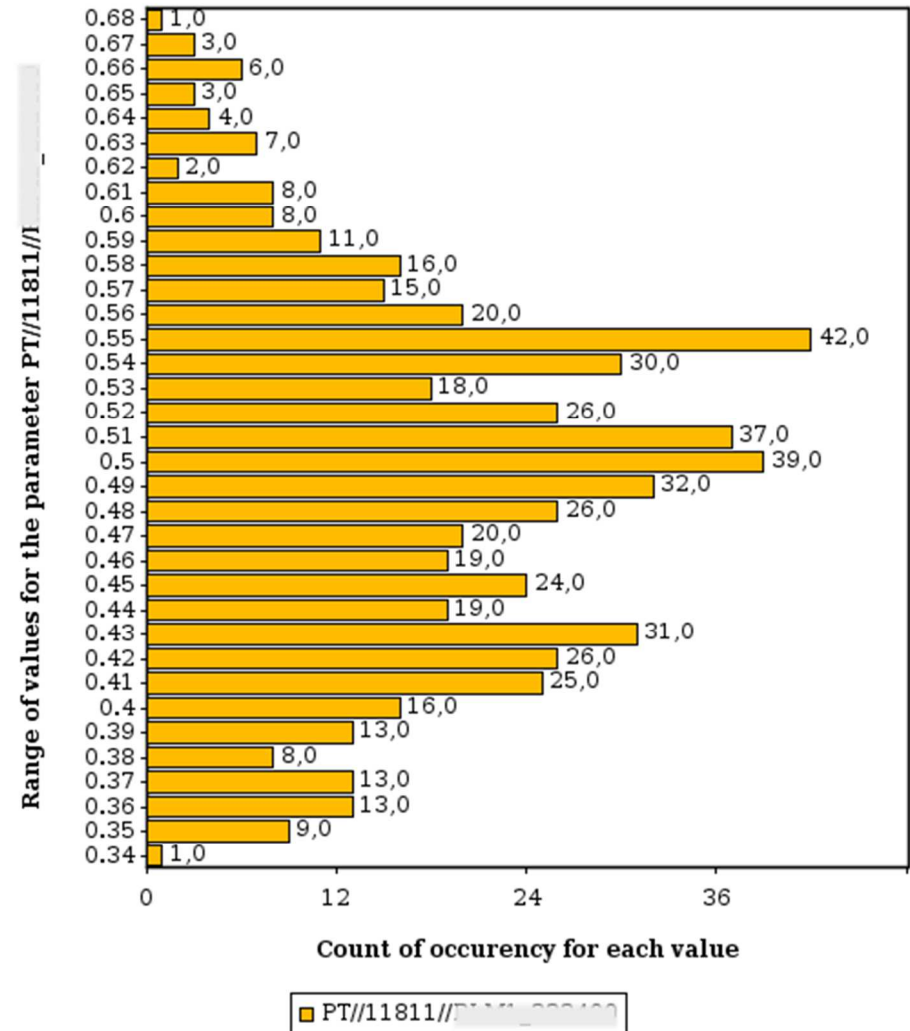
	Parameter Name	Original Number of Occurrence
Show	PT//11810//	1379
Show	PT//11811//	591
Show	PT//5187//	1629
Show	PT//5251//	1729
Show	PT//5252//	1729
Show	PT//5253//	1729
Show	PT//5254//	1729
Show	PT//5255//	1729
Show	PT//5256//	1729
Show	PT//5257//	1729

Data Normalization Overview of value

Shown Norm Value	Original Value
0.68	58.737449999999995
0.67	58.425743333333333
0.66	58.114036666666664
0.65	57.80233

Result Column Histogram

Histogram overview of the parameter PT//11811//





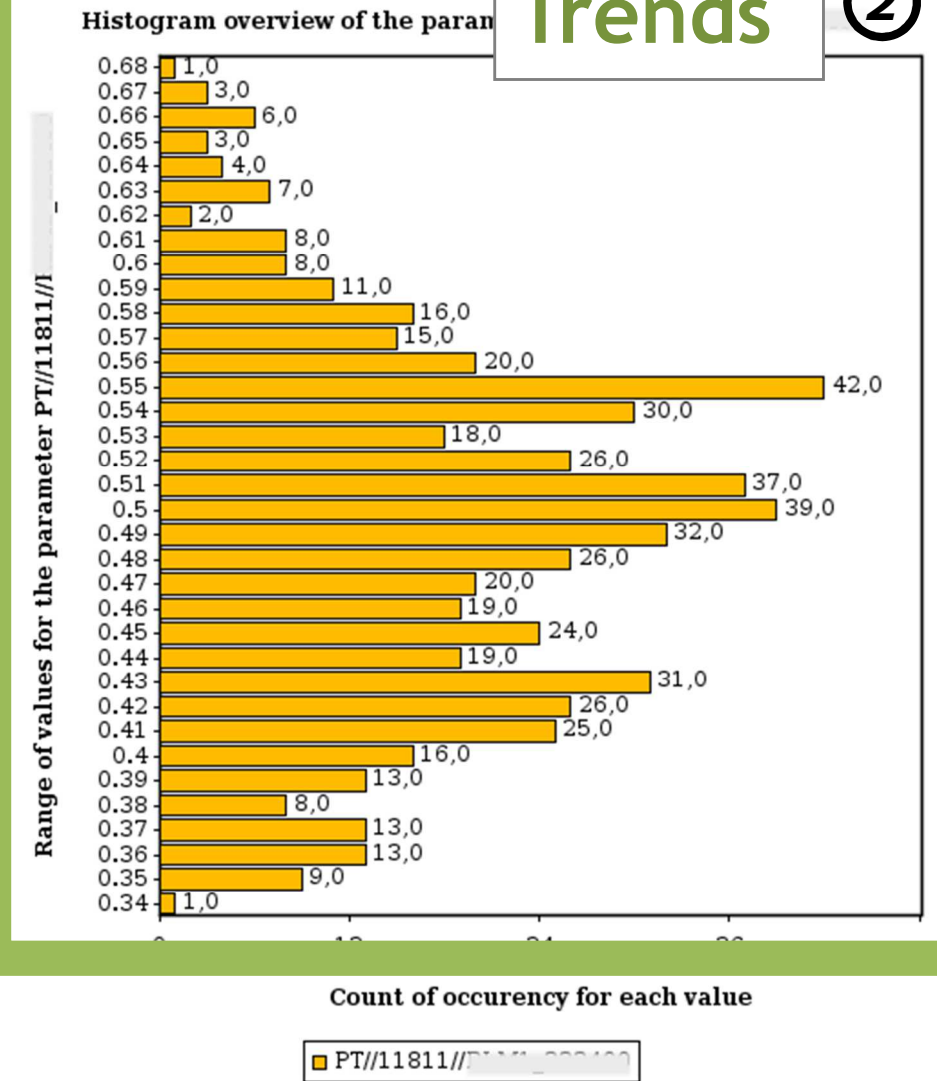
	SetId	Result Col	Events ①
Show result	1212	YIELD	

	Parameter Name	Original Number of Occurrence
③	Pot. Features	
Show	PT//11811//I	591
Show	PT//5187//I	1629
Show	PT//5251//I	1729
Show	PT//5252//I	1729
Show	PT//5253//I	1729
Show	PT//5254//I	1729
Show	PT//5255//I	1729
Show	PT//5256//I	1729
Show	PT//5257//I	1729

Data Normalization Overview of value

Shown Norm Value	Original Value
0.68	58.737449999999995
0.67	58.425743333333333
0.66	58.114036666666664
0.65	57.80233

Result Column Histogram **Trends ②**



Reduktion und Klassifizierung

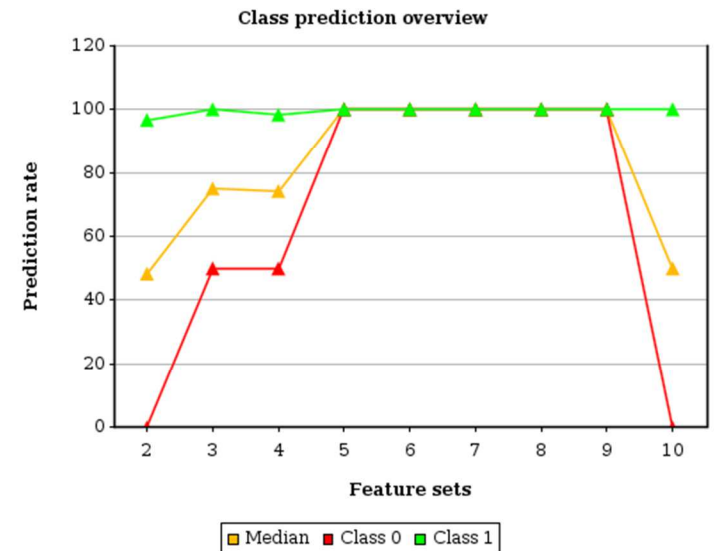
Analysis result overview per number of features :

	Feature Set with number of elements	Feature Selection internal criterion value	Average of predicted classes while empty classes are taken as 0
Show feature set	2	0.846154	48.33
Show feature set	3	0.846154	75.0
Show feature set	4	0.846154	74.16
Show feature set	5	0.846154	100.0
Show feature set	6	0.846154	100.0
Show feature set	7	0.846154	100.0
Show feature set	8	0.846154	100.0
Show feature set	9	0.846154	100.0
Show feature set	10	0.846154	50.0

Analysis result for the selected Feature Set 5:

Class	Upper class border	Class description	Member vectors in the class	Correct predicted vectors for the class	Prediction rate
0	1	High	2	2	100
1	0.5	Low	60	60	100

Class prediction overview:



Members of the feature set:

PT//10917// _____), PT//10919// _____),
 PT//11013// _____), PT//11024// _____),
 PT//12131// _____)

Staying the same:

PT//10919// _____), PT//11013// _____),
 PT//11024// _____), PT//12131// _____),

New in this set:

PT//10917// _____),

Removed since the last set:

Reduktion und Klassifizierung

Analysis result overview per number of features :

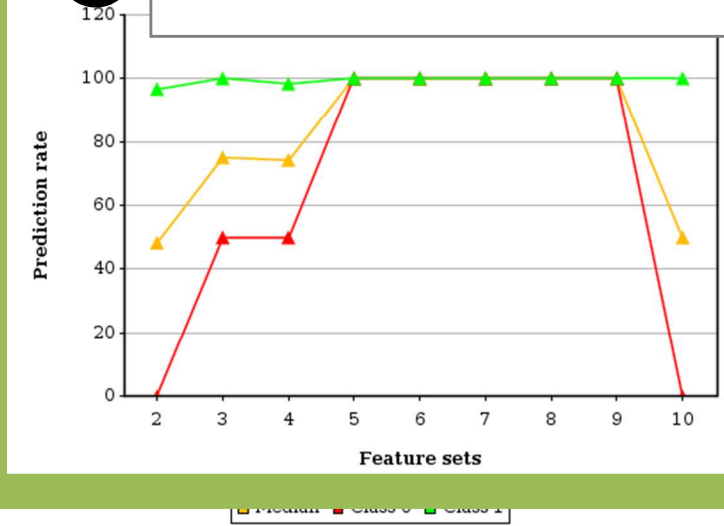
① Features Evaluieren

Show feature set			
Show feature set	3	0.846154	75.0
Show feature set	4	0.846154	74.16
Show feature set	5	0.846154	100.0
Show feature set	6	0.846154	100.0
Show feature set	7	0.846154	100.0
Show feature set	8	0.846154	100.0
Show feature set	9	0.846154	100.0
Show feature set	10	0.846154	50.0

Analysis result for the selected Feature Set 5:

Class	Upper class border	Class description	Member vectors in the class	Correct predicted vectors for the class	Prediction rate
0	1	High	2	2	100
1	0.5	Low	60	60	100

② Beste Resultate



Members of the feature set:

PT//10917// , PT//10919// ,
 PT//11013// , PT//11024// ,
 PT//12131//

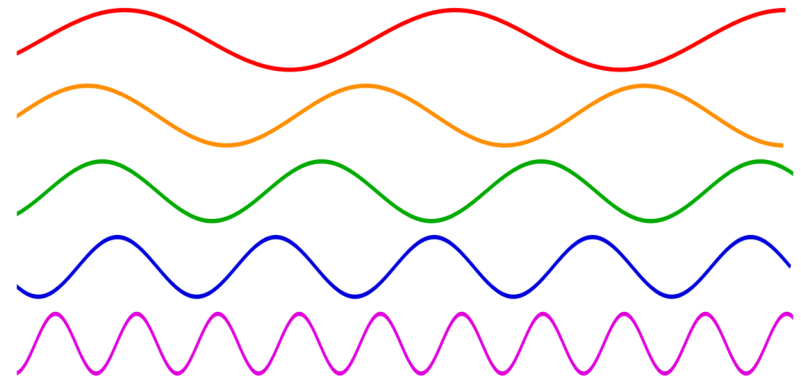
Staying the same:

PT//10919// , PT//11013// ,
 PT//11024//

New in the set:

PT//10917//

③ Feature Sets



Ein Effekt - wiederkehrend.

② SCHWANKUNGEN IDENTIFIZIEREN UND ERKLÄREN DURCH FREQUENZANALYSEN

Problemstellung

Wiederkehrende Effekte
beeinflussen die
Variation von
Produktionsparameter,
die Quelle ist unklar.

Wie lassen sich
wiederkehrende Effekte
isolieren und sichtbar
machen?

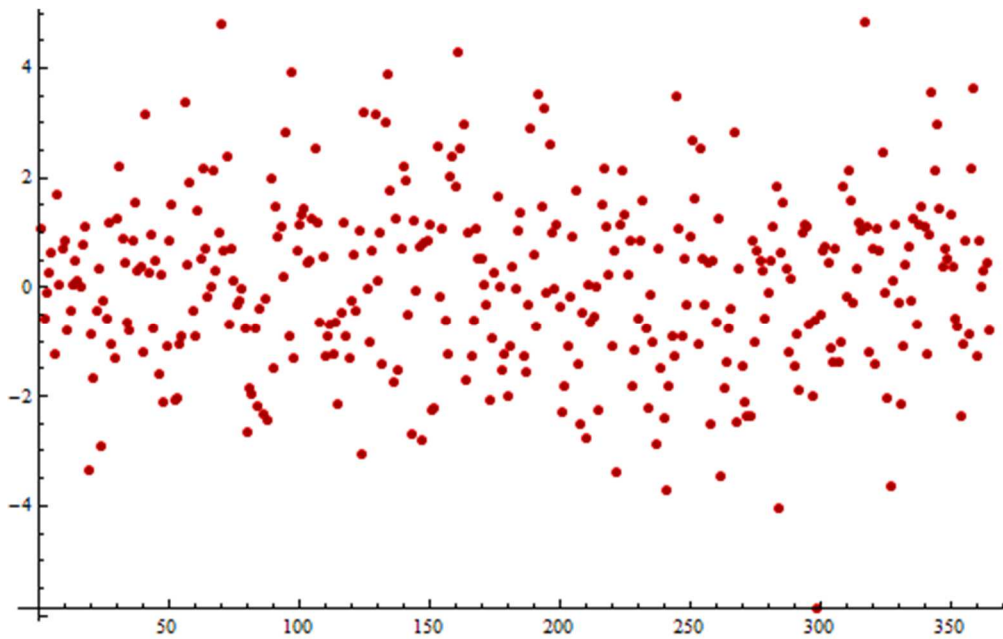
Lösung

Produktionsparameter
als Überlagerung von
Frequenzen betrachten.

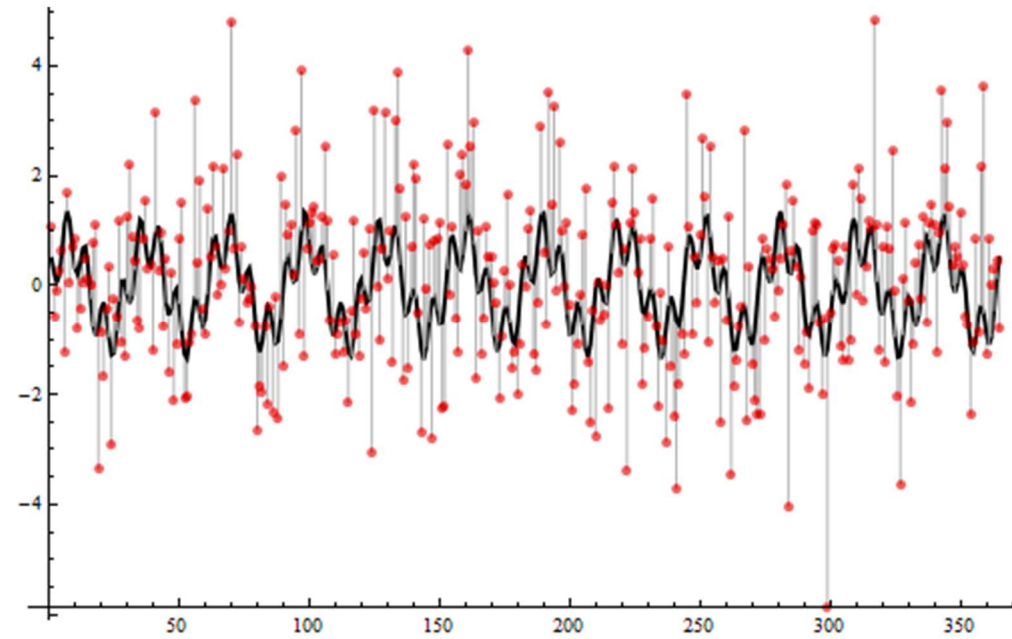
Parameter transformieren
und signifikante
Frequenzen darstellen.

Erkannte Zyklen
untersuchen.

Schwankungsanalyse - Kontinuierliche Trends

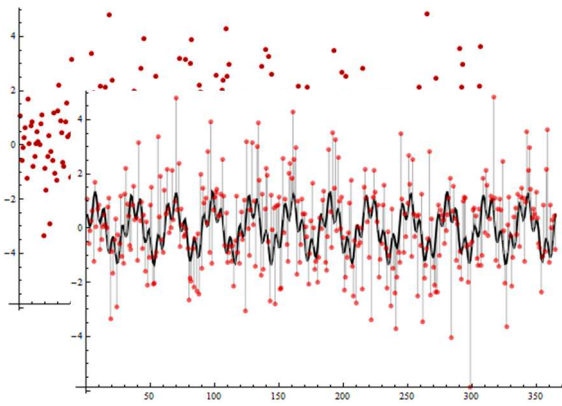


Was wir sehen.

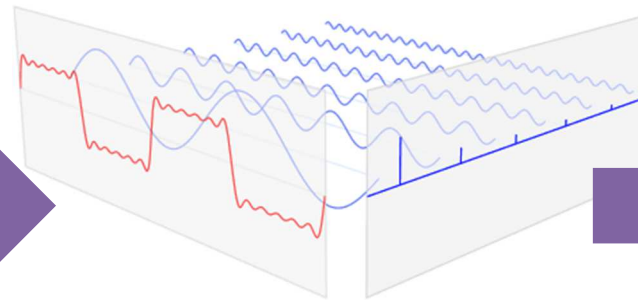
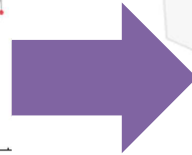


Was darin steckt.

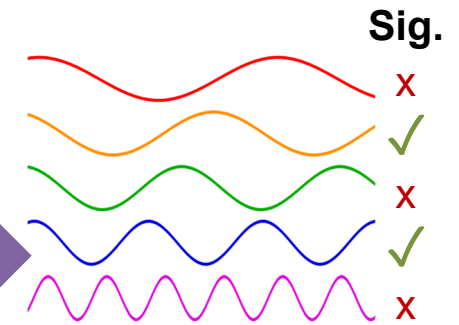
Schwankungsanalyse - Der Prozess



**Variable zu Signal
transformieren**

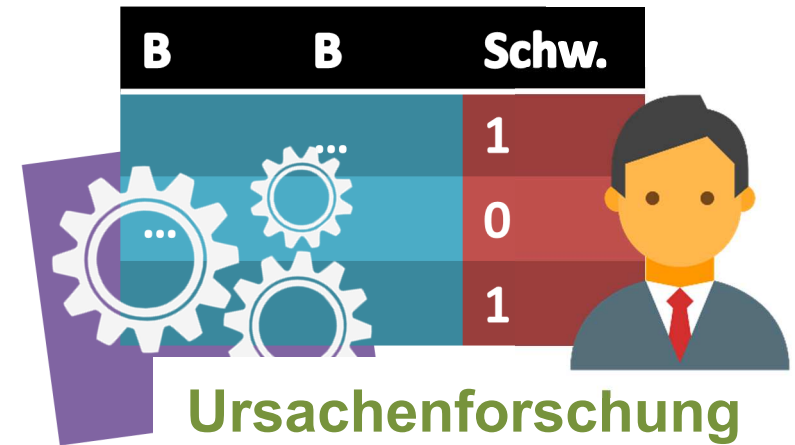
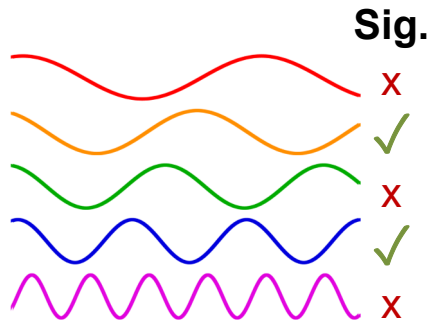


**In Unterschwingungen
zerlegen**



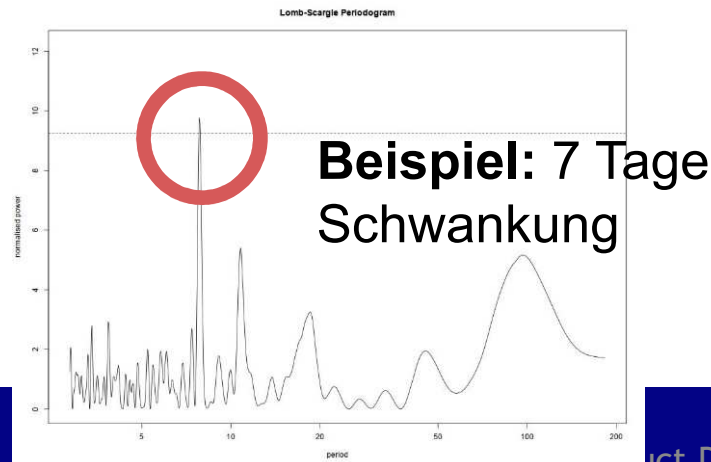
**Sind Frequenzen
signifikant?**

Schwankungsanalyse - Interpretieren



**Ursachenforschung
mit signifikanten
Schwankungen**

1. Korrelationen
2. Vergleichen mit bekannten Fällen im gleichen Zeitraum
3. ...



Maßgeschneiderte Lösungen für KMUs? Ja!

- **KMU relevante Lösungen** auch in Zeiten von Big Data möglich
- Einstiegshürden können in universitärer **Kooperation** - auch mit Absolventen - genommen werden



Zukunftsvision Yield-Prediction

- Bisherige Erkenntnisse und Methoden bündeln und **Fehler und Entwicklungen vorhersagen**



**Thank you for your
attention**