A red line graphic that starts as a vertical line on the left, then turns right to form a horizontal line. It has a small rectangular notch cut out of the horizontal line, positioned under the first 'E' of the 'FREE' text.

**FREE**

**Frühzeitige Erkennung und  
Entscheidungsunterstützung für kritische  
Situationen im Produktionsumfeld**

**Big Data zur Unterstützung von Anlagenfahrern in der  
chemischen Industrie**



**Deutsches Zentrum  
für Luft- und Raumfahrt e.V.  
Projekträger im DLR**

GEFÖRDERT VOM



**Bundesministerium  
für Bildung  
und Forschung**

# Chemische Industrie – Ein Fall für Big Data



## ■ High Volume, e.g.:

- > 300 GB metrische Daten p.a. in einer einzelnen Raffinerie
- 400 GB Alarmdaten p.a. in einer petrochemischen Anlagen

## ■ High Velocity, e.g.:

- 66.000 Sensoren mit Abtastraten von 1s – 60s in einer Raffinerie

## ■ High Variety, e.g.:

- Zeitreihen, Log-Daten, unstrukturierter Text, Videodaten

## ■ Low Veracity, e.g.:

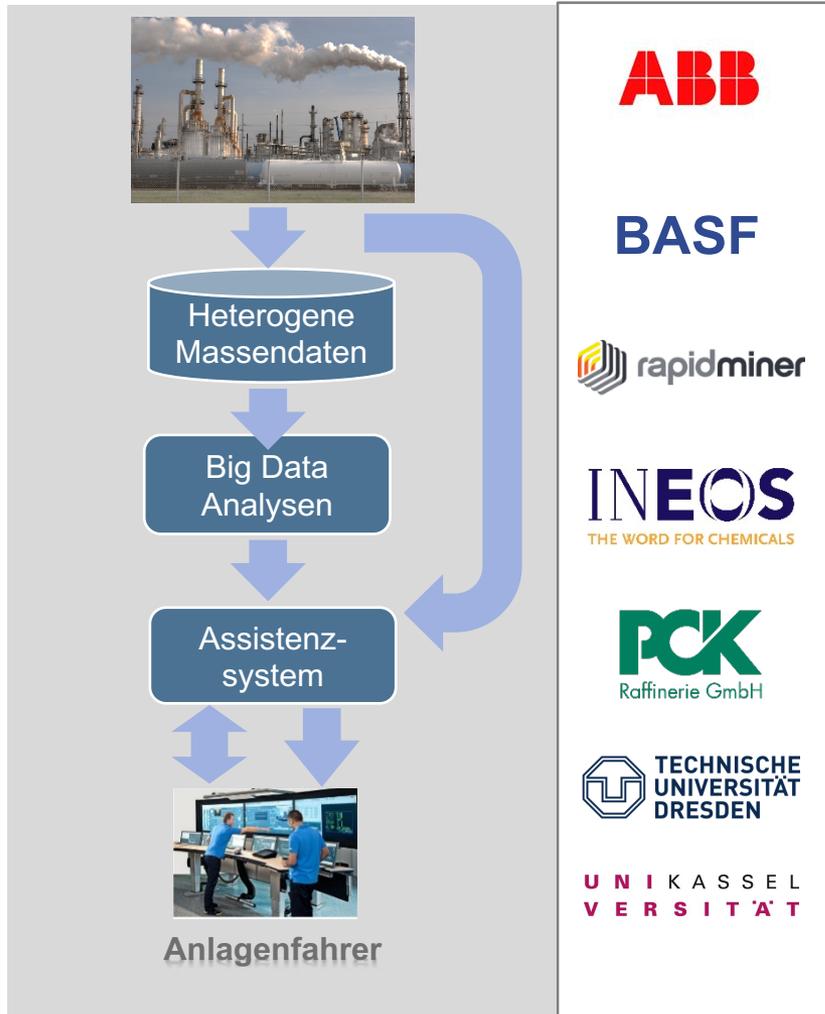
- Zeitsynchronisation, fehlerhafte Messwerte, verlorene Daten

# Chemische Industrie – Herausforderungen für Big Data Projekte



- Anspruchsvolle Fragestellungen für die Datenanalyse – Mehr maschinelles Lernen als Statistik
- Datensammlung nicht optimiert für Big Data Analysen
- Hohe Aufwände zur Datenexploration durch Datensilos mit unstrukturierten Referenzen
- Hohe Aufwände in der Datenaufbereitung durch für den Analysten unklare Zusammenhänge

# FEE im Überblick



## Ziel: Assistenzfunktionen für Anlagenfahrer

- Frühzeitige Warnungen
- Ad-hoc Analyse
- Entscheidungsunterstützung

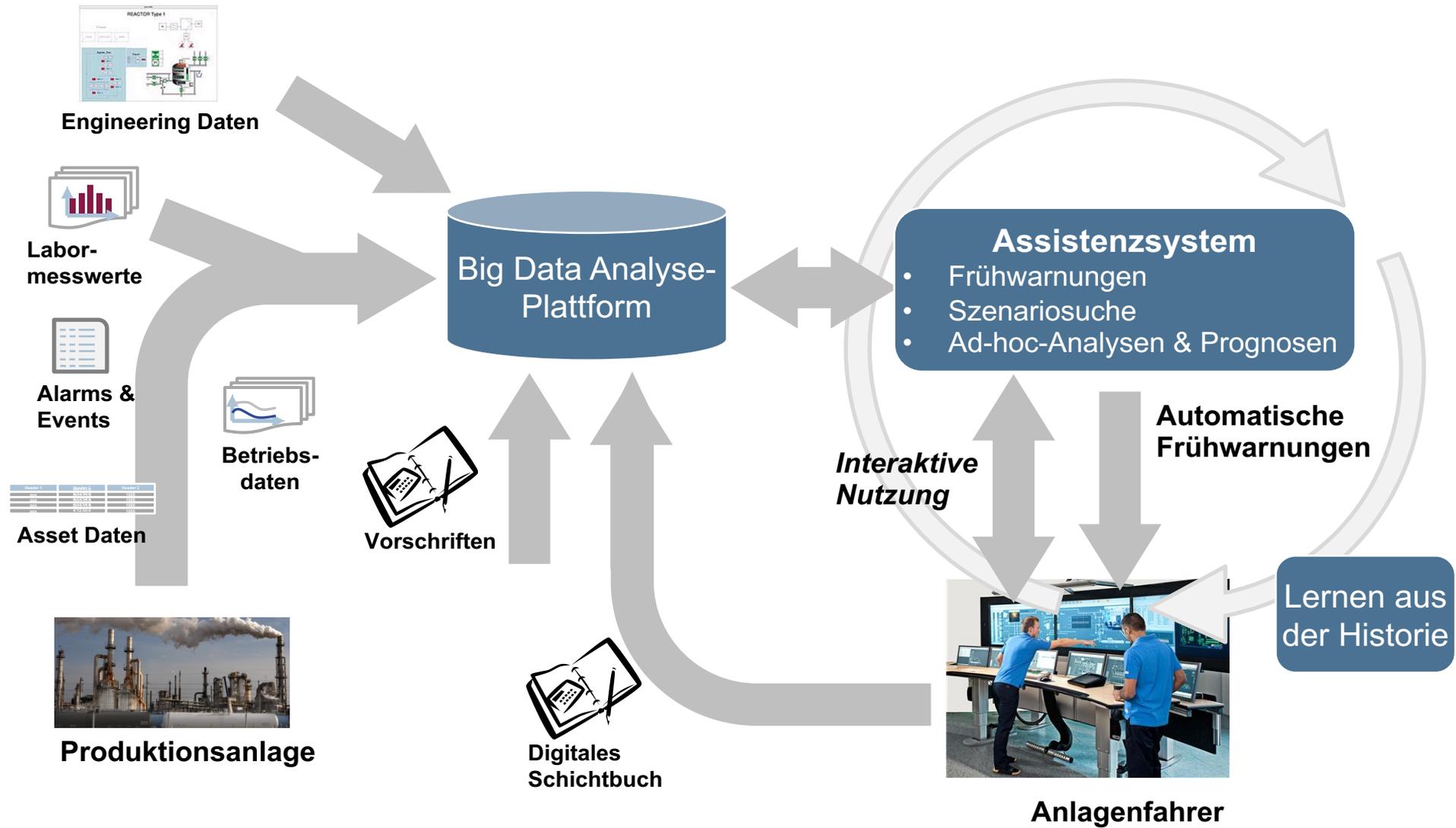
## Ansatz: Integrierte Analyse von Anlagendaten

- Betriebsmessungen, Engineering Daten, Elektronische Schichtbücher ...

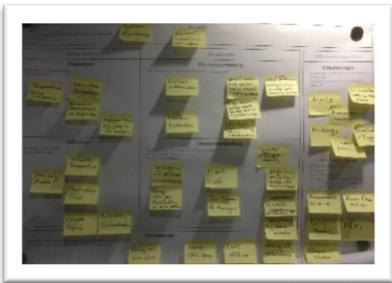
## Forschungsthemen

- Algorithmenentwicklung
- Indexierung von und Suche in Prozessdaten
- Einbindung in Echtzeit-Anlagenbetrieb
- Big Data Technologien und Architektur
- Nutzerzentrierte Interaktionskonzepte

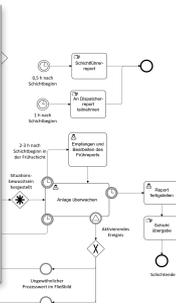
# FEE - Daten- und Systemlandschaft



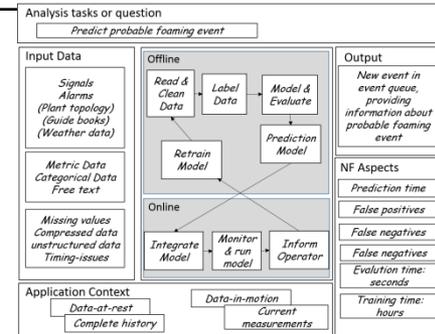
## FEE – Vorgehen im Projekt



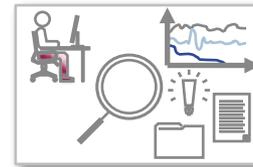
1. Szenarioidentifikation



2. Papier Prototypen



3. Analyse Workflows & Nicht-Funktionale Anforderungen

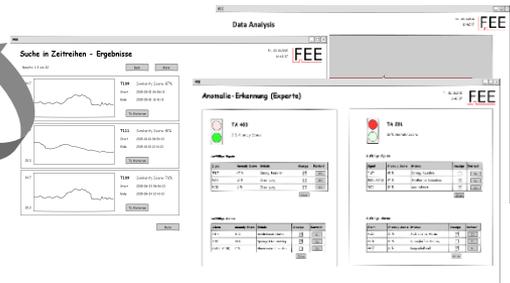


4. Proof-of-Concept Data Analytics

Quantifizierbare Aussagen		
SUS-Score	Befragung mit Hilfe des System-Usability-Scale-Fragebogens (SUS)	Summe aus 10 Fragebögen (0-100)
Einzelne Items		
	Ich denke, dass ich dieses System gerne regelmäßig nutzen würde.	[1,5]
	Ich fand das System unnötig komplex.	[1,5]
	Ich denke, das System war leicht zu benutzen.	[1,5]
	Ich denke, ich würde die Unterstützung einer fachkundigen Person benötigen, um das System benutzen zu können.	[1,5]
	Ich fand, die verschiedenen Funktionen des Systems waren gut integriert.	[1,5]
	Ich hätte das System für zu inkomplex.	[1,5]
	Ich glaube, dass die meisten Menschen sehr schnell lernen würden, mit dem System umzugehen.	[1,5]
	Ich fand das System sehr unverständlich zu benutzen.	[1,5]
	Ich fühle mich bei der Nutzung des Systems sehr sicher.	[1,5]
	Ich musste viele Dinge lernen, bevor ich mit dem System arbeiten konnte.	[1,5]



6. Integration und Deployment



5. Verfeinerte Mock-ups

# FREE

Big Data zur Unterstützung von  
Anlagenfahrern

Szenario – From Big Data to Smart Data



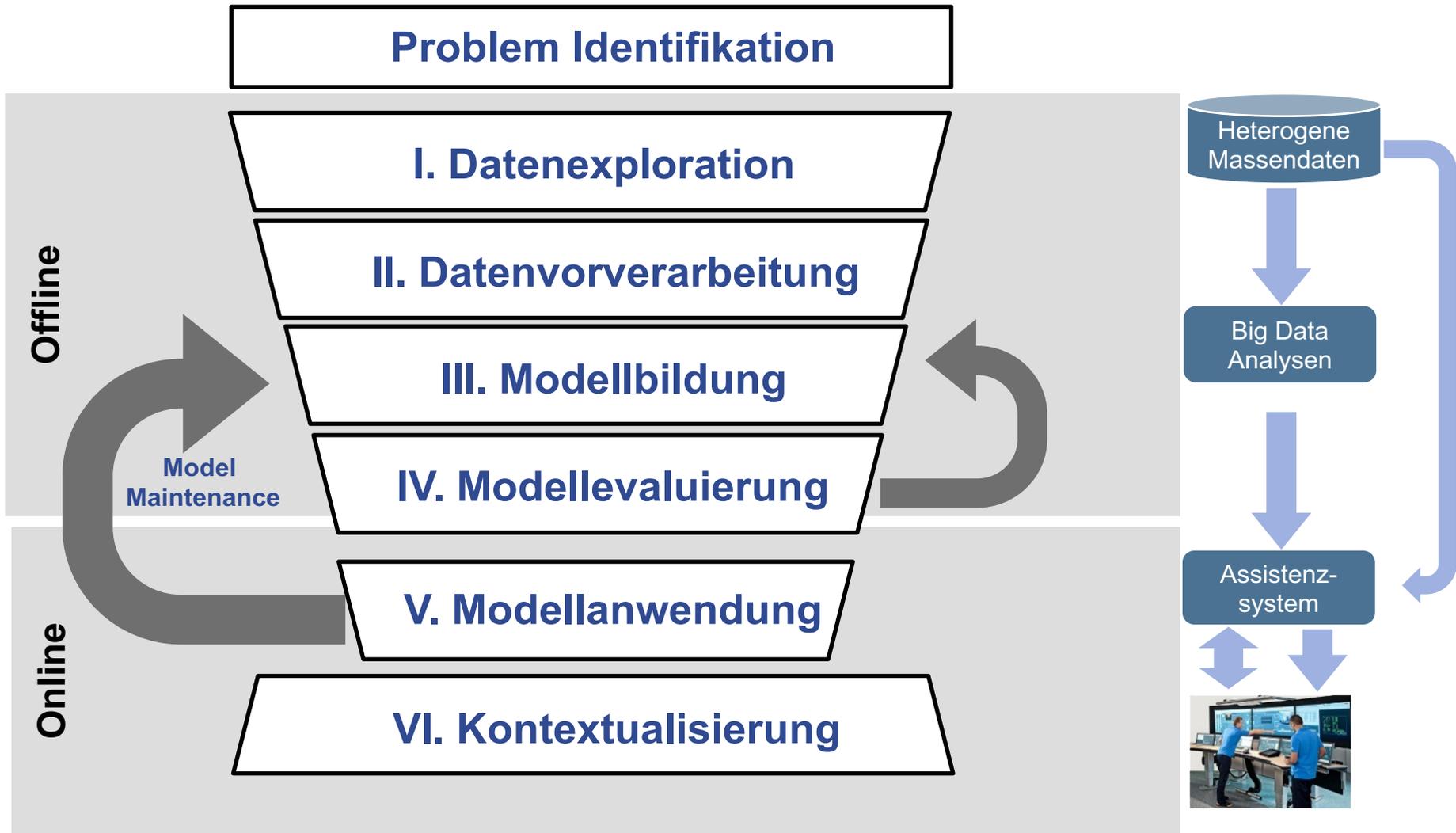
Deutsches Zentrum  
für Luft- und Raumfahrt e.V.  
Projekträger im DLR

GEFÖRDERT VOM



Bundesministerium  
für Bildung  
und Forschung

# Lebensphasen einer FEE Funktion



## Ist-Zustand:

### Wer:

- Anlagenfahrer in der Leitwarte und im Feld

### Was:

- Schaumbildung in einer Prozesskolonne führt zu erhöhten Druck und Risiko des Überlaufs. Anti-Schaummittel muss rechtzeitig von Hand zugegeben werden.

### Wie:

- Überwachung der relevanten Signale durch Anlagenfahrer in der Leitwarte

### Probleme:

- (1) Risiko die Schaumbildung nicht rechtzeitig erkannt wird
- (2) Schaumbildung zeigt sich sehr kurzfristig in Signalen – Eingriff erfolgt immer unter Zeitdruck
- (3) Unerfahrene Anwender erkennen evt. die Situation nicht oder wissen nicht, wie einzugreifen ist

## Soll-Zustand:

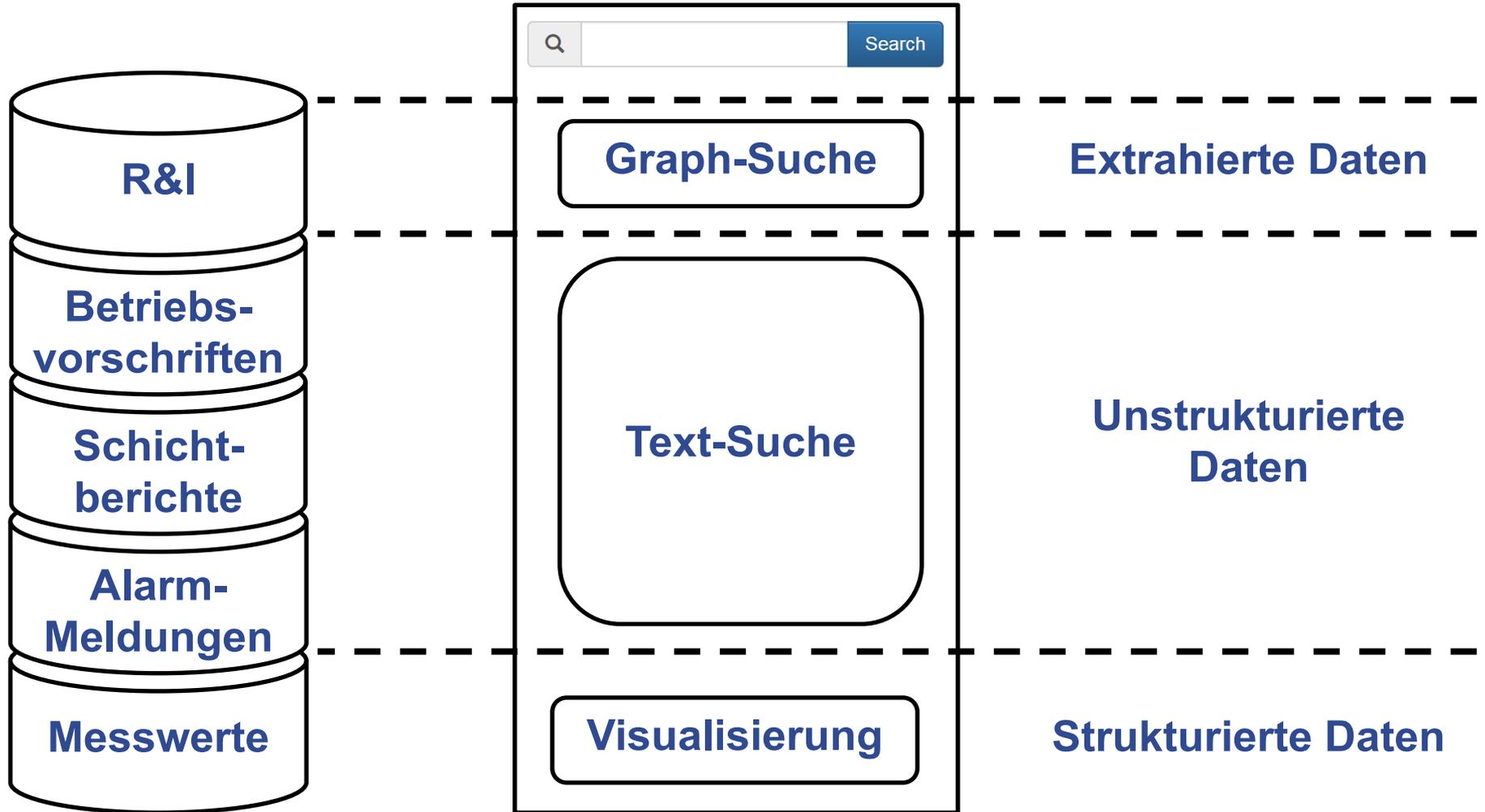
### FEE Support:

Frühzeitige Information über eine sicher oder wahrscheinliche Schaumbildung in naher Zukunft.

### Wunsch:

- (1) Rechtzeitige Information – spätestens 30 Minuten vor dem Eintreten des Ereignis
- (2) Eindeutige Handlungsanweisung, Diagnostizität sollte nicht notwendig sein
- (3) Hohe Trefferquote, wenige Fehllarme

# Hybride Datenexploration



# Textsuche / Kontextualisierung

🕒 01.04.2014 - 30.04.2014

☰

Anlagenkontext

Textsuche

Alle

Alarme

Betriebsvorschriften

Schichtberichte

Messstellenliste

**Ereignis**  
K433 Übrerriss ,Antischaum dosiert,A433 a.B.

**Produktionsaufgabe**  
SC4 - Scotanlage CA4

**NÜ**  
X

**Zeitpunkt**  
2014-04-22T13:05:00

**M**  
X

**EreignisKlasse**  
INFORMATION

**A**  
X

**Bereich**  
AV-2

**Ereignisbeginn**  
22.04. 13:05

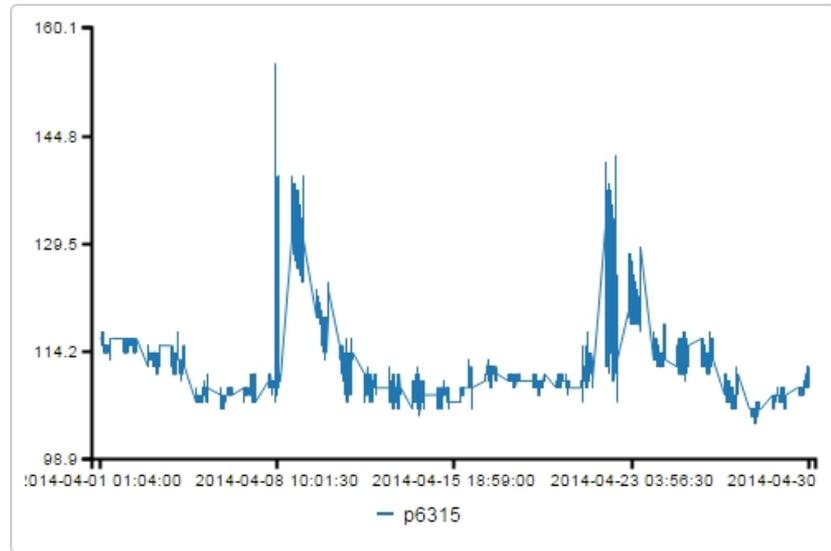
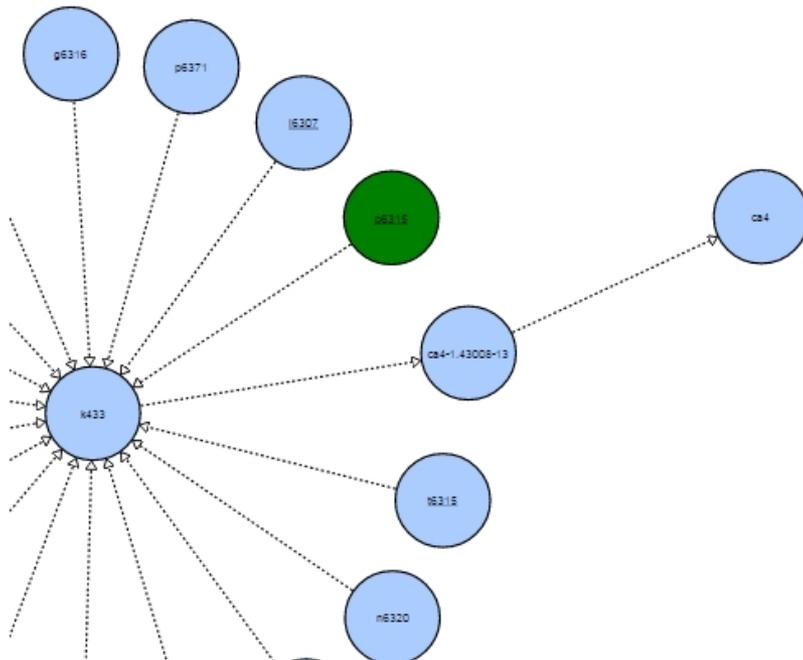
```

t-5: NaOH zur K431 dosiert, Eingangsarmatur NaOH B.L. wieder geschlossen
t-4: A431 a.B., gespült und entleert, spült mit N2,SC4 - Scotanlage CA4,X
t-3: A431 Filterkerzen wechseln,SC4 - Scotanlage CA4,X,2014-04-21T20:19:00
t-2: A431 Filterkerzen gewechselt und i.B.,SC4 - Scotanlage CA4,X,2014-04-22T12:55:00,X
t-1: LZ6309/A überprüft i.O.,SC4 - Scotanlage CA4,X,2014-04-22T12:55:00,X
t: K433 Übrerriss ,Antischaum dosiert,A433 a.B.,SC4 - Scotanlage CA4,X,2014-04-22T13:05:00,X
t+1: P433A Saugsiebkontrolle,SC4 - Scotanlage CA4,X,2014-04-22T15:56:00,1
t+2: P433B Saugsiebkontrolle,SC4 - Scotanlage CA4,X,2014-04-22T16:01:00,1
t+3: P433B Scheibe in Entleerung Saugseite stecken,SC4 - Scotanlage CA4,X
t+4: A433 wieder i.B.,SC4 - Scotanlage CA4,X,2014-04-22T16:23:00,X,INFORM
          
```

Einfacher Zugang durch Volltextsuche

# Textsuche / Kontextualisierung (2)

📅 01.04.2014 - 30.04.2014



**Grafische Exploration auf Basis automatisch abgeleiteter Topologien**

# Tool-gestützte Datenvorverarbeitung

The screenshot displays the 'Data Quality' tool interface. On the left, a sidebar lists various data processing options: Raw Data, Formatting, Messages, Missing Values Options, Interpolation Options, Technically Correct Data, Consistent Data, Tidy Data: Review, Tidy Data: Restructure, and About. The main area is divided into three panels: 'Example: Original Sensor' showing a line graph with significant gaps, 'Example: Linearly interpolated Sensor' showing the same data with straight lines connecting the points, and 'Example: Cubic spline interpolated Sensor' showing the data with a smooth curve. Below these, a 'Redundant Sensors' table lists sensor pairs with their correlation coefficients. A 'Max gap size (minutes) to interpolate' slider is also visible.

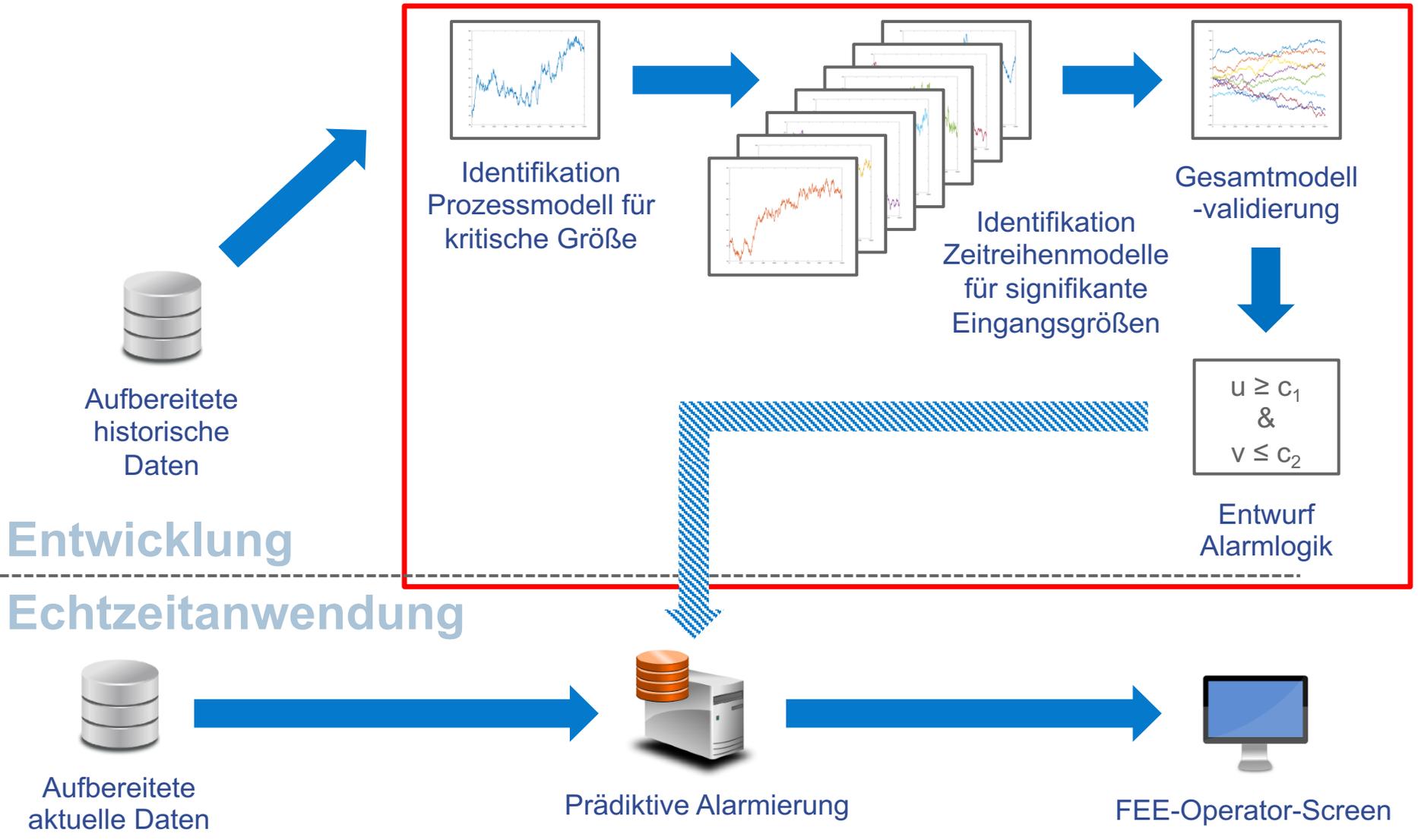
**Identifizieren und entfernen redundanter Signale**

**Behandlung von Lücken in Sensordaten**

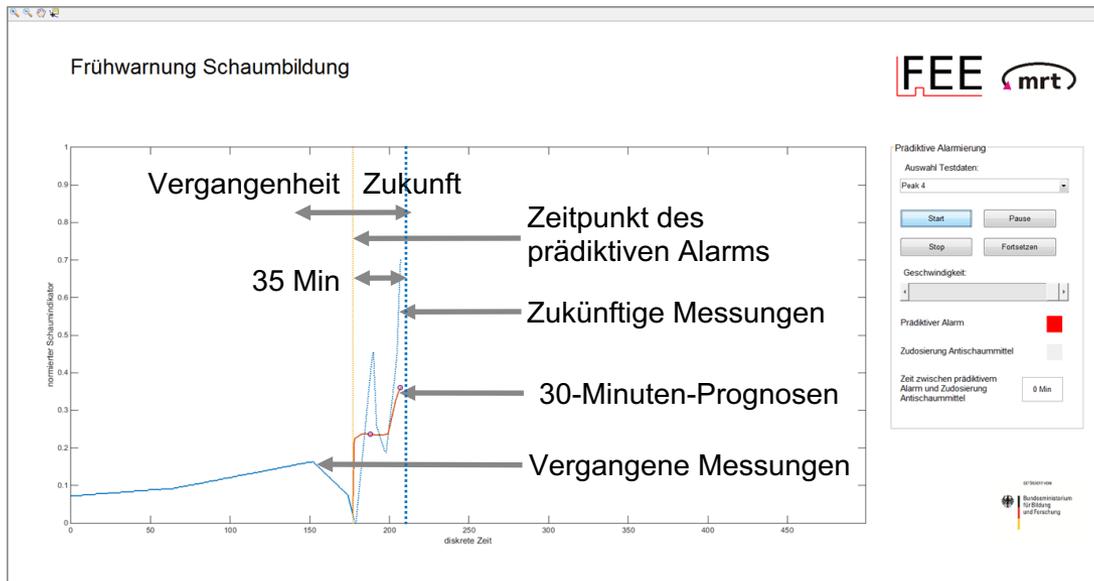
row	column	cor	p	p-adjusted
1	SENSOR_ID_1	SENSOR_ID_2	1	0
2	SENSOR_ID_1	SENSOR_ID_3	1	0
3	SENSOR_ID_1	SENSOR_ID_4	1	0
4	SENSOR_ID_1	SENSOR_ID_5	1	0
5	SENSOR_ID_1	SENSOR_ID_6	1	0
6	SENSOR_ID_1	SENSOR_ID_7	1	0
1081	SENSOR_ID_1	SENSOR_ID_8	1	0
1176	SENSOR_ID_1	SENSOR_ID_9	1	0
1222	SENSOR_ID_1	SENSOR_ID_10	1	0
1223	SENSOR_ID_1	SENSOR_ID_11	1	0

**Beschleunigung typischer Tätigkeiten des Data Cleansing!**

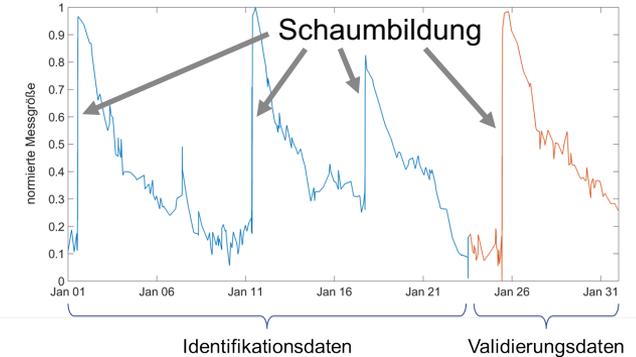
# Modellbildung, -validierung und -anwendung



- Automatisierte Selektion signifikanter Eingangssignale und Modellterme für ARX-Prozessmodell
- Automatisierte Selektion signifikanter Modellterme für AR-Zeitreihenmodelle
- Gesamtvalidierung mittels iterierter Mehrschrittprädiktion
- Einfache Alarmlogik auf Prädiktion (Schwellenwerte für Signalamplitude und Signalsteigung)
- Prädiktive Alarmierung aus Engineering-Sicht:



- Kritische Prozessgröße:



- Zeitliche Auflösung: 1 Min
- Anzahl Messwerte pro Signal: 44641
- Anzahl Eingangssignalkandidaten für Prozessmodell: 29
- Anzahl signifikanter Eingangssignale: 7
- Vorzeitigkeit des prädiktiven Alarms: 35 Minuten
- Keine Falschalarme

# FREE

Big Data zur Unterstützung von  
Anlagenfahrern

Szenario – Anomalie-Erkennung: Big Data für seltene  
Ereignisse



Deutsches Zentrum  
für Luft- und Raumfahrt e.V.  
Projekträger im DLR

GEFÖRDERT VOM



Bundesministerium  
für Bildung  
und Forschung

# Szenario: Analyse auffälliger Signale (Anomalie-Erkennung)

## Ist-Zustand:

### Wer:

- Anlagenfahrer in der Leitwarte (und Prozess-Ingenieure)

### Was:

- Überwachung des Prozesses in ruhigen Situationen

### Wie:

- Durchsuchen der Operator Screens und Trenddisplays auffälligen Signalverläufen
- Wird eher in ‚ruhigen‘ Prozesssituation durchgeführt

### Probleme:

- (1) Risiko einen signifikanten Signalverlauf zu übersehen
- (2) Kein breites Monitoring in stressigen Situationen
- (3) Schwierigkeiten unerfahrener Anlagenfahrer, die Besonderheit von Signalverläufen zu bewerten

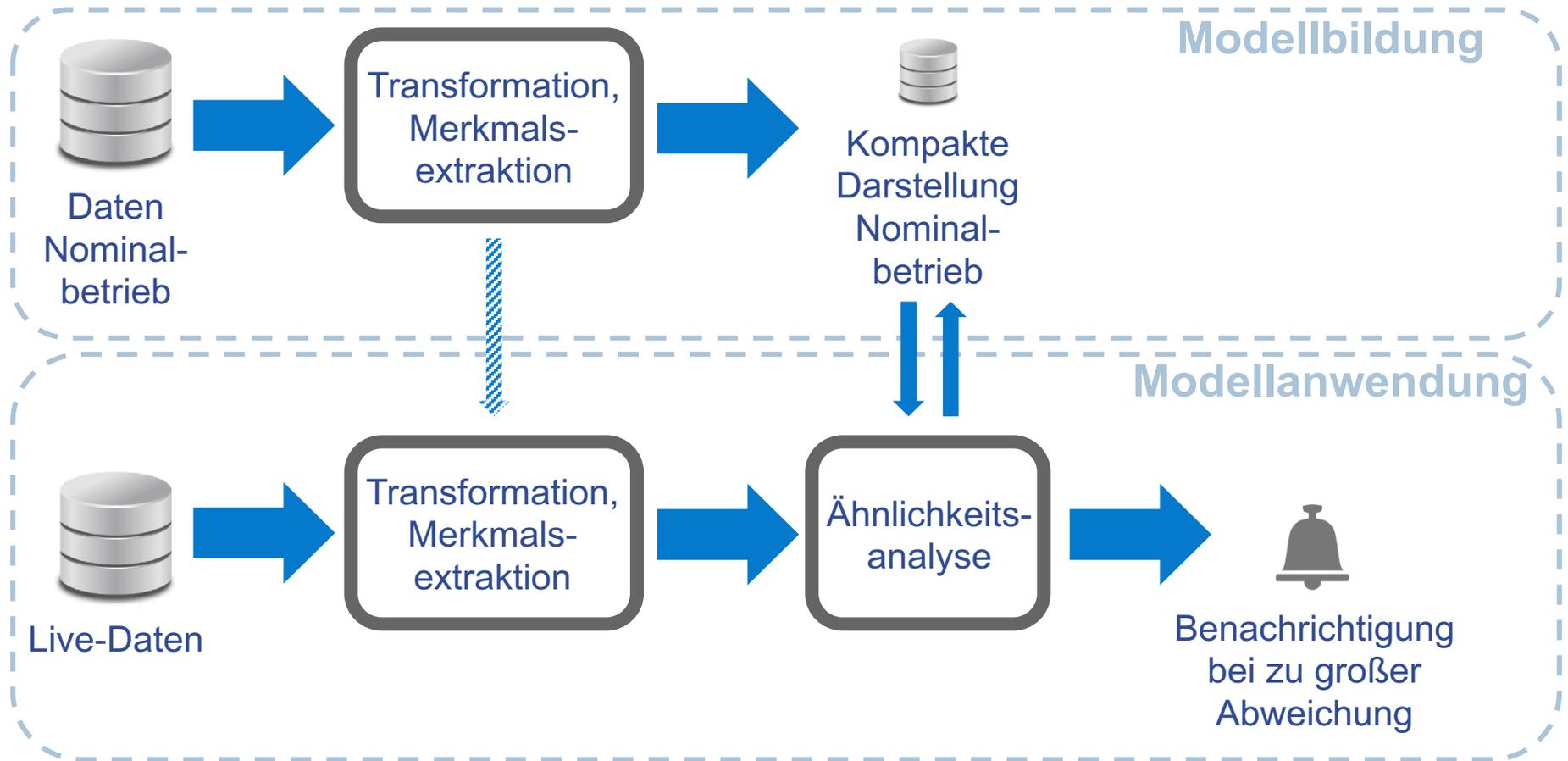
## Soll-Zustand:

### FEE Support:

Identifikation auffälliger Signalverläufe und Bereitstellung relevanter Informationen zur Diagnose

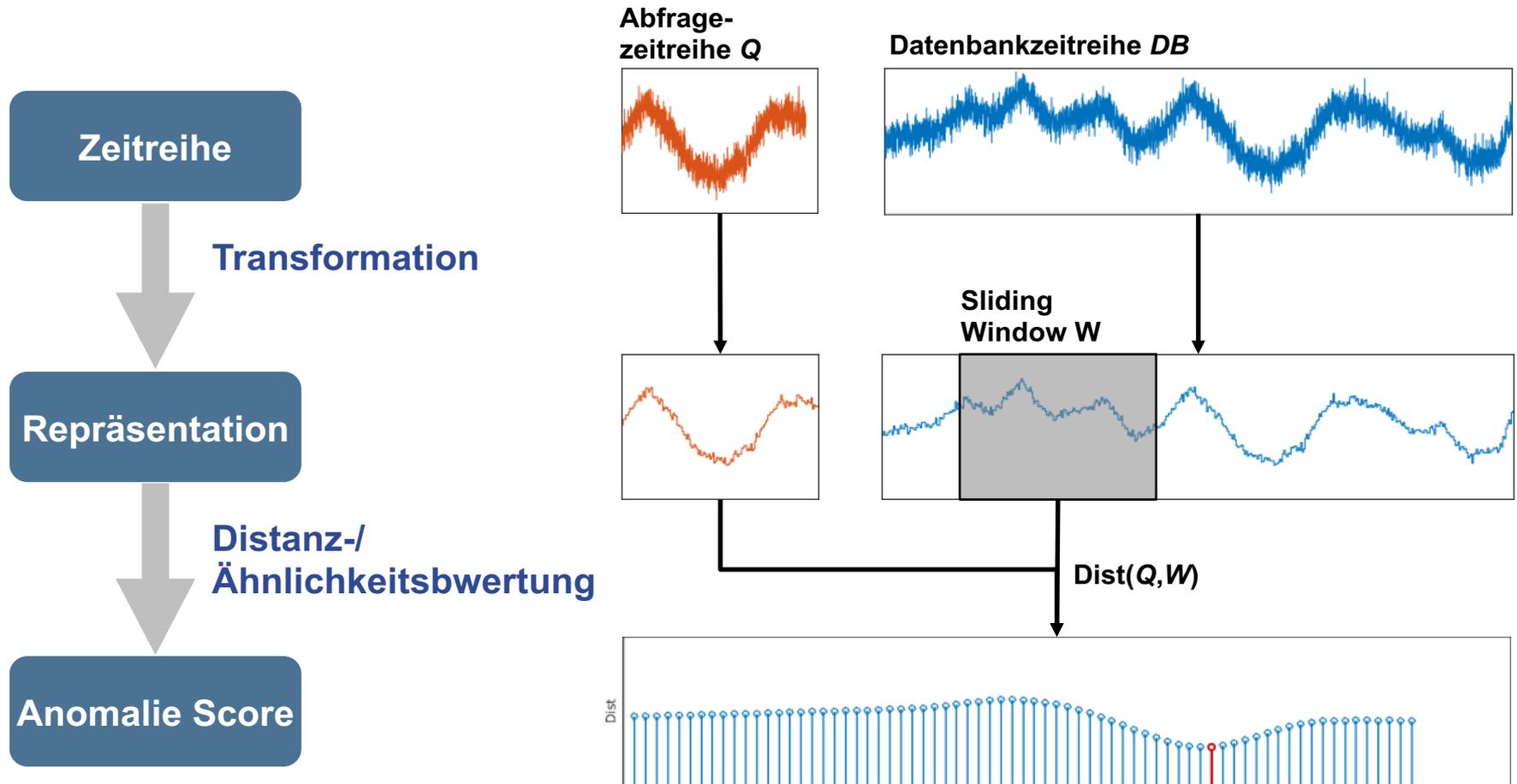
### Wunsch:

- (1) Schneller visueller Eindruck über Auffälligkeiten im Prozess
- (2) Verbindung zu historischen ‚normalen‘ und ‚abnormalen‘ Signalverläufen darstellen
- (3) Herstellen eines erweiterten Kontexts (relevante Alarme, Operator Notes, Dokumente, etc.)



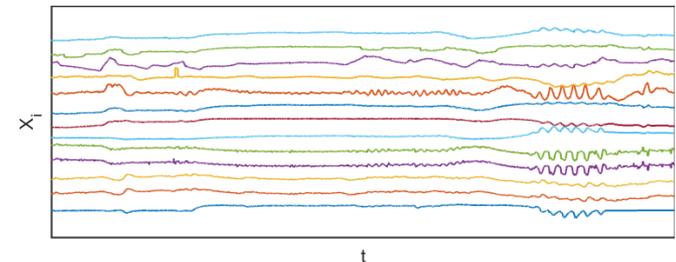
# Subsequence Matching basierte Anomaliedetektion

Die Distanz zwischen einer Abfragezeitreihe und der zu dieser ähnlichsten Subsequenz in einer Datenbankzeitreihe wird als Anomalie-Score verwendet.

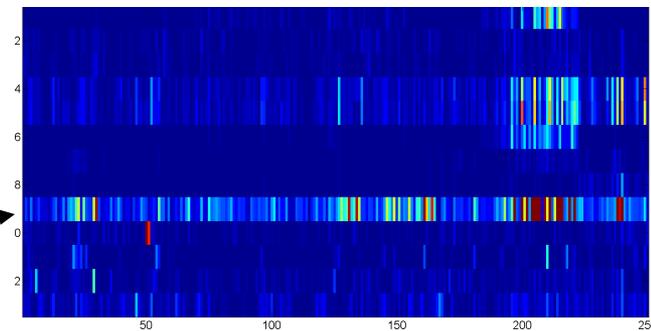
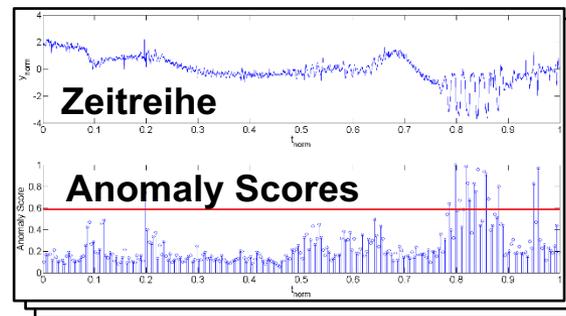


# Fallstudie: Oszillationsdetektion

- Kontinuierlich betriebene Butadienanlage
- Eine (bekannte) einmalig auftretende Anomalie
- **Großes Datenvolumen:** 1000 Messstellen mit Abtastintervall von 1 min über einen Zeitraum von 2 Jahren
- **Heterogenität:** Drücke, Durchflüsse, Füllstände, Analysedaten, Temperaturen; Kompressionsrate variiert mit der Zeit und von Zeitreihe zu Zeitreihe.
- **Instationarität:** Häufige Lastwechsel
- **Datenselektion:**
  - Datensatz ohne Expertenwissen: Durch Elimination redundanter und überwiegend konstanter Zeitreihen reduziert auf 104 Messstellen.
  - Datensatz mit Expertenwissen: 13 Messstellen (dargestellt)



- Visualisierung der berechneten Anomaly Scores der Zeitreihen in einer Heatmap

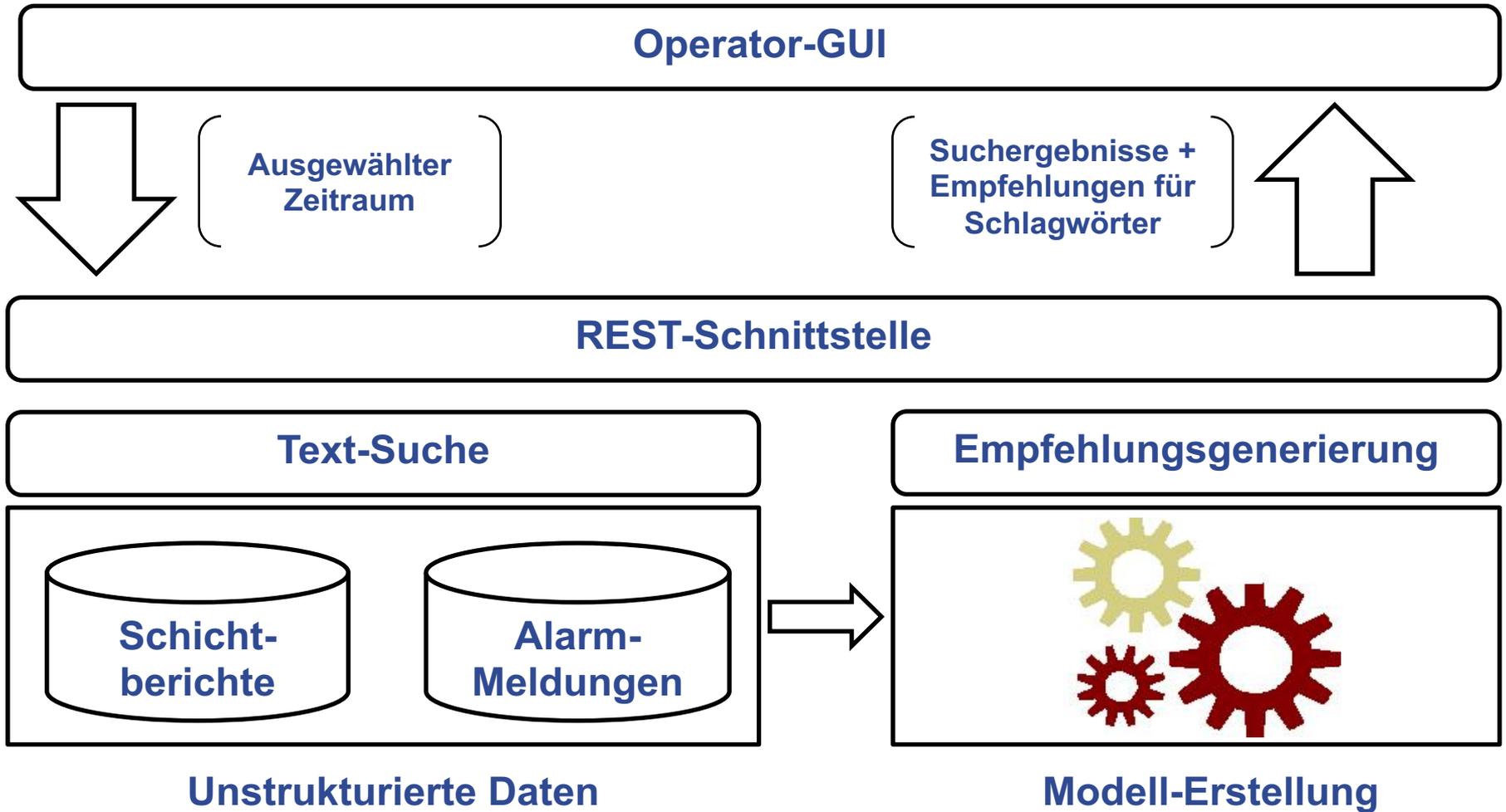


- Viele Information im Betrieb liegen in unstrukturierter Form vor

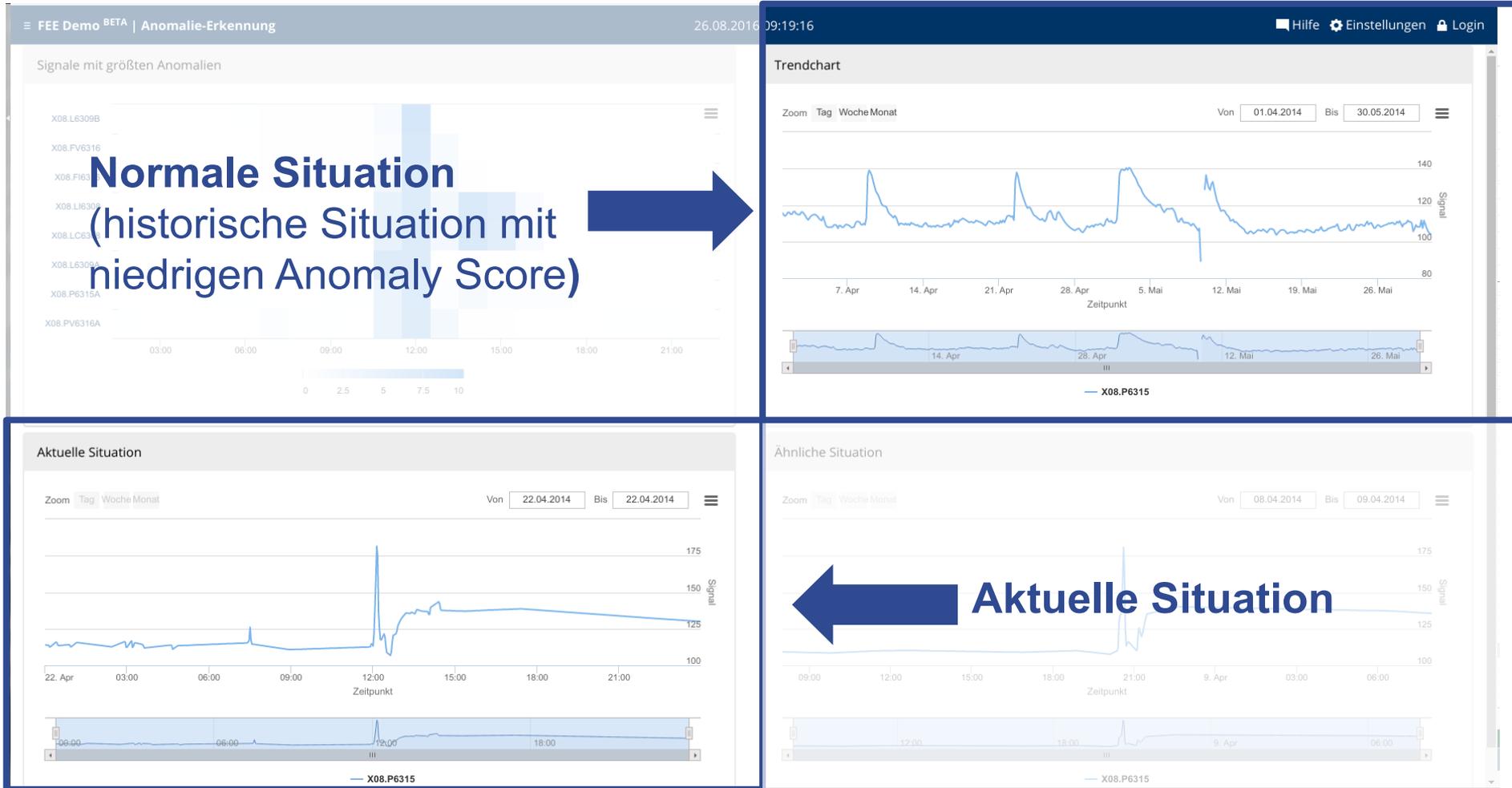


- Ziel: Unterstütze den Operator beim Finden dieser Informationen durch gezielte Empfehlung von Suchbegriffen für einen Kontext

Antischaum  
Durchfluss  
Desorber  
Kolonne  
Pumpe 324  
Kopfdruck









# Operator Schnittstelle zur Anomalie-Erkennung (2)

FEE Demo BETA | Anomalie-Details 26.08.2016 09:19:49 Hilfe Einstellungen Login

**Aktuelle Situation**

Zoom Tag Woche Monat Von 22.04.2014 Bis 22.04.2014

Signal  
150  
100  
Zeitpunkt  
22. Apr 03:00 06:00 09:00 12:00 15:00 18:00 21:00

X08.P6315

**Ähnliche Situationen: 1 Suchergebnis**

Zoom Tag Woche Monat Von 08.04.2014 Bis 09.04.2014

Signal  
150  
100  
Zeitpunkt  
09:00 12:00 15:00 18:00 21:00 9. Apr 03:00 06:00

X08.P6315

**Referenzsituation**

Zoom Tag Woche Monat Von 14.04.2014 Bis 14.04.2014

Signal  
150

Dokumente Alarms Operator Notes

rk7\_f8140  
xia lochblende<sup>rot</sup>  
ca2\_x7853<sup>v313</sup> p325 u3206  
ha3\_f2847\_2 mix<sup>logik</sup> qv3105  
teilstromfilter<sup>f8140</sup>  
ea1\_f1035

**SUCHERGEBNISSE**

qv3105	>
lochblende	>
p325	>
u3206	>
rot	>
ca2_x7853	>
xia	>
mix	>
logik	>
v313	>



# Operator Schnittstelle zur Anomalie-Erkennung (2)

FEE Demo BETA | Anomalie-Details 26.08.2016 09:19:49 Hilfe Einstellungen Login

Aktuelle Situation

Zoom Tag Woche Monat Von 22.04.2014 Bis 22.04.2014

Signal 150  
100  
Zeitpunkt 22. Apr 03:00 06:00 09:00 12:00 15:00 18:00 21:00

X08.P6315

Ähnliche Situationen: 1 Suchbegriffe

**Relevante Suchbegriffe** →

Zoom Tag Woche Monat Von 08.04.2014 Bis 09.04.2014

Signal 150  
100  
Zeitpunkt 09:00 12:00 15:00 18:00 21:00 9. Apr 03:00 06:00

X08.P6315

Referenzsituation

Zoom Tag Woche Monat Von 14.04.2014 Bis 14.04.2014

Signal 150

Dokumente Alarms Operator Notes

rk7\_f8140  
xia lochblende<sup>rot</sup>  
v313  
ca2\_x7853 p325 u3206  
ha3\_f2847\_2 mix logik  
teilstromfilter qv3105  
ea1\_f1035

SUCHERGEBNISSE

qv3105	➤
lochblende	➤
p325	➤
u3206	➤
rot	➤
ca2_x7853	➤
xia	➤
mix	➤
logik	➤
v313	➤



## Zusammenfassung und Ausblick

---

### ■ Was wurde gezeigt

- Übertragung von Datenanalysemethoden auf den Kontext der chemischen Industrie
- Herausforderungen einer Big-Data Architektur für Produktionsanlagen
- Lösungsansätze für zwei typische Szenarien (Ereignisvorhersage und Anomalie-Erkennung)

### ■ Nächste Ziele

- Bearbeitung von weiteren Referenzszenarien
- Weiterentwicklung der Algorithmen und Übertragbarkeit der Methoden auf andere Anlagen zeigen
- Demonstration der Funktionen im direkten Anlagenumfeld