

2018 SMART Maintenance Conference

04. September 2018

# *Prognosen für Gleislage und Rollkontaktermüdungen – Eine technische Sicht*

 SBB CFF FFS

 ELCA

# Vorstellung

---

***Katharina Mellert, Leiterin Datenlabor Fahrweg SBB***  
*katharina.mellert@sbb.ch*

***Zacharias Kull, Manager Business Intelligence/BigData-Gruppe ELCA***  
*zacharias.kull@elca.ch*

*Co-Autor: Ingolf Nerlich, Fachexperte Fahrbahn SBB*

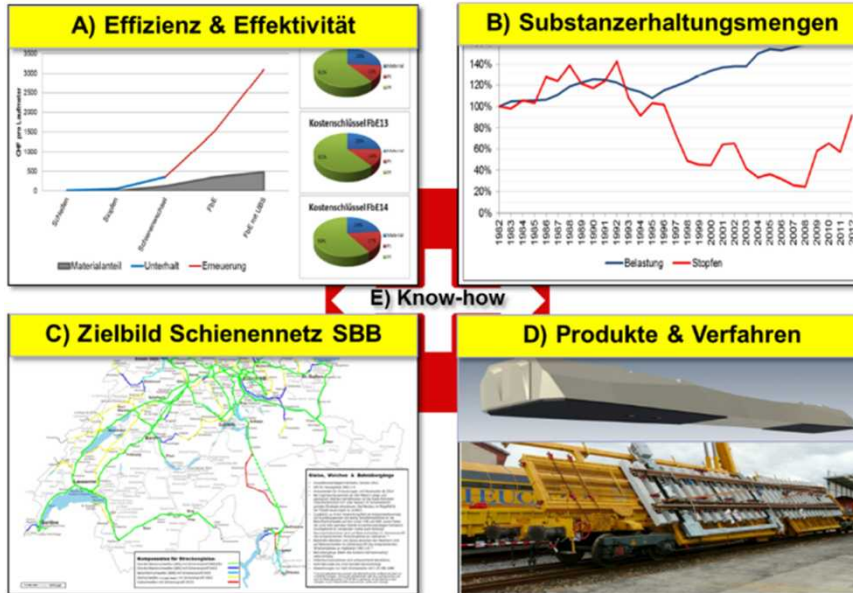
# Agenda

---

- Ziele des Anlagenmanagements und unser Weg dahin
- Stopfprognose: Erfahrungen und Architektur
- Rollkontaktermüdungsprognose: Erfahrungen und Architektur
- Blick in die Zukunft: Anforderungen an Plattformarchitektur und Frameworks



# Ziele des Anlagenmanagements und unser Weg dahin



## Programm Anlagenbewirtschaftung 4.0



- Predictive Maintenance für die Fahrbahn
- Steigende Ausgaben stabilisieren
- Fakten- und prognosebasierte Ableitung des Gesamtbedarfs an Substanzerhaltungsmassnahmen.

## Datenlabor Fahrweg



Entwickelt Prototypen zur Qualitätsverbesserung, Prognosen und Darstellung



Forschung zu Wechselwirkungen Rad-Schiene

LCC-Optimierung: Bei gegebener Verfügbarkeitsanforderung die kostengünstigste Instandhaltungsstrategie finden.



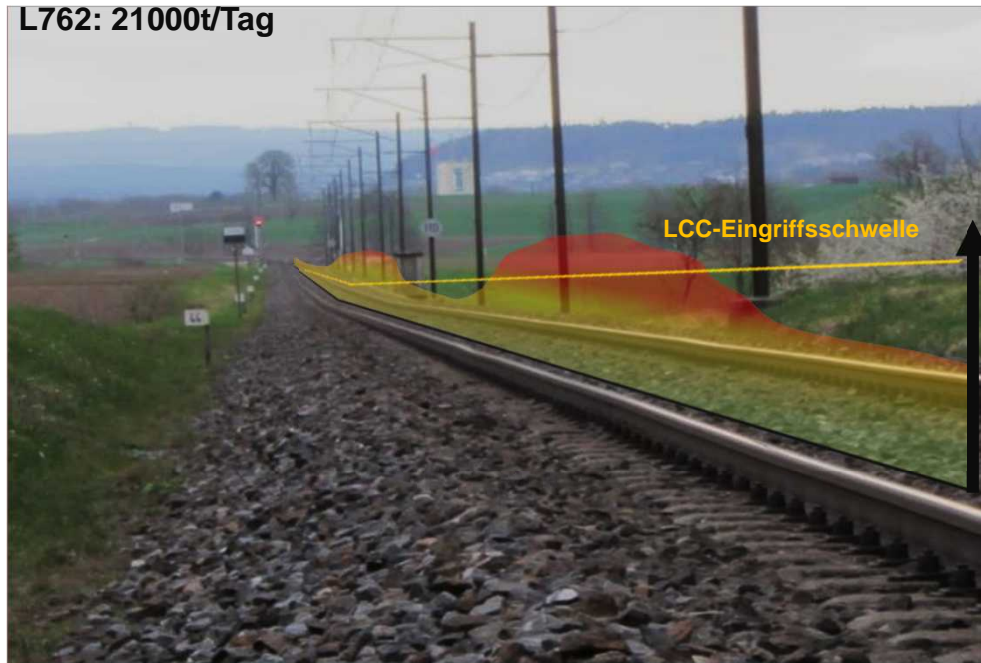
A photograph of two railway workers in orange safety gear (hard hats and high-visibility jackets) working on a trackside electrical cabinet. The worker in the foreground is using a red-handled tool to work on the internal wiring of the cabinet. The background shows railway tracks and infrastructure. The text "Stopfprognose: Erfahrungen und Architektur" is overlaid in white on the left side of the image.

# Stopfprognose: Erfahrungen und Architektur

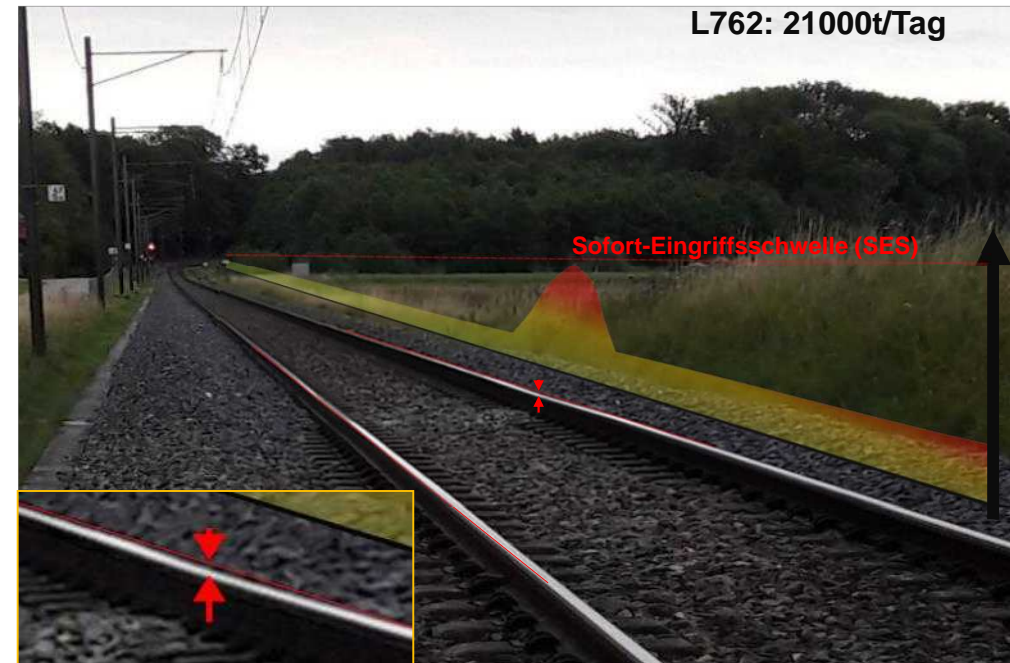


# Beispiel Stopfen

## Gleisqualität und Einzelfehler

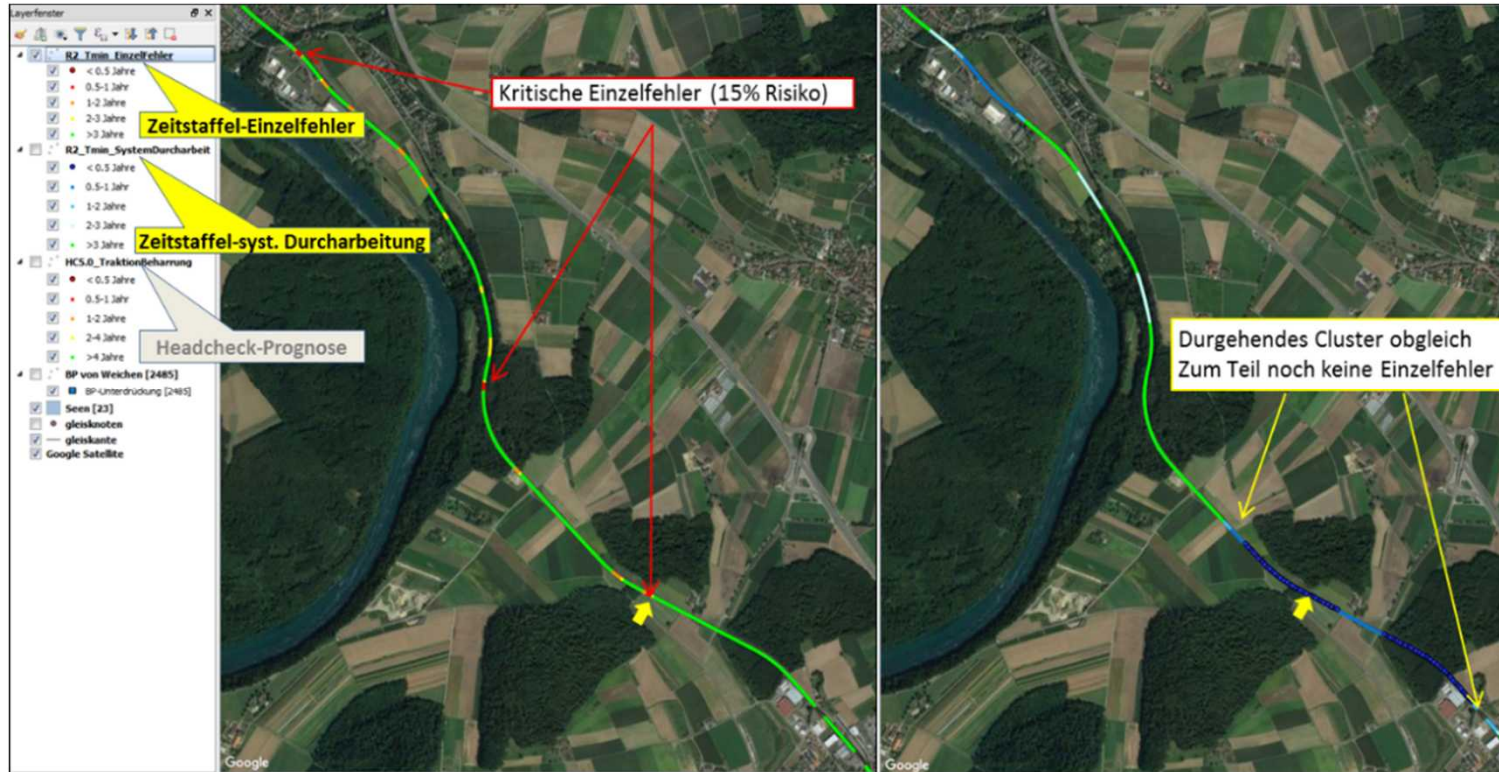


**Gleislage- Qualität (Lageunruhe)**  
- LCC aber nicht verfügbarkeitsrelevant



**Gleislage-Einzelfehler**  
- **Verfügbarkeitsrelevant**, Soforteingriffsschwelle (SES)

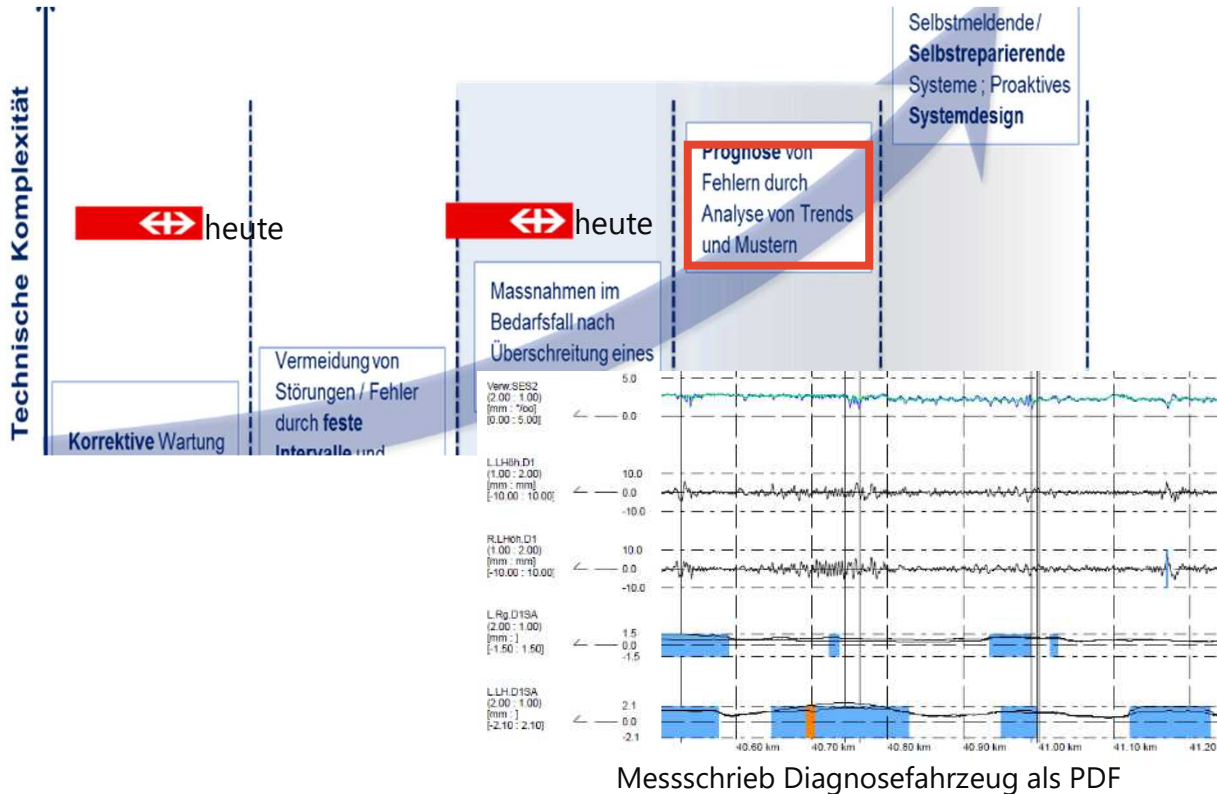
# Ziel: Prognose von Einzelfehlern und Gleislagequalität



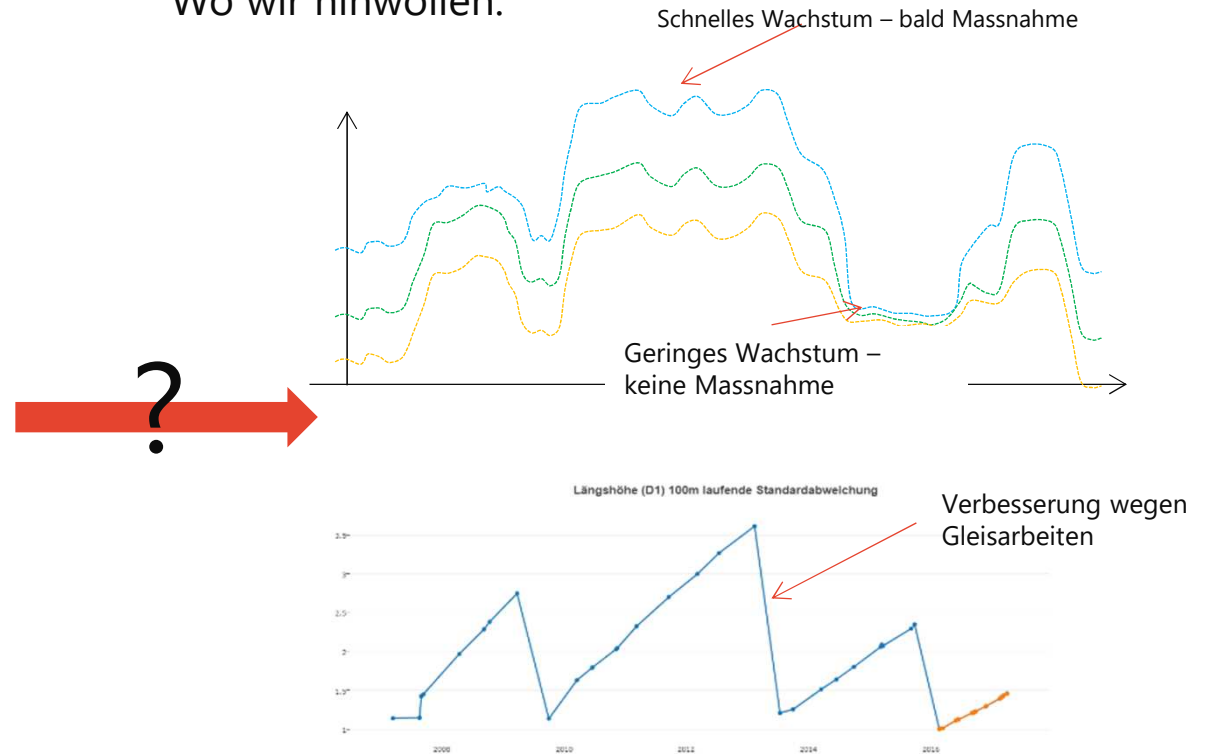
→ **Kenntnis von lokalen Fehlern hilft bei der Planung von systematischem Stopfen**

# Ziel: Einsparung durch Predictive Maintenance

Wo wir heute stehen:



Wo wir hinwollen:

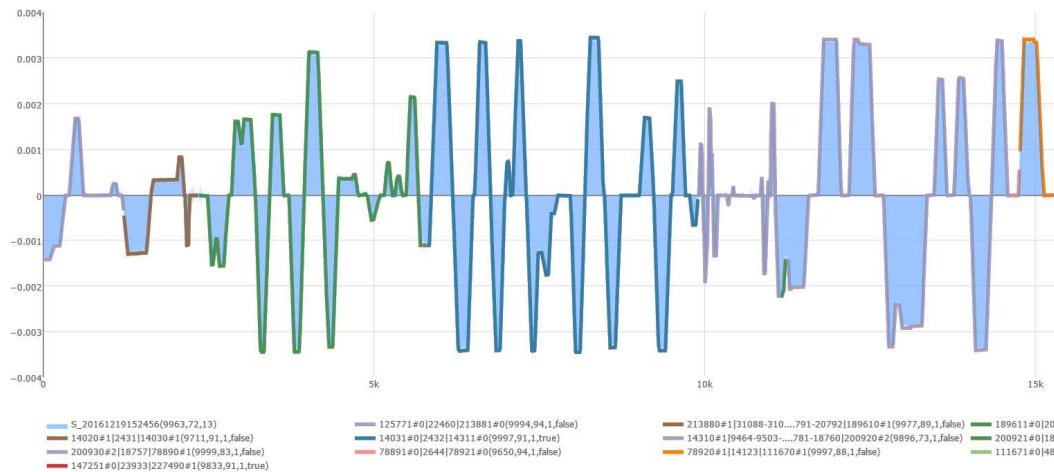


Anforderungen:  
Hohe Datenqualität auch bei historischen Daten:  
**Verortung der Messdaten** und **Gleisarbeiten**  
müssen bekannt sein.



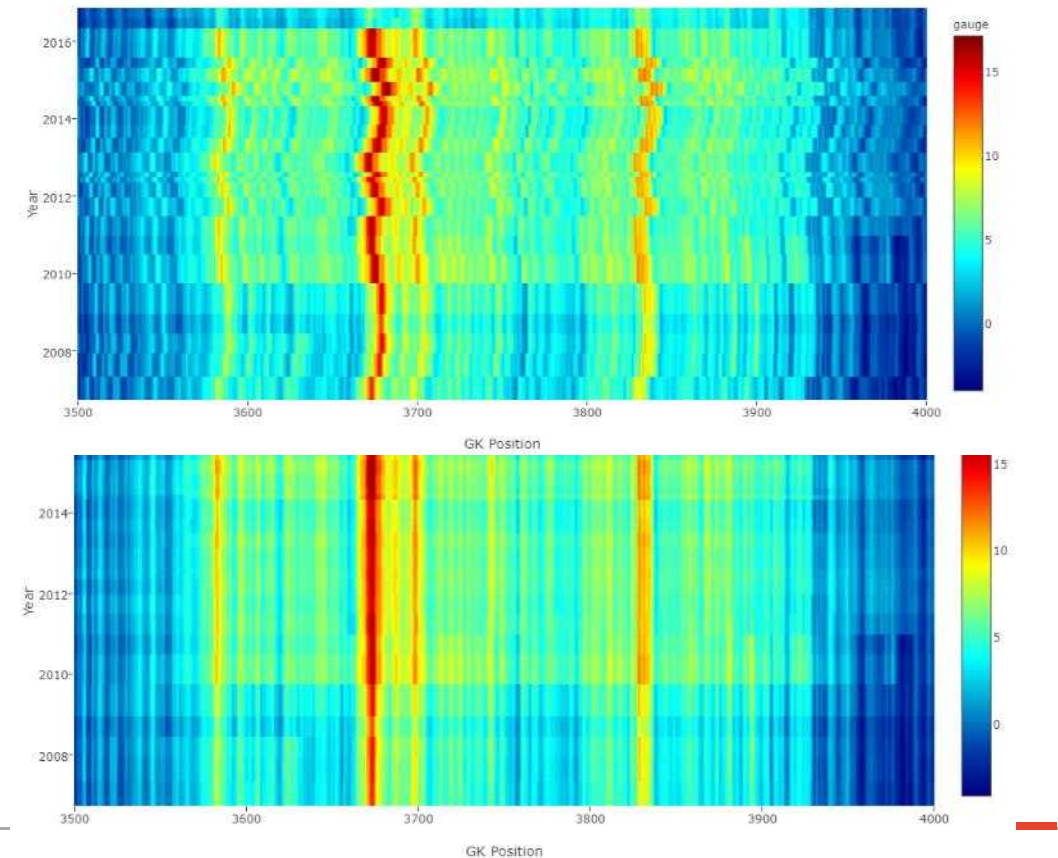
# Verortung – absolut und relativ

Finde den richtigen Fahrweg – durch Vergleich von gemessener zu Sollkrümmung

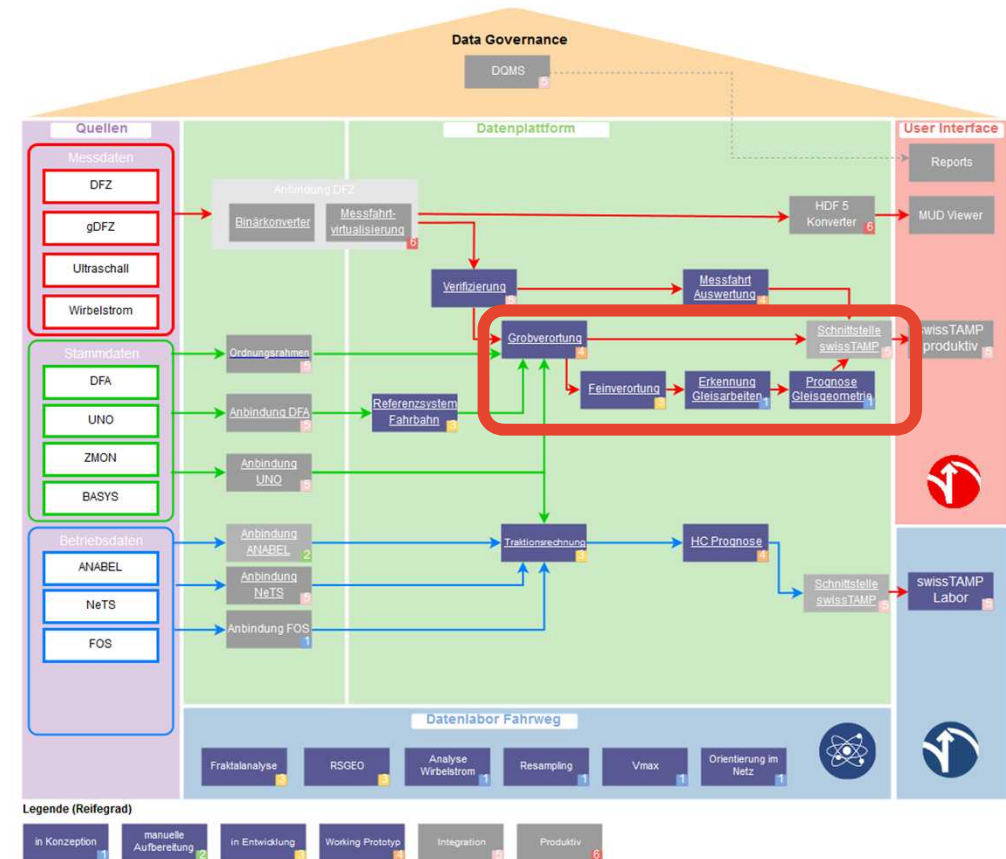
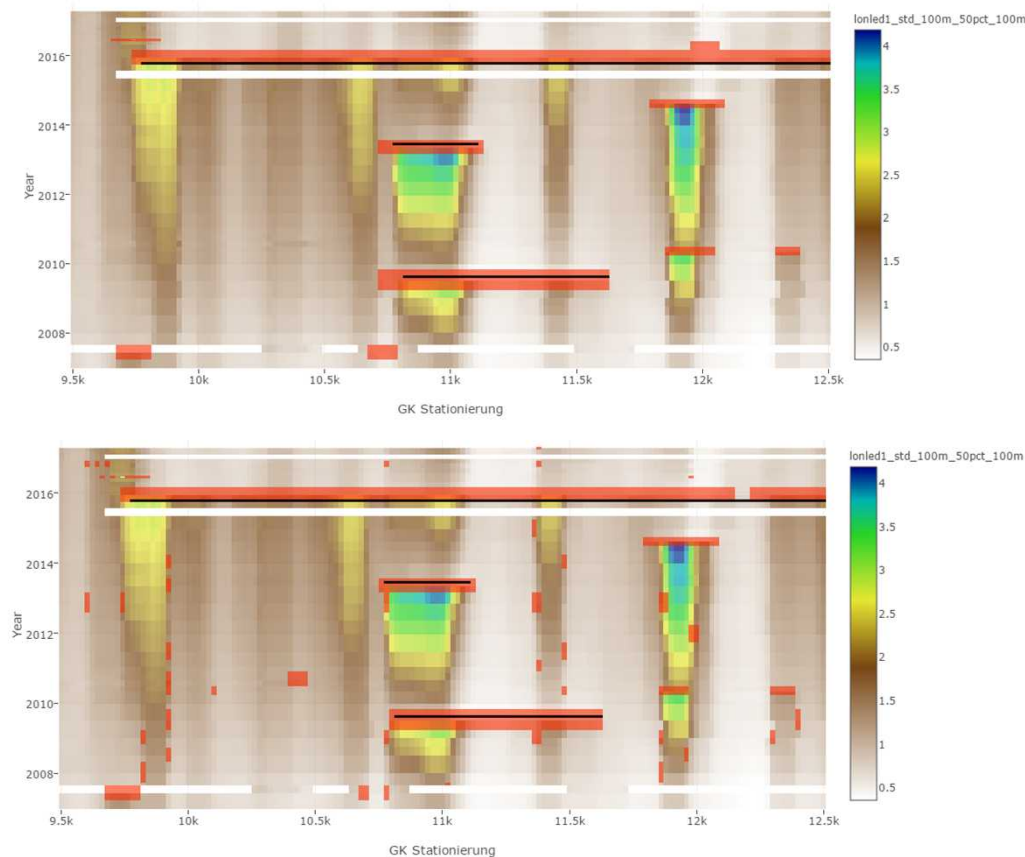


Anmerkung:  
Durch Investitionen an der Quelle (Messfahrzeug) bald nicht mehr notwendig.

Relative Verortung um Prognose zu ermöglichen – durch Korrelation



# Gleisarbeiten erkennen – Gesamtarchitektur



Aktuelle Weiterentwicklungen:

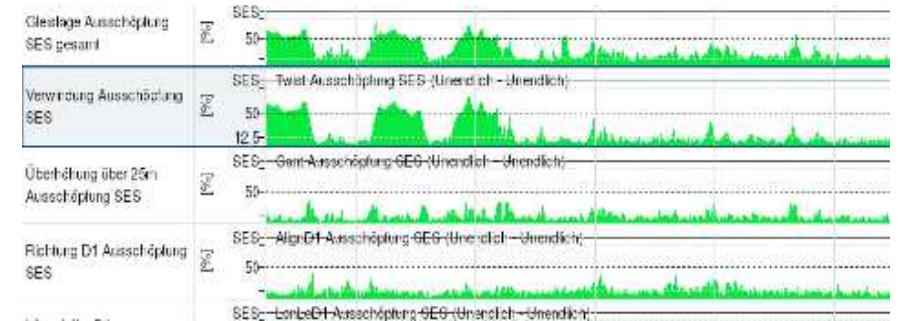
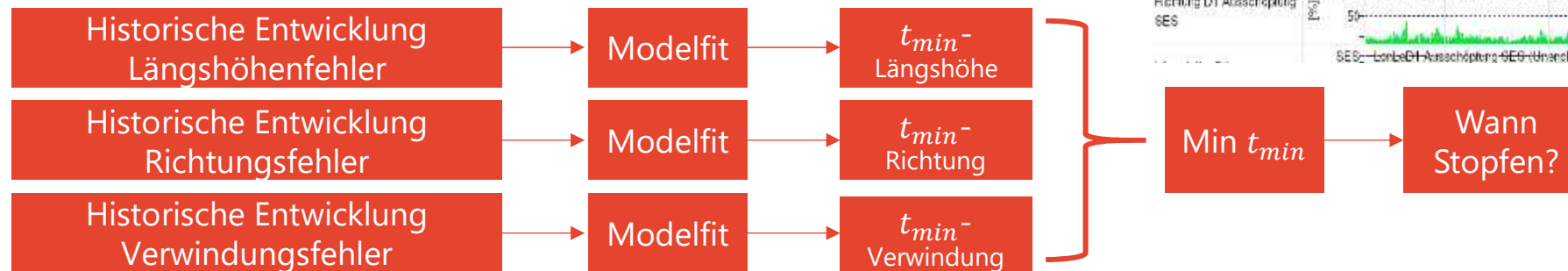
- Abgleich mit Befahrungen um Sperrungen zu erkennen
- Algorithmischer Abgleich mit DfA
- Maschineneinsatzpläne und -daten prüfen

# Vereinfachungen: Ausschöpfung und $t_{min}$

## Statische Sicht

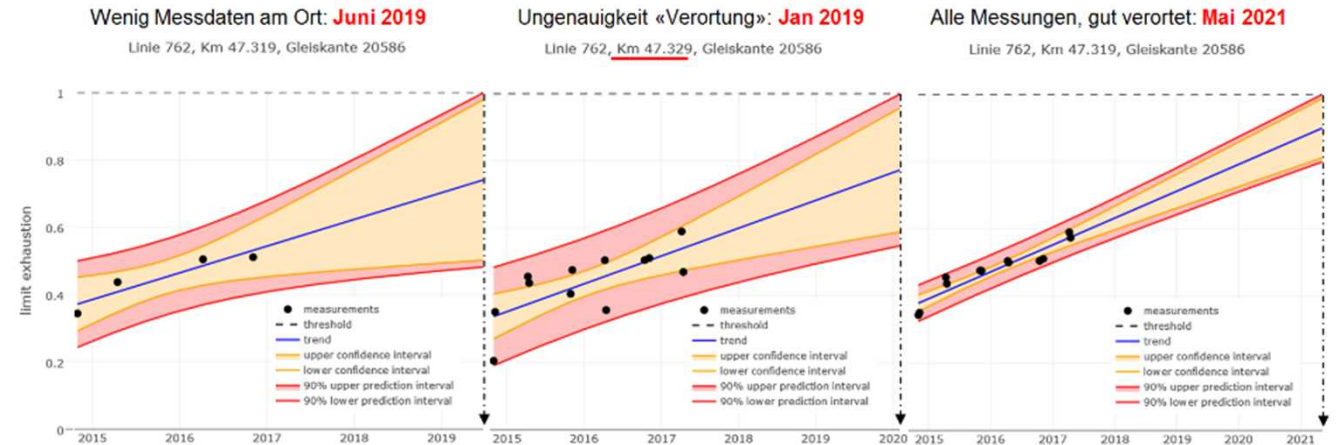
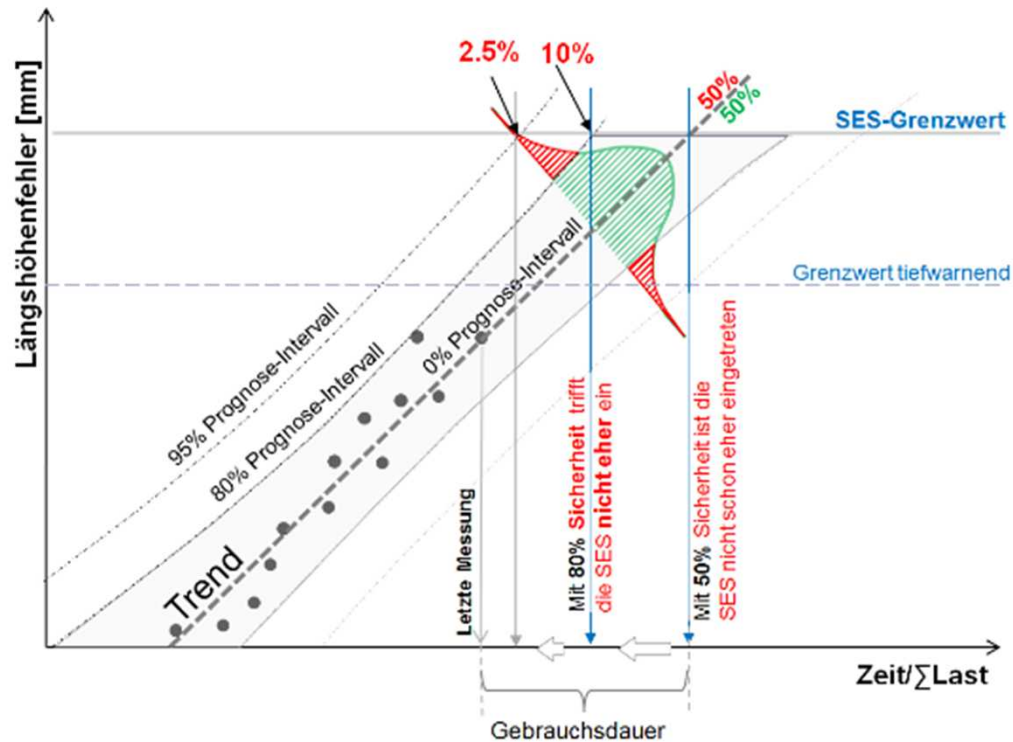


## Dynamische Sicht - Prognose





# Eine Frage des Risikos - Einflussfaktoren



➔ Investition in CTM und Datenqualität lohnt sich



# Rollkontaktermüdung (RCF)



# Was ist Rollkontaktermüdung?

---

- Rolllt ein Rad über die Schiene, dann können die dabei auftretenden Spannungen nach der Hertz'schen Theorie abgeschätzt werden.
- Die Berührfläche ist annähernd elliptisch. Die Radlast erzeugt Normalspannungen, zusätzlich treten unter der Schienenoberfläche Schubspannungen auf.
- Die Schubspannungen erreichen in einer Tiefe von etwa 6 mm ihren Größtwert. Die Normalspannungen erreichen bei einer 22,5 t-Achse Werte von bis zu 1300 N/mm<sup>2</sup> [87].
- Diese Spannungen liegen über den üblichen Festigkeiten von Schienenstählen. Damit wird die Schiene unter jeder Überrollung an der Berührfläche plastisch verformt.

(2010 Handbuch Gleis Lichtberger)



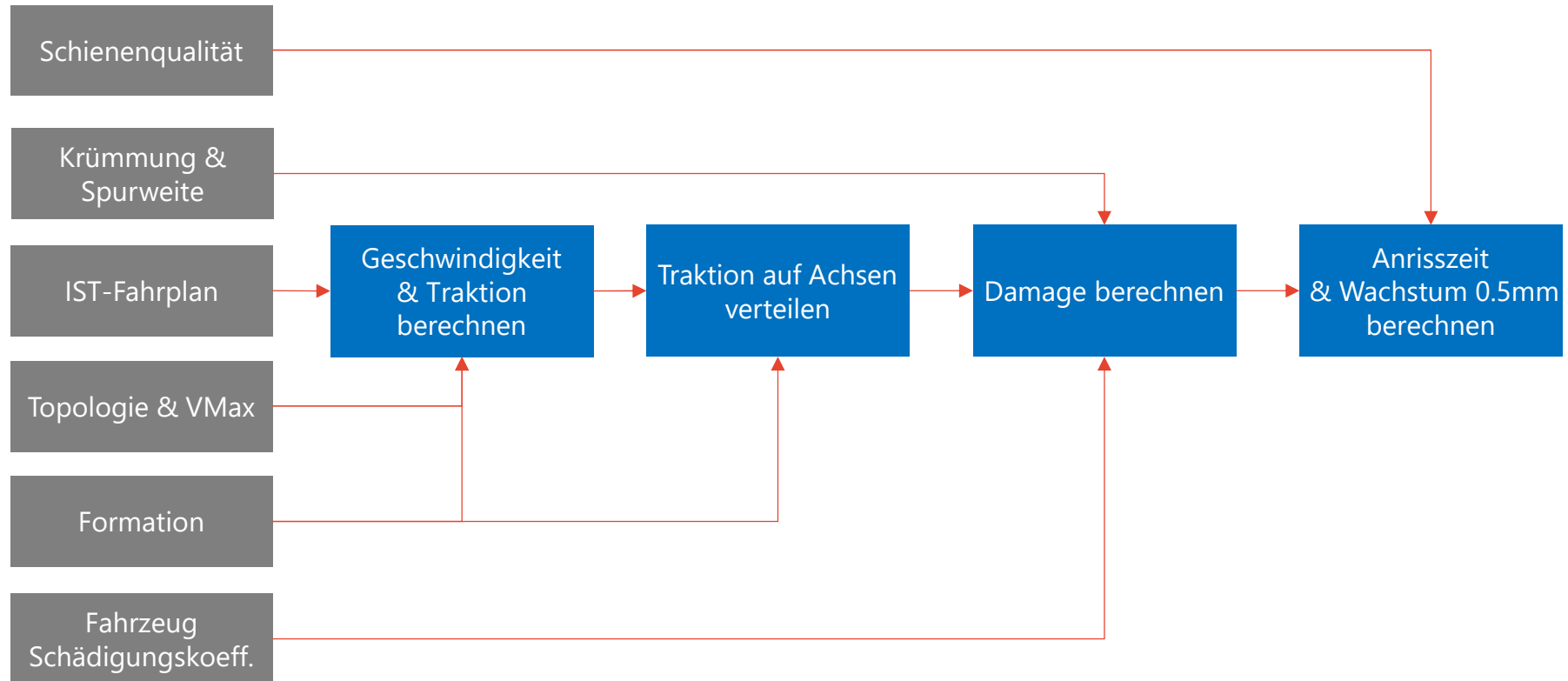
# Squat & Head-Check

## Können diese Rollkontaktermüdungen prognostiziert werden?

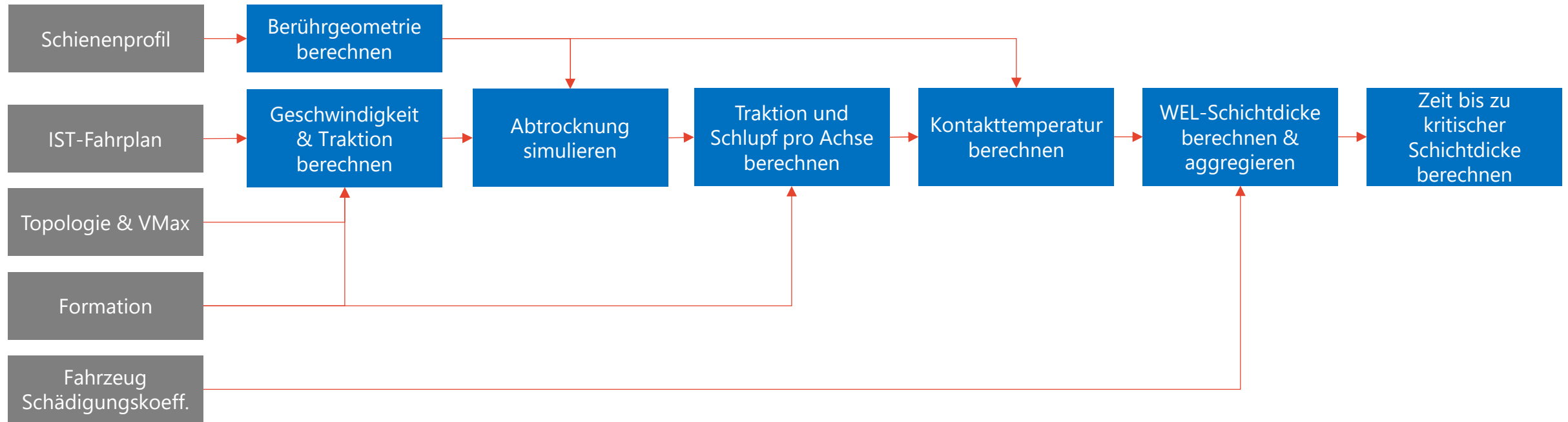


Trümmerbruch Hatfield/England, Okt. 2000,

# Berechnungslogik Head-Check

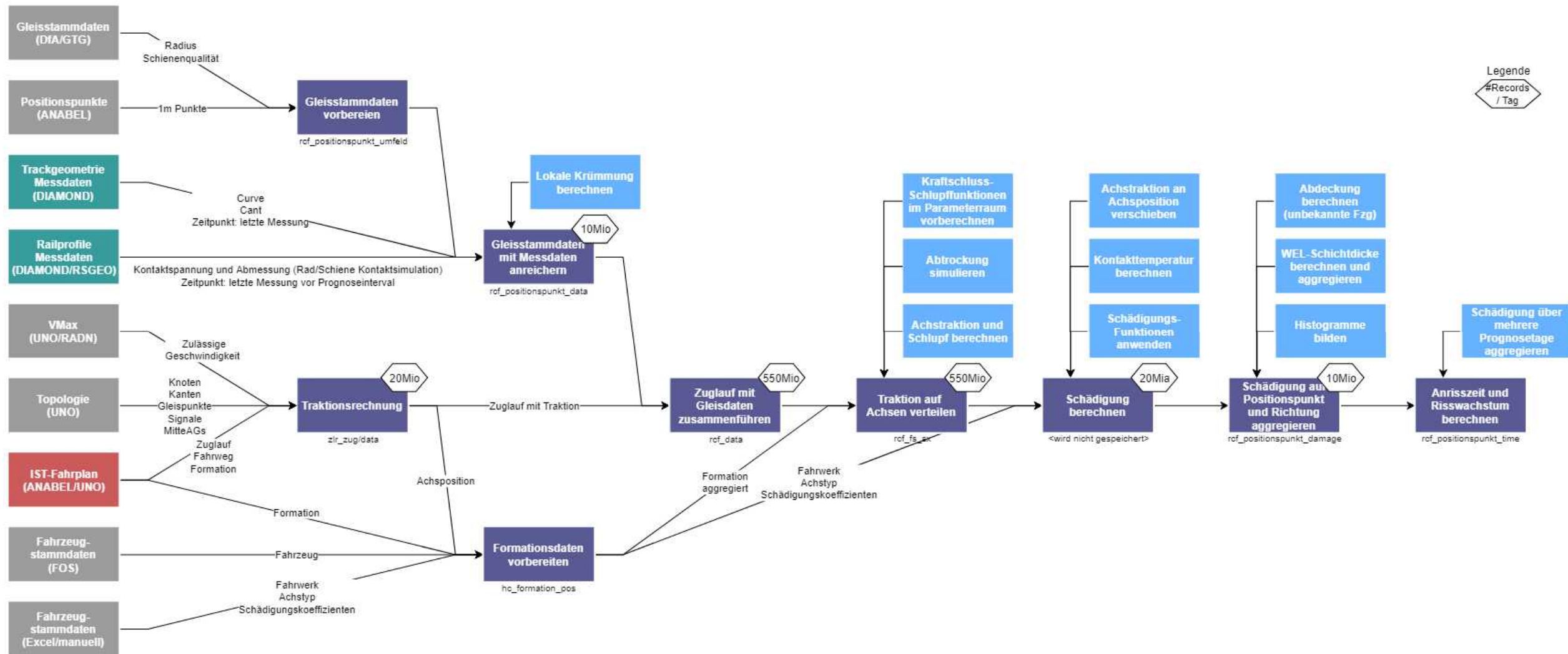


# Berechnungslogik White Etching Layer

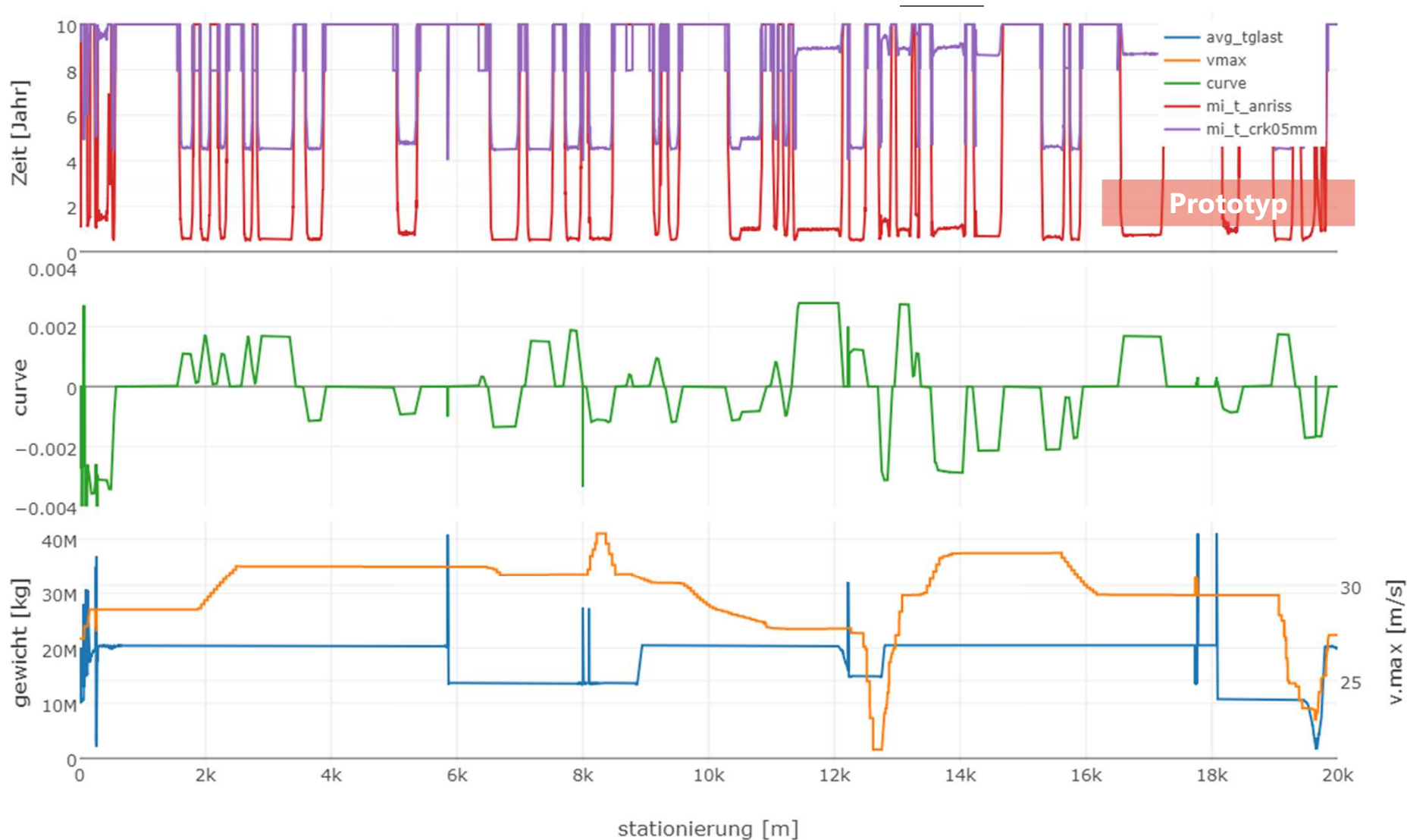




# Datenfluss Prognose Rollkontaktermüdung: eine Datenschlacht

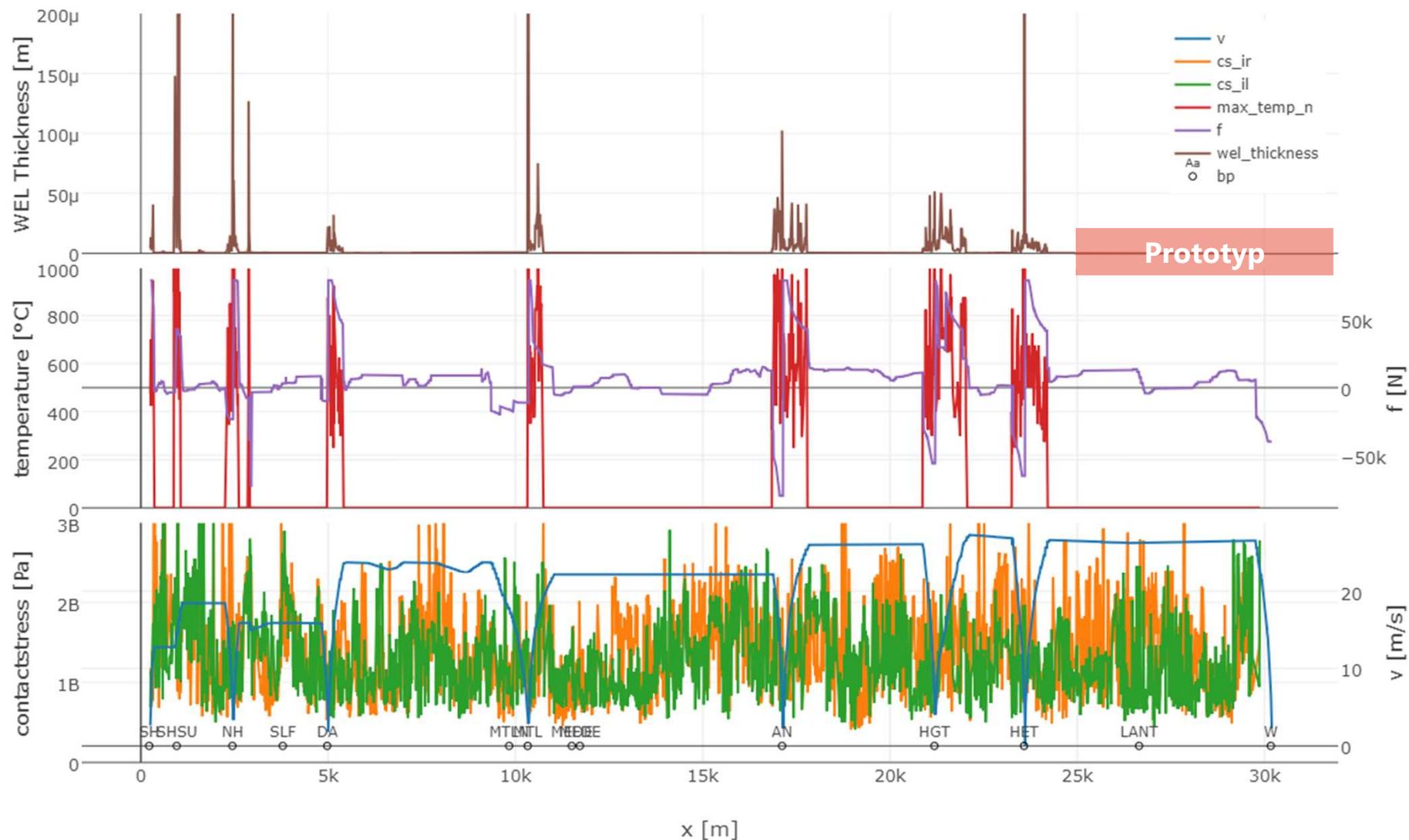


# HC: Head-Check Zeit bis kritischer Risstiefe



Berechnete Anrisszeit und Zeit bis Risse 0.5mm tief sind, hochgerechnet aus einem IST-Fahrplan Tag.

# WEL/Squats: White Etching Layer Kontakttemperatur und Schichtdicke



Berechnete WEL-  
Schichtdicke für einen Zug

Berechnete Temperatur  
aus Reibenergie und  
Kontaktfläche



# Herausforderung Performance

## Performance für Prognoserechnung mit einem Tag

- 20 Milliarden Records/Tag = 1.4TB Daten (pro Achse und Meter ein Schädigungspunkt)
- 200 Cores -> 15h Rechenzeit (skaliert horizontal)

## Lessons Learned

- BigData Technologie bringt Faktor 1000x
- Wenn bei 500Mio Records ein Funktionsaufruf 300ms dauert, sind das 27'000h Berechnungszeit, das geht auch mit BigData nicht
- Old-School Performance Tuning ist nötig:
  - > Collection-Typ sorgfältig wählen
  - > Mehrfach verwendete Zwischenresultate vorberechnen
  - > Schritte im Datenfluss sorgfältig designen
  - > Profiling, Hardwarebeschleunigung, Bugs...

A railway worker in a bright orange safety suit and helmet is kneeling on gravel tracks, working on a rail joint with a tool. A wooden toolbox lies on the ground nearby. In the background, two other workers in similar gear stand near railway infrastructure. The scene is overlaid with a semi-transparent dark blue filter and white text.

# **Infrastruktur und Framework**

## **Daten effizient aufbereiten und bereitstellen**

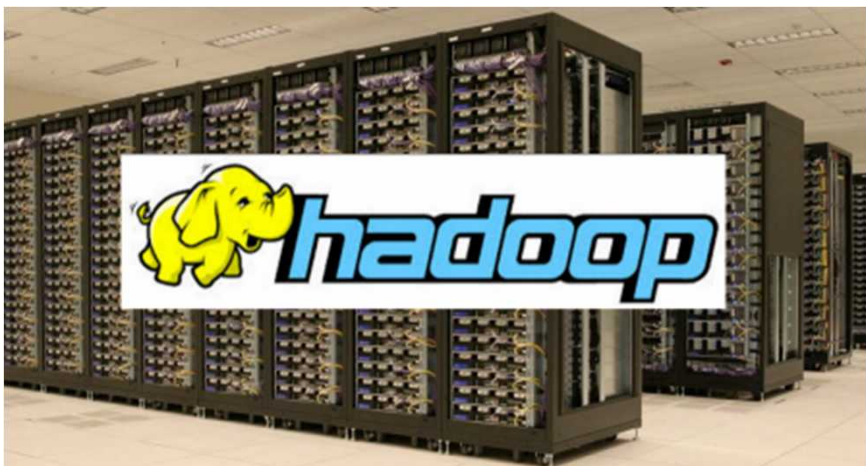
# Data & Analytics Platform (SBB/Datenlabor Fahrweg)



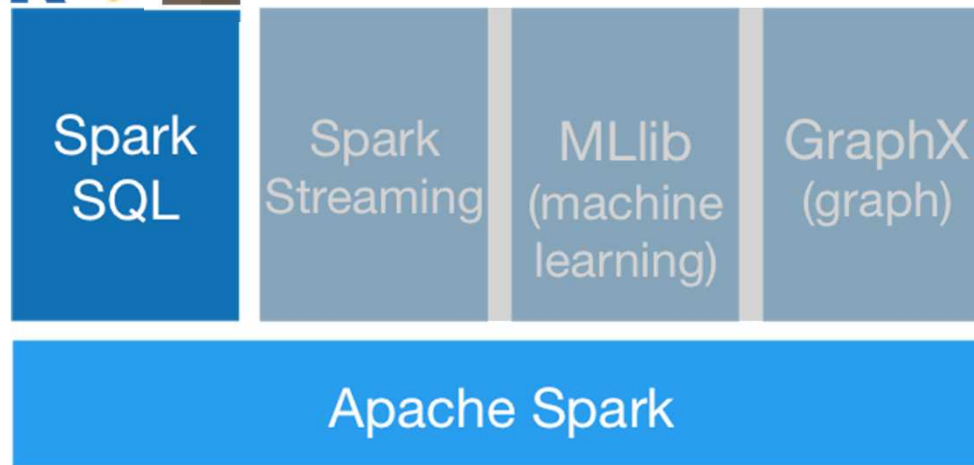
## SPARK NOTEBOOK

Webbasiertes Data Science Tool  
für Spark/Scala und Visualisierungen

cloudera



16 data-nodes, 420 computing cores  
2TB RAM, 160TB Netto HDFS storage



Strukturierter Datalake  
mit versionierten Daten der  
relevanten Quellsysteme

SQL





# Min. 80% des Aufwands ist Datenaufbereitung

March 2015

data preparation means:

- find data sources
- interpret, ingest, clean and harmonize data
- check data quality
- enhance data (sometimes with advanced algorithms)
- industrialize data feeds (keep data up-to-date)

March 2017

Mai 2017

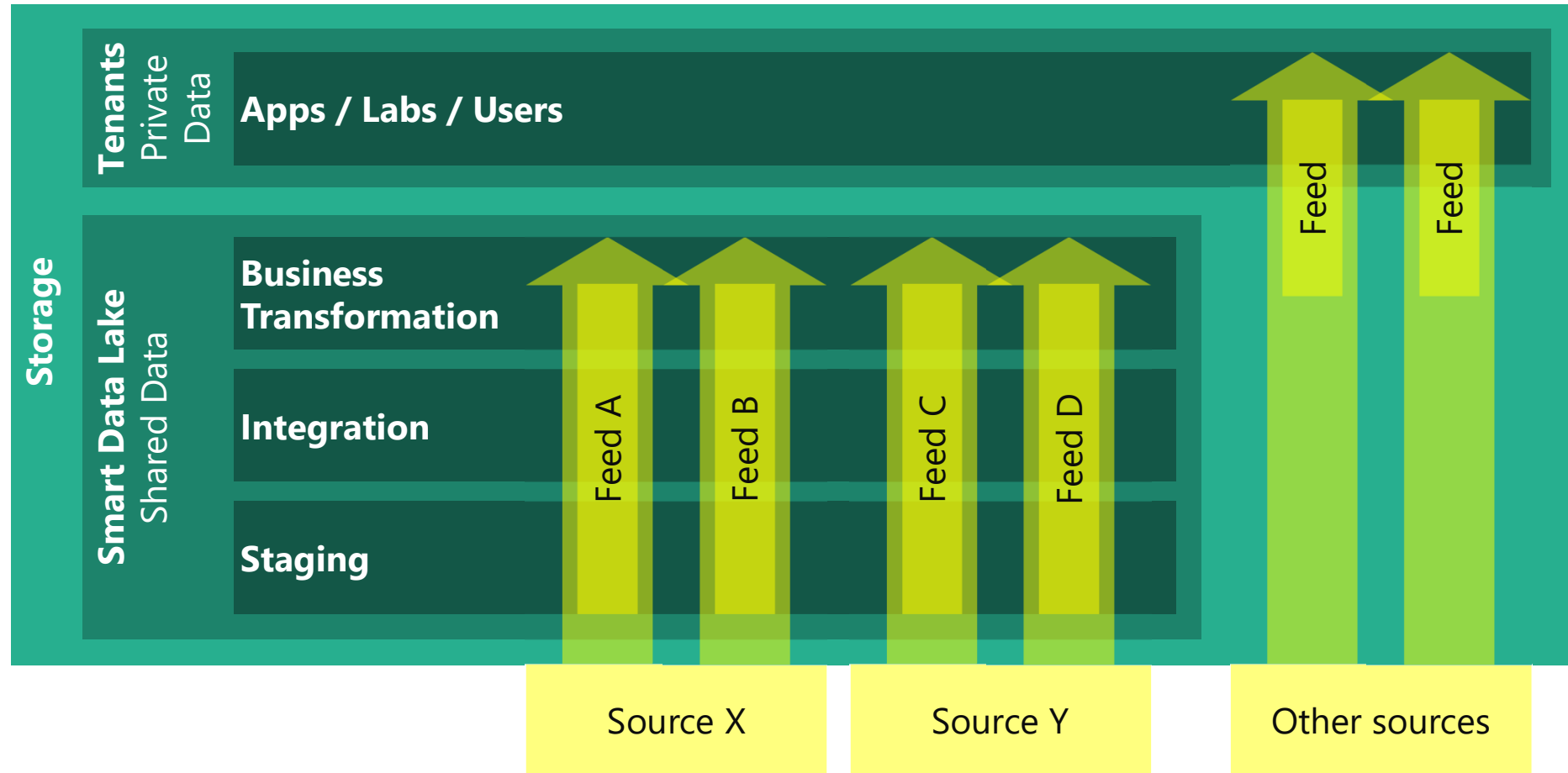
Data preparation

Data preparation

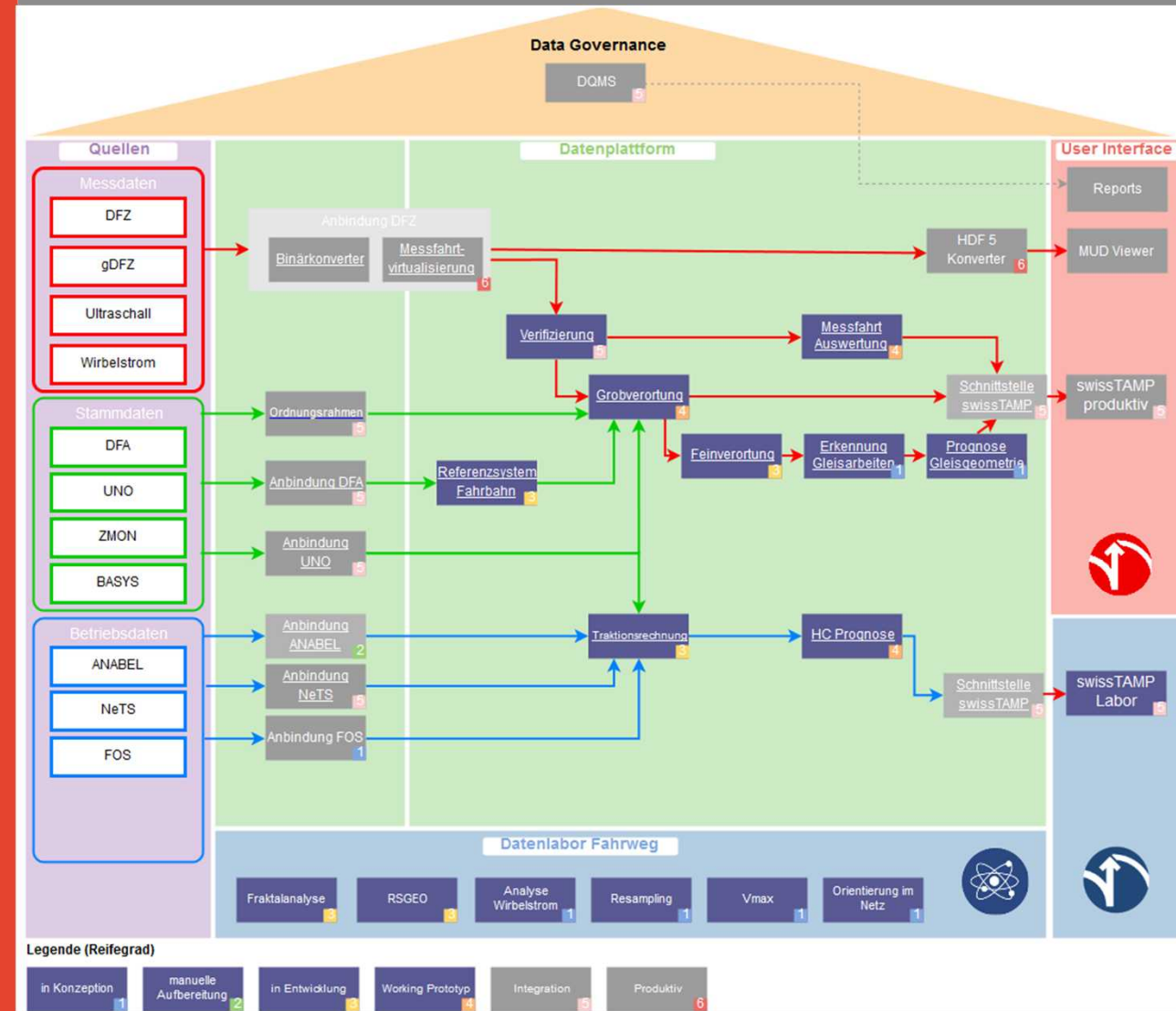
## Data Science:

PoC for statistical prediction of defects done in 2 month.  
Result where promising, but further data preparation is needed...

# Aufbereitete Daten bereitstellen -> Smart Data Lake

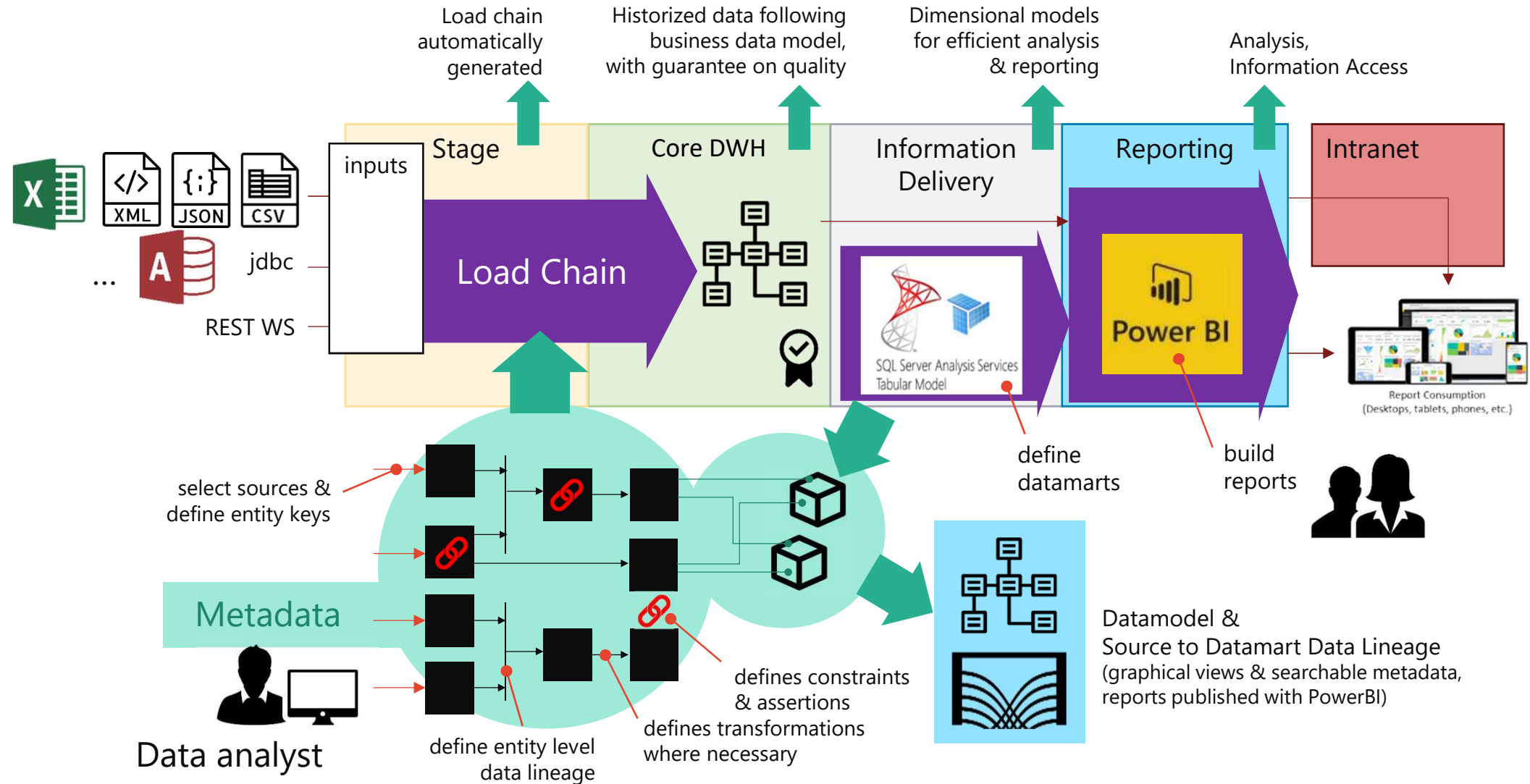


Und wie können wir die Komplexität der Abhängigkeiten in den Griff bekommen?





# Von DWH's lernen -> Automatisierung über Metadaten



# Baustein Sourcing Framework: Automatisierung über Metadaten

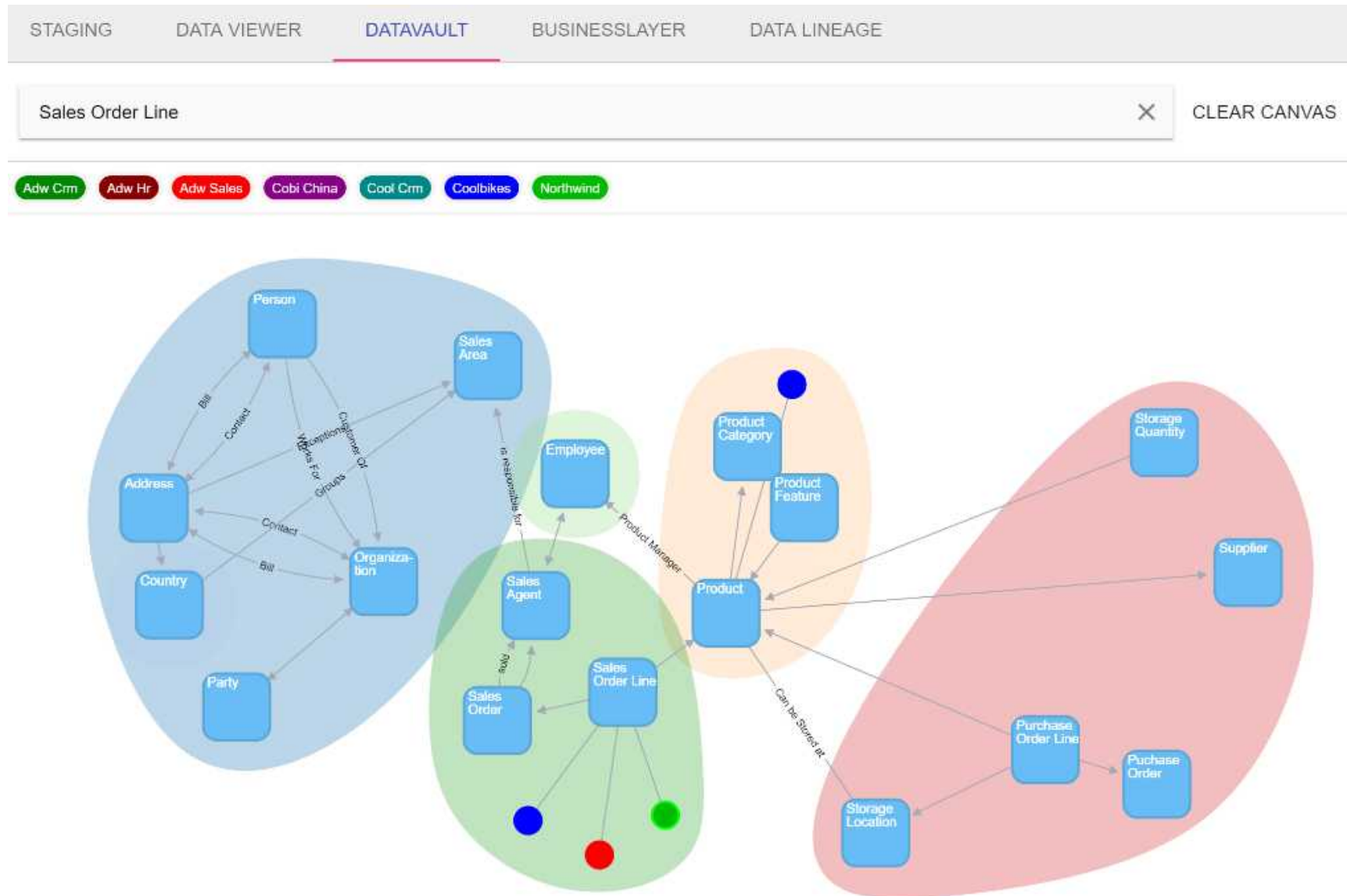
Deklarativer meta-data getriebener ETL approach:  
Beispiel Konfiguration vom SBB Sourcing Framework:  
-> Hive Table lesen und nach JDBC schreiben.

```
feeds=[  
  {  
    name = test  
  
    input = [{  
      type = hive  
      db = app_datalake  
      tables = [{name = btl_uno_h_gleiskante}]  
    }]  
  
    output = [{  
      action {  
        type = copy  
        name = First  
      }  
      type = jdbc  
      db = GSU_ANABEL  
      tables = [{name = anabel_test}]  
      jdbc-driver = oracle.jdbc.OracleDriver  
      jdbc-url = "jdbc:oracle:thin:yyy/xxx@ (DESCRIPTION= (ADDRESS= (PROTOCOL=tcp..."  
    }]  
  }  
]
```

SBB Sourcing Framework Features:

- Sources für JDBC, Webservice, MQ, CSV, Excel...
- Generic Actions für Historization, Deduplication
- Hooks für Custom Logic
- Schema Evolution
- Security (Kerberos, ACL's)

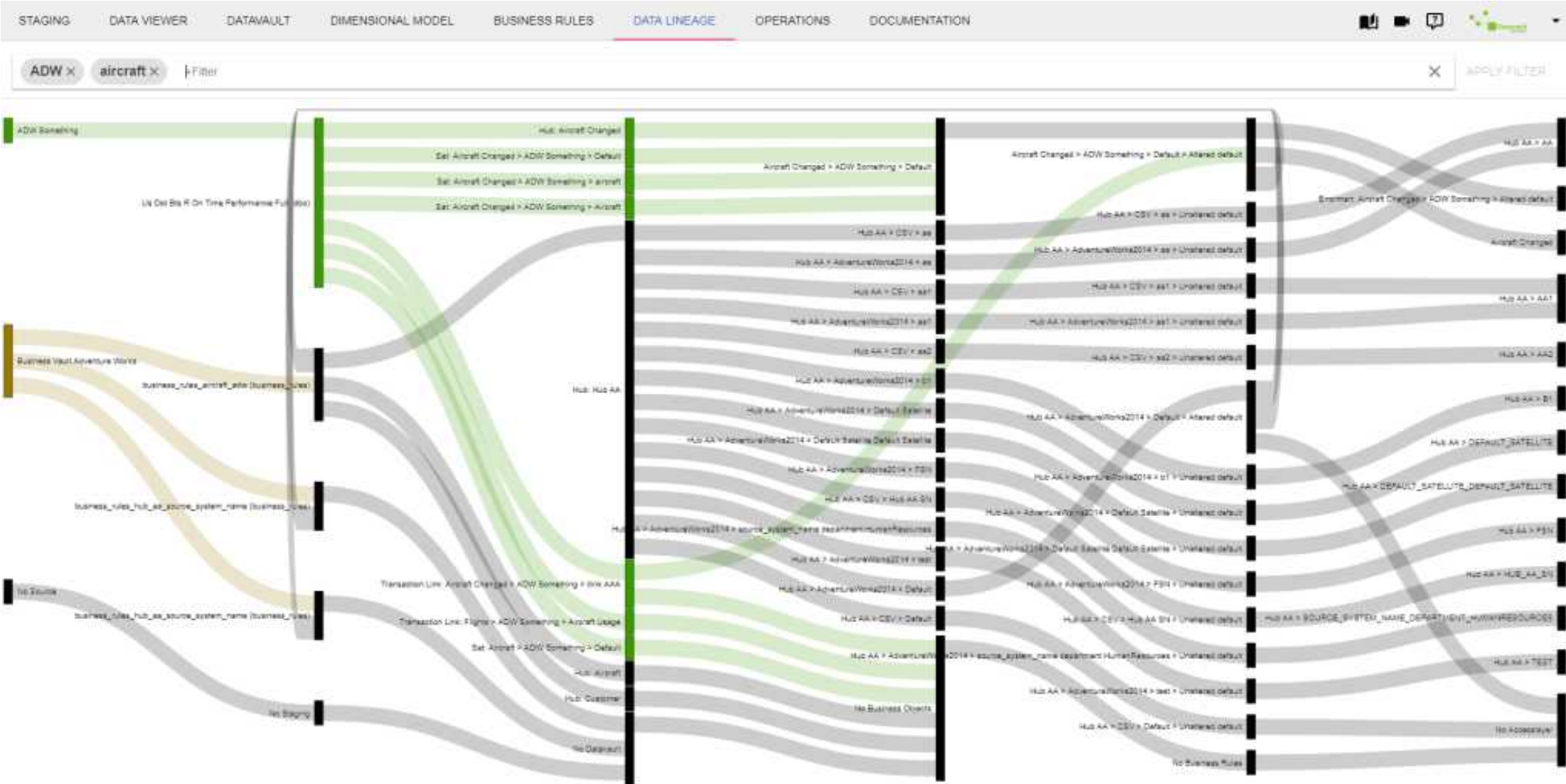
# [Future] Baustein «Data Catalog» -> Self-service



Screenshot from Datavault Builder



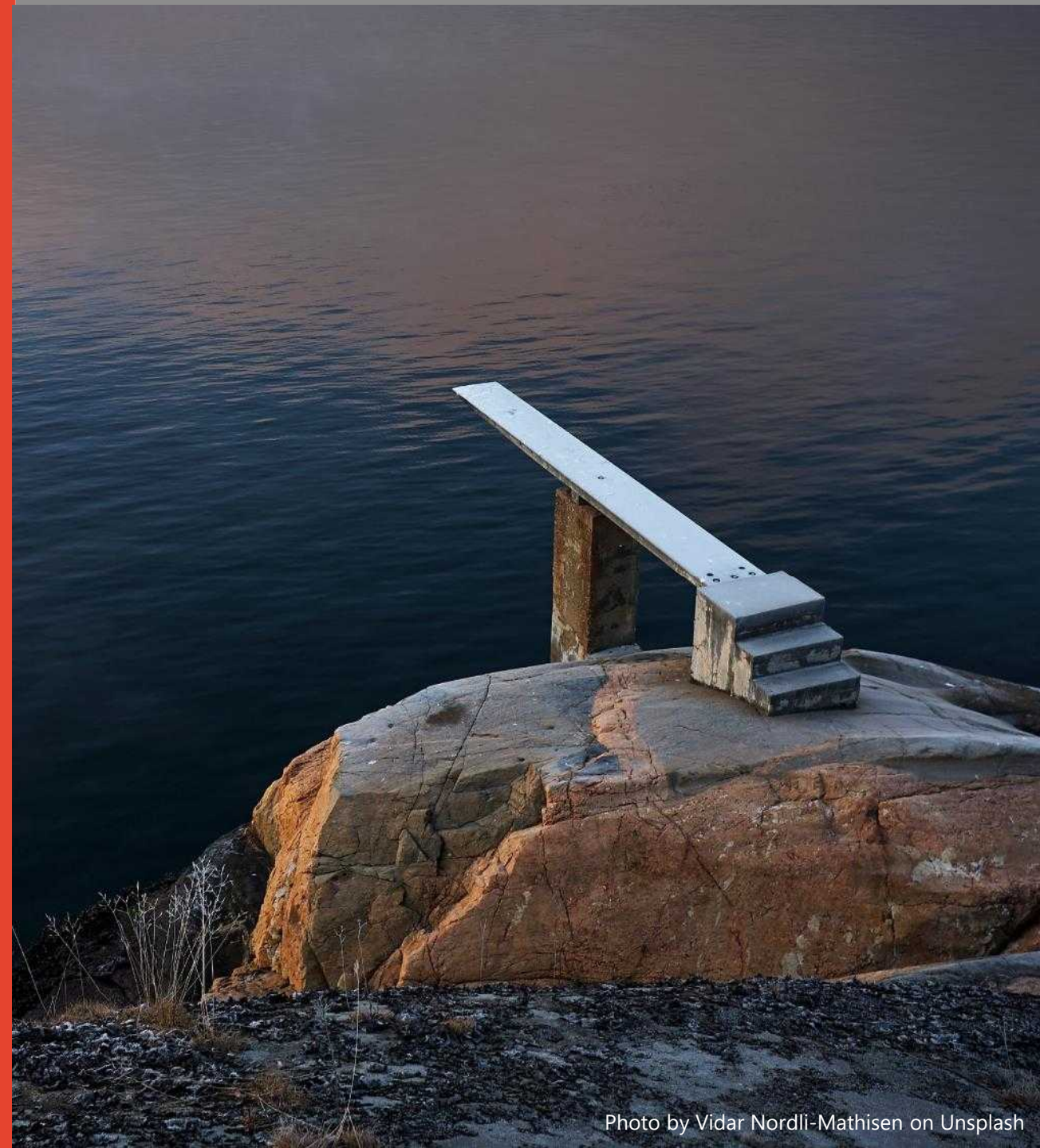
# [Future] Baustein «Workflow, Lineage und Überwachung» -> effizienter Betrieb



Screenshot from Datavault Builder

# Smart Data Lake Builder

ELCA & SBB-IT überlegen sich, das Smart Data Lake Framework als **Open Source** zu veröffentlichen und gemeinsam weiter zu entwickeln.







## Fazit

Komplexität erkennen und managen

Abhängigkeit von Stammdaten und deren Qualität berücksichtigen – und investieren

Produktiv setzen herausfordernd – technologisch und organisatorisch

Leistungsfähige Plattform notwendig für Datenmenge und Berechnungen





Fragen?