Big data analytics and IoT in railway as enabler for digital rail services

Professor Diego Galar

Head of Maintenance & Reliability, Tecnalia

Lulea *University* of technology

Big data analytics approaches O&M

Globalization and Globally Available Resources

Billions of mobile devices accessing the Web

Access to streams of information in the Real Time

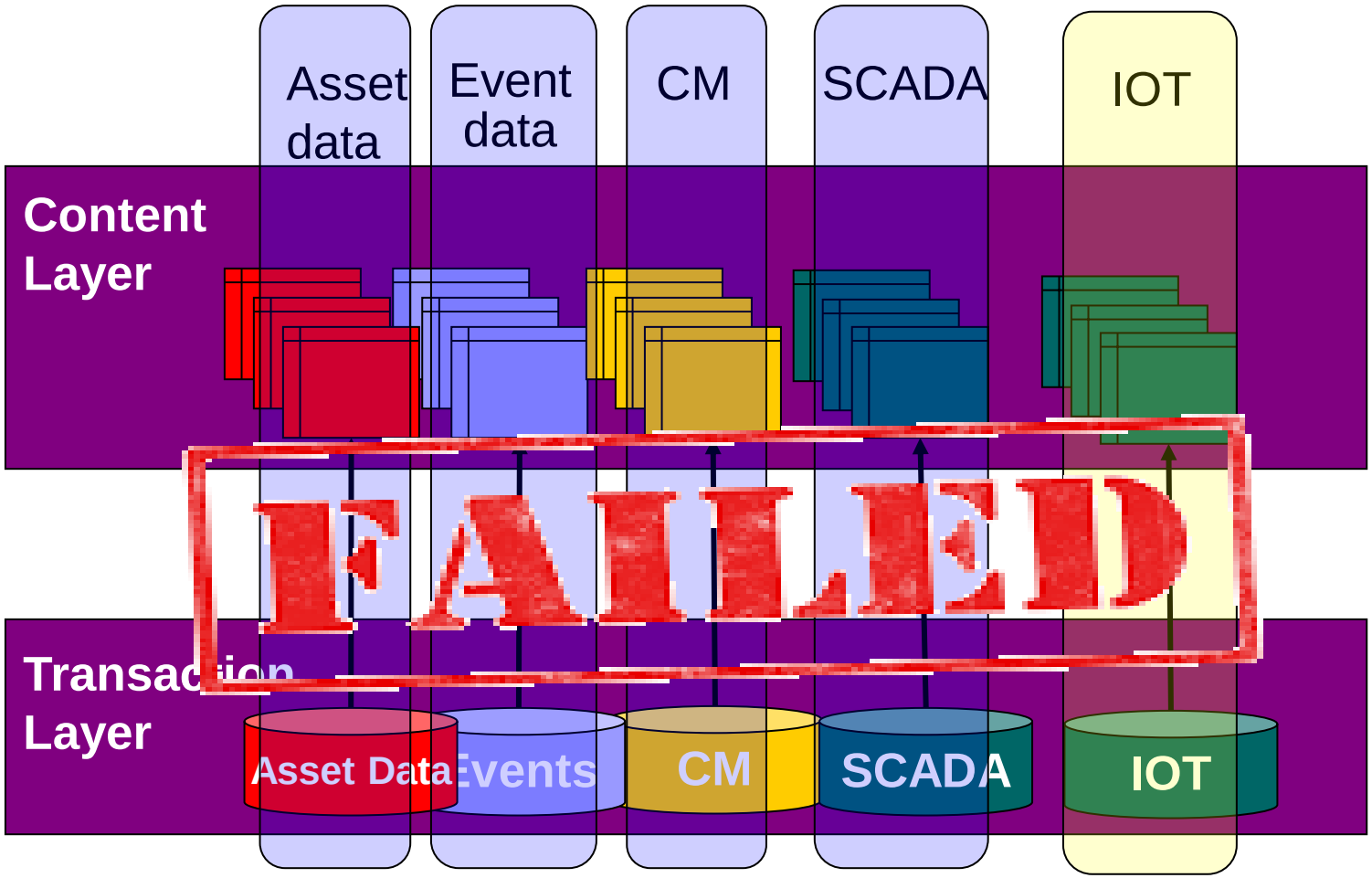
New Forms of Collaboration

New possibilities.
New complexities.
New risks.

Big Data in railway for O&M is mostly machine generated data

Volume | Velocity | Variety | Variability

What data we gather and what connections are needed?



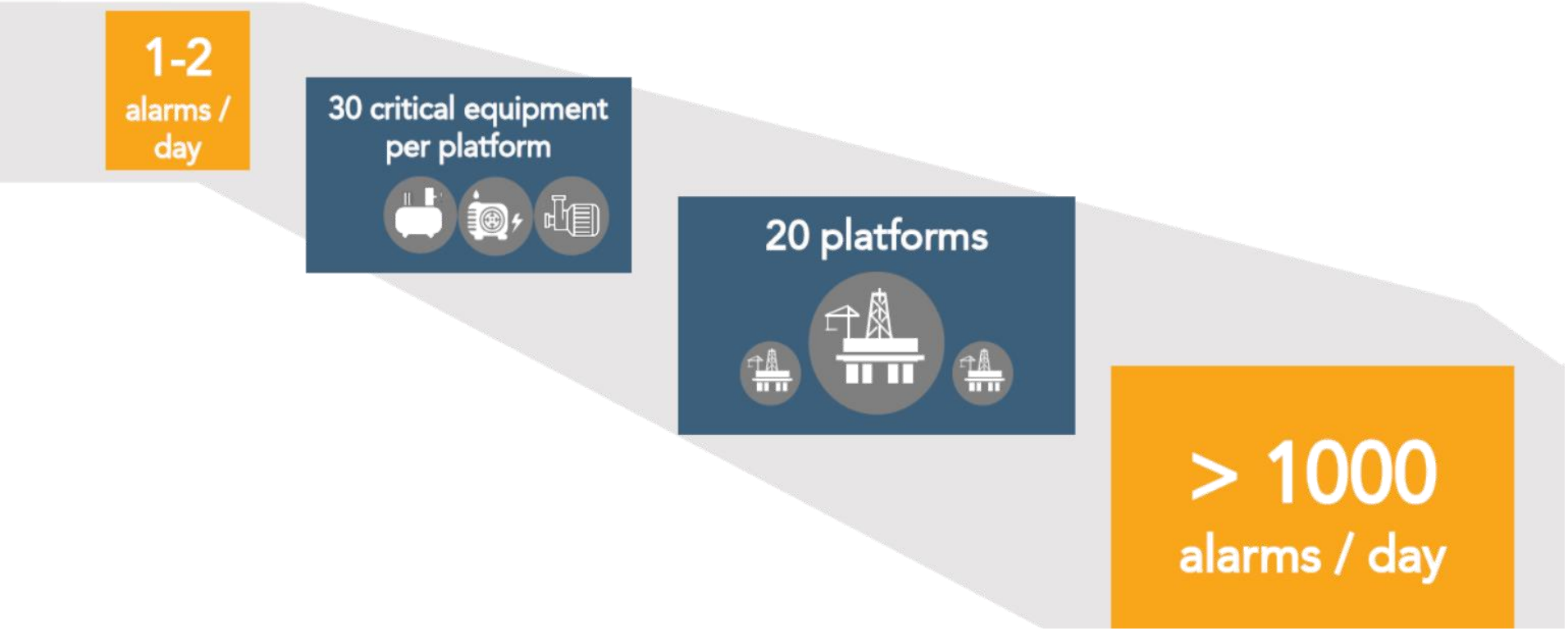
Silos of data by functional area

tecnalia Inspiring Business

LULEÅ UNIVERSITY OF TECHNOLOGY

Pervassive computing

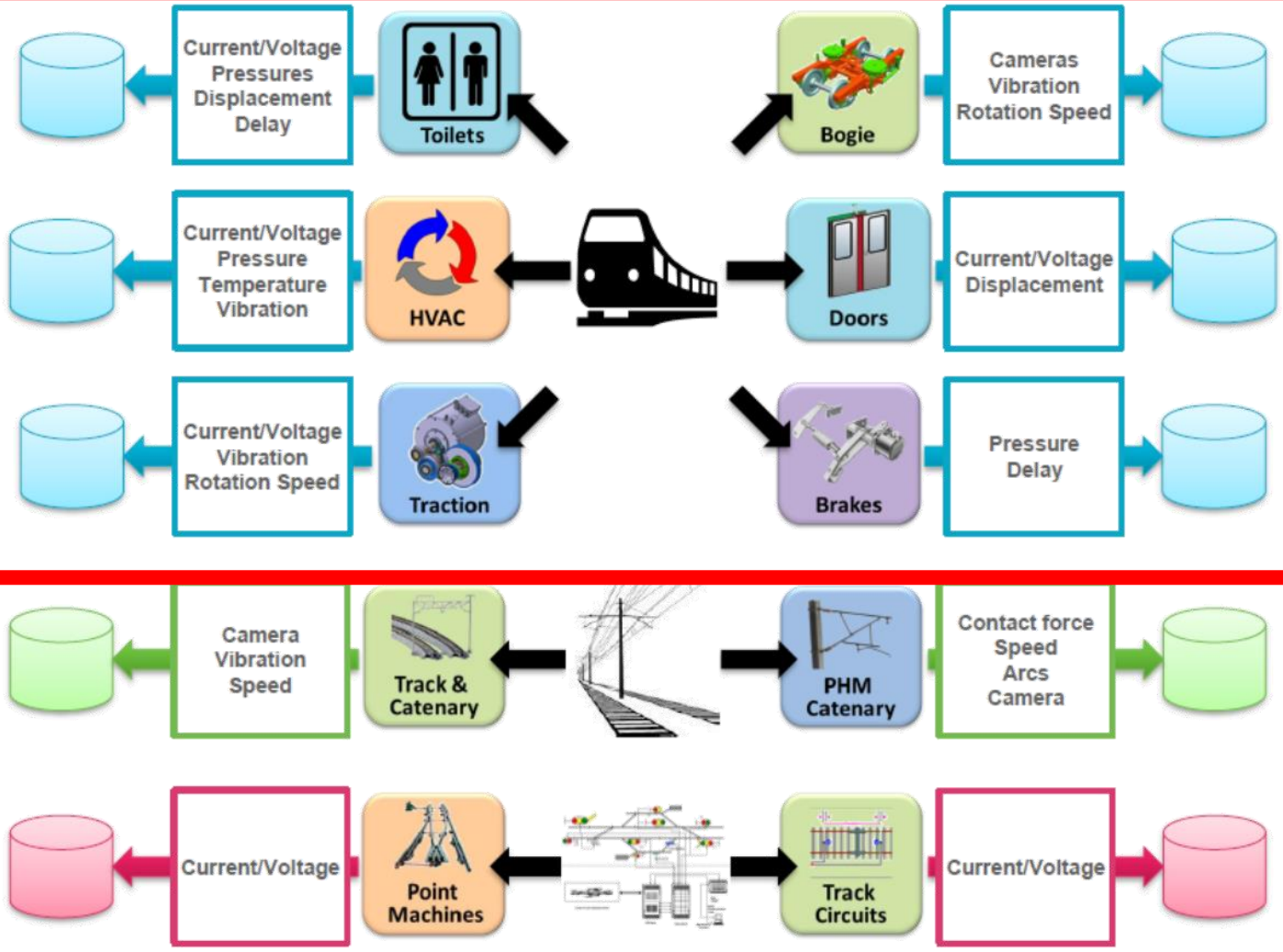
Alarm fatigue



tecnalia Inspiring Business

LULEÅ UNIVERSITY OF TECHNOLOGY

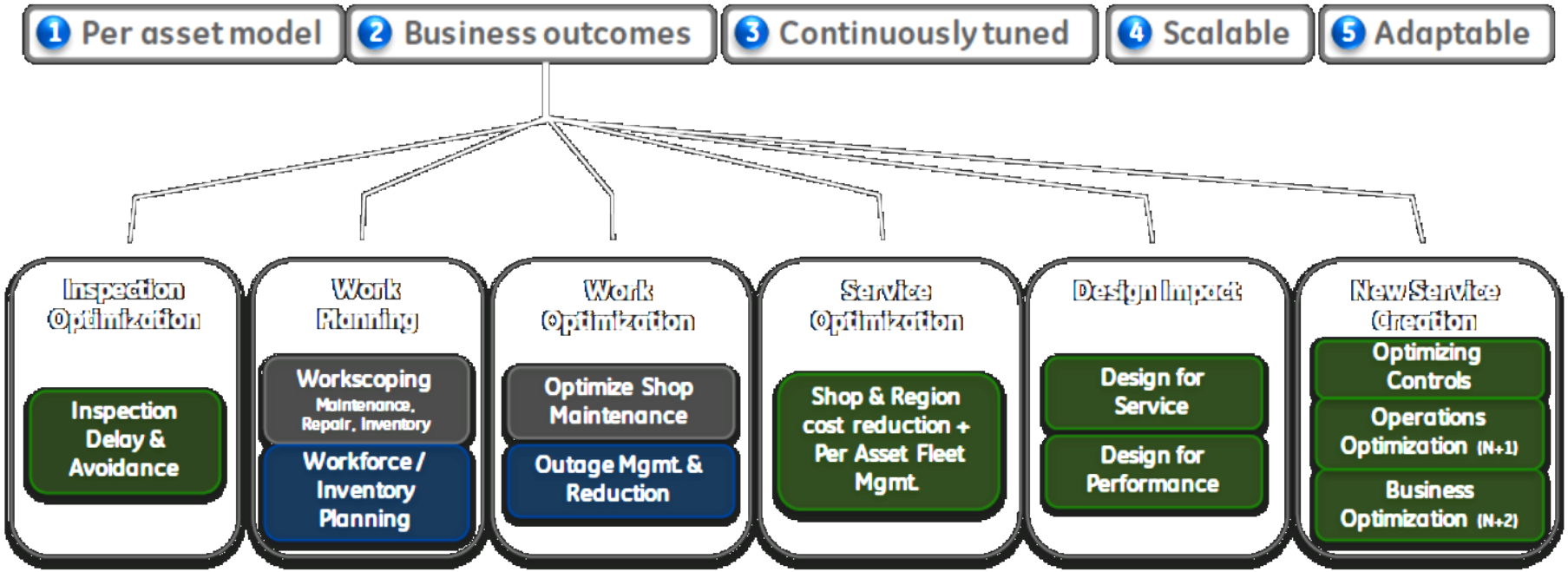
Dealing with most subsystems of a Railway network



tecnalia Inspiring Business

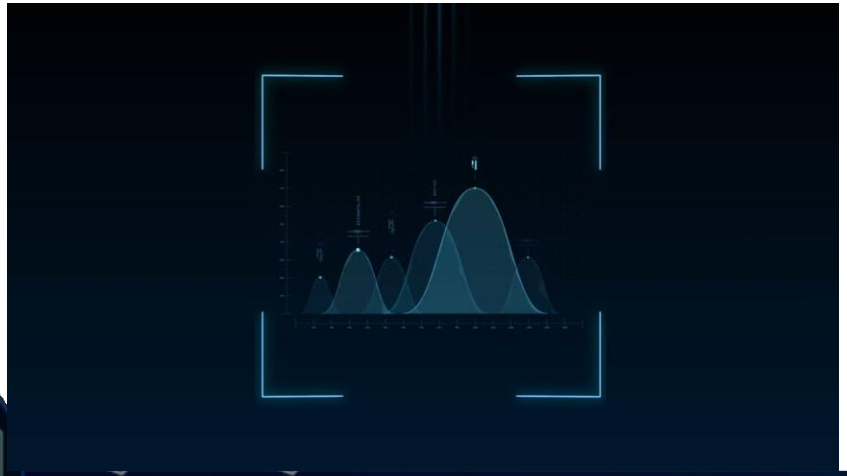
Digital Twin – Definition

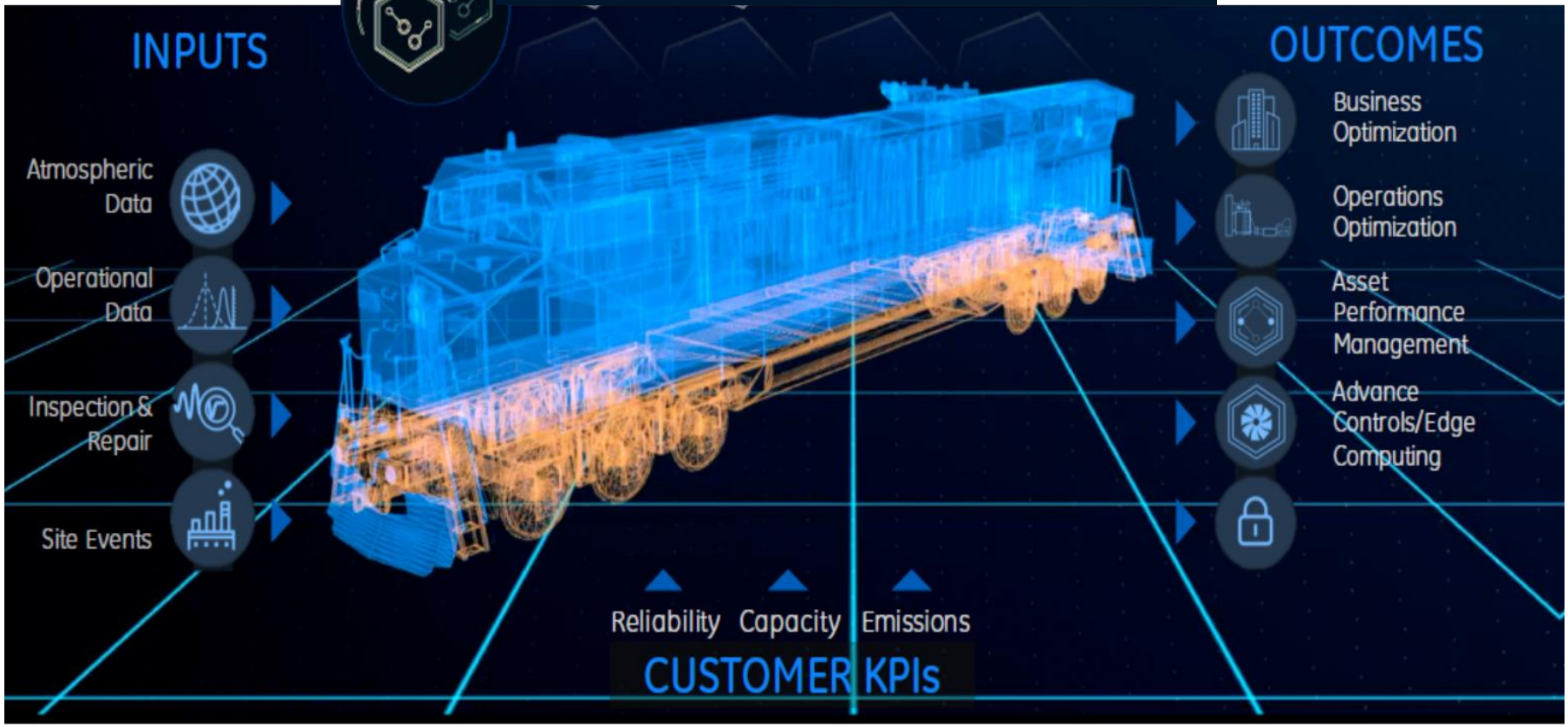
Engineering models that **continuously increase insights** into each asset to deliver specific **business outcomes**



tecnalia Inspiring Business

LULEÅ UNIVERSITY OF TECHNOLOGY

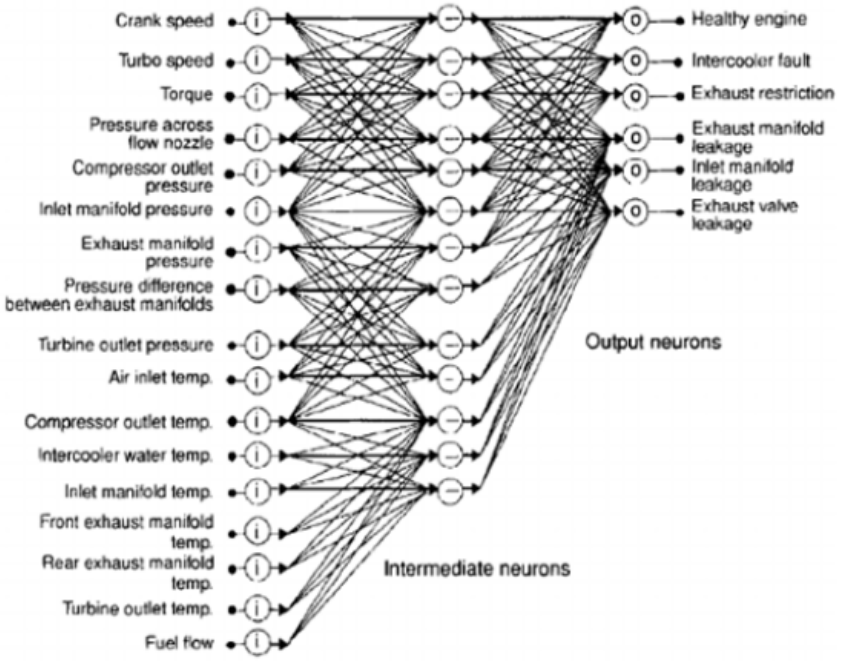


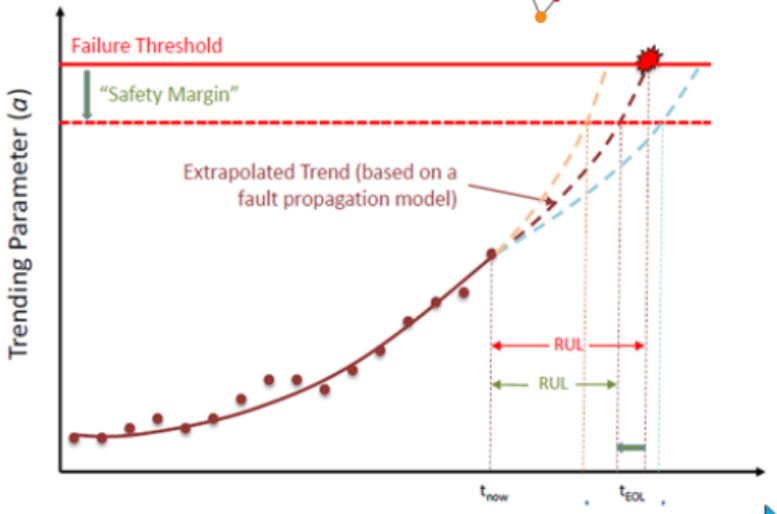


tecnalia Inspiring Business

LULEÅ UNIVERSITY OF TECHNOLOGY

Digital twin 1.0

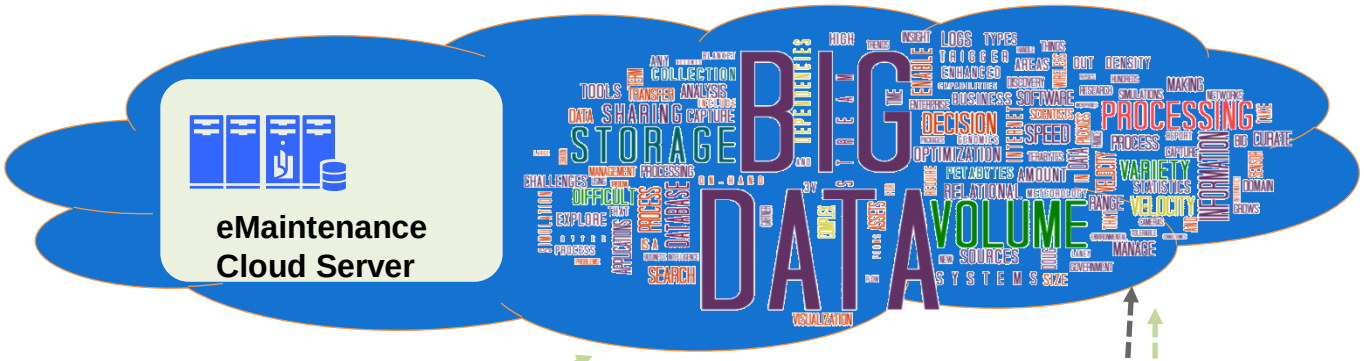


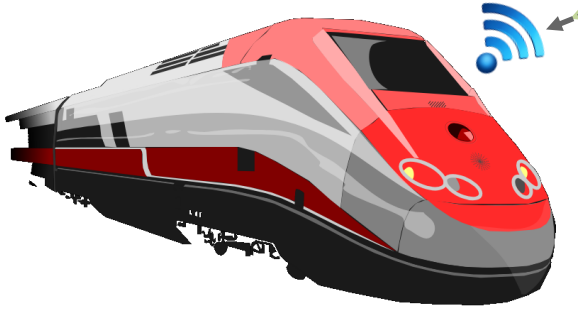


Diagnostics

Prognostics

Digital twin 1.0





On board Wireless System

Machine Maintenance Analytics

Data → Information → Knowledge

tecnalia Inspiring Business

LULEÅ UNIVERSITY OF TECHNOLOGY

Smarter: Anomaly detection

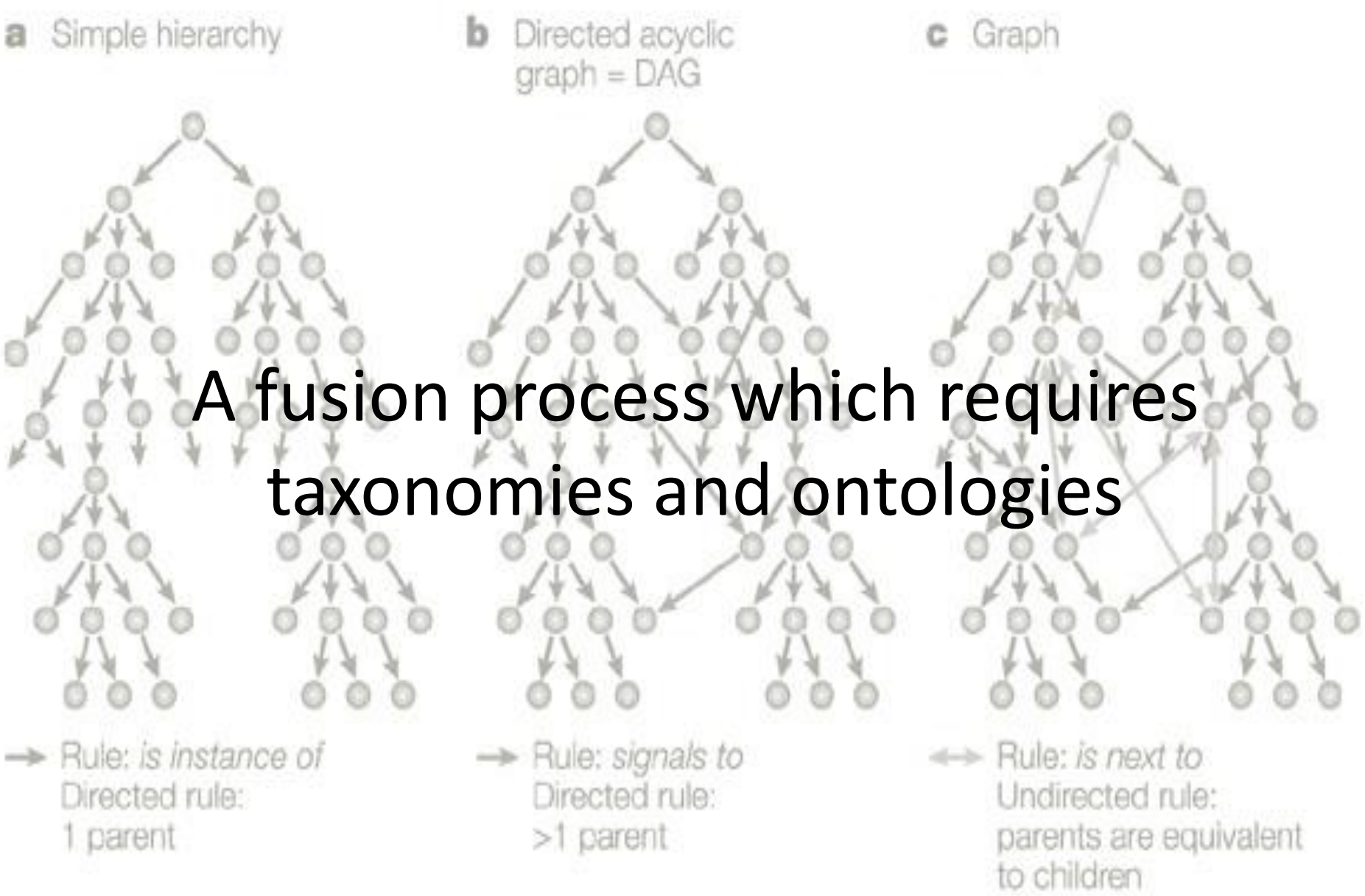
100+ sensors and derived sensors

> 99 % normal behaviour

One virtual sensor of "normality"

Deploy model, connect to live data, send notifications

What about IT systems?



A fusion process which requires taxonomies and ontologies

Taxonomy vs. Ontology

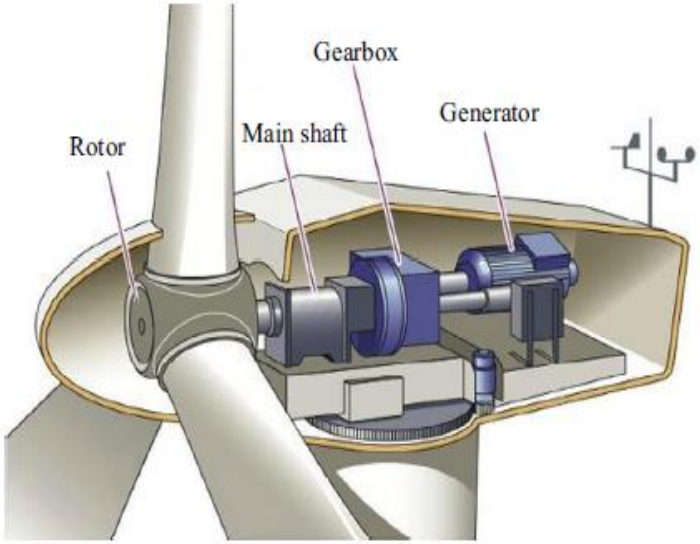
Taxonomies:

- Usually are a single, hierarchical classification within a subject
- Primarily focused on "is-a" relationships between classes
- Limited in inferencing potential due to lack of relational expressiveness.

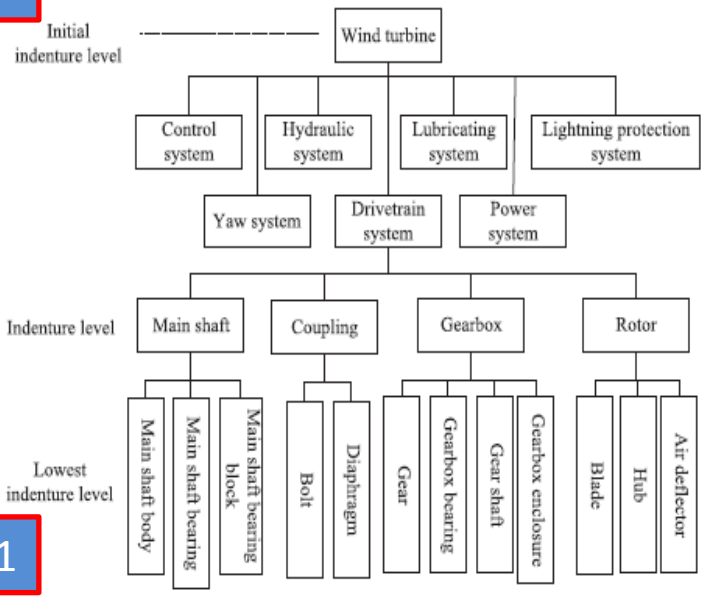
Ontologies:

- Subsume taxonomies.
- Include attributes with cardinality and restricted values.
- Unlimited relationships between entities.
- Superior inferencing support due to relational expressiveness.

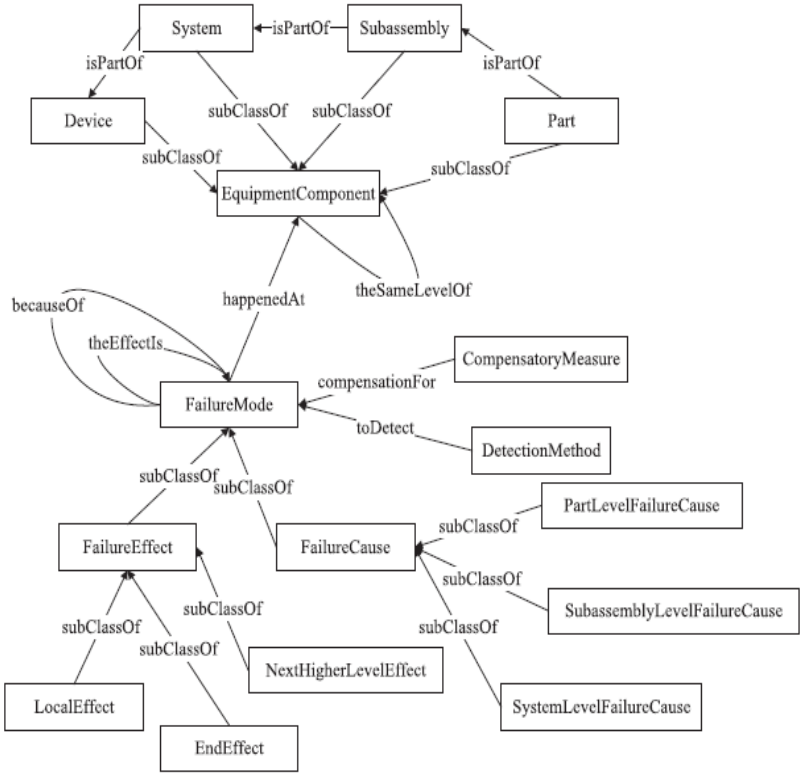
Taxonomies and ontologies



1



1



2

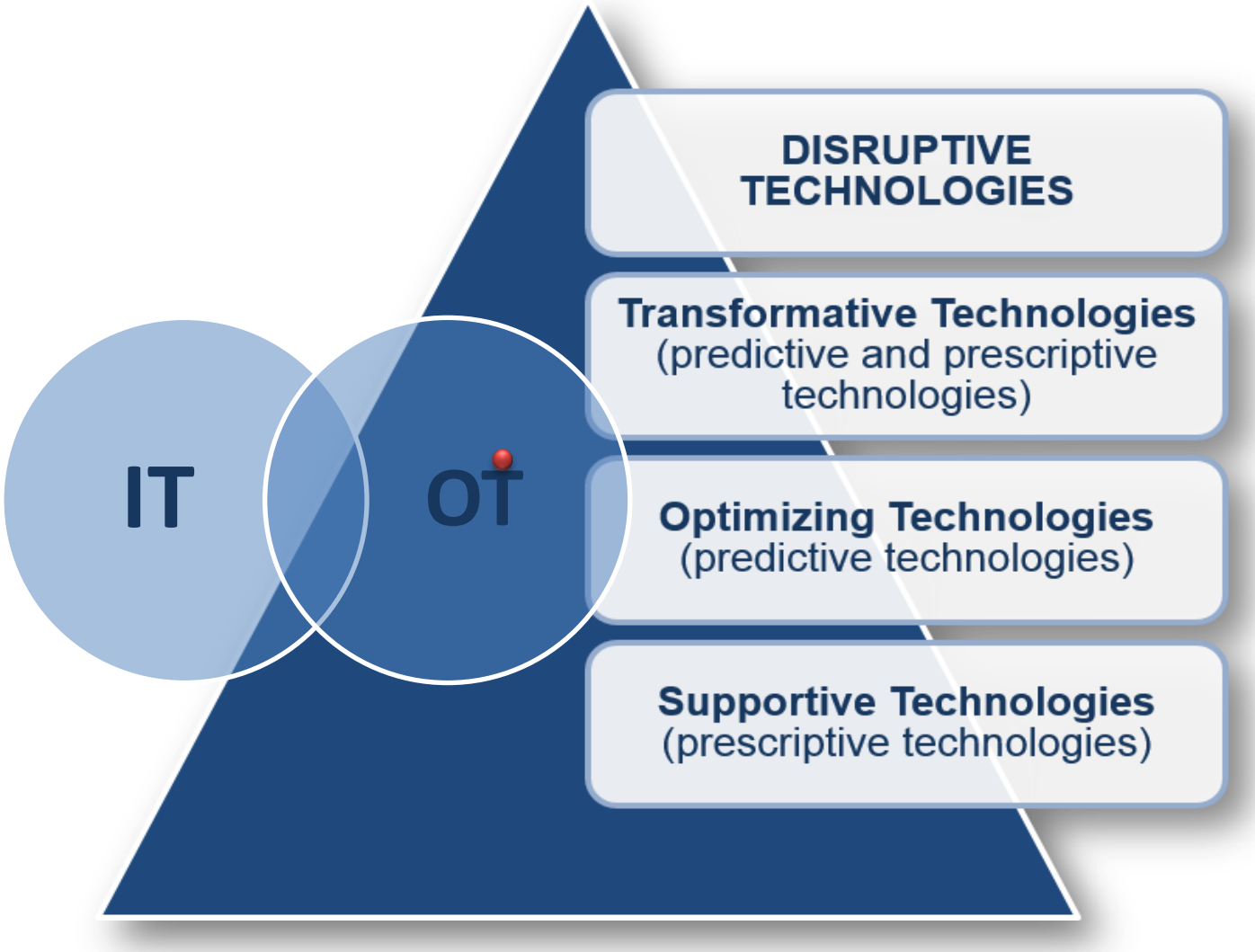
Rule-1

FailureMode(?x) ^ hasHappened(?x, true) ^ Device(?y) ^ happenedAt(?x, ?y) ^ FailureMode(?z) ^ theEndFffectIs(?z, ?x) ^ FailureMode(?a) ^ theHighEffectIs(?z, ?a)?theDirectFailureCauseIs(?x, ?a) ^ hasHappened(?a, true)

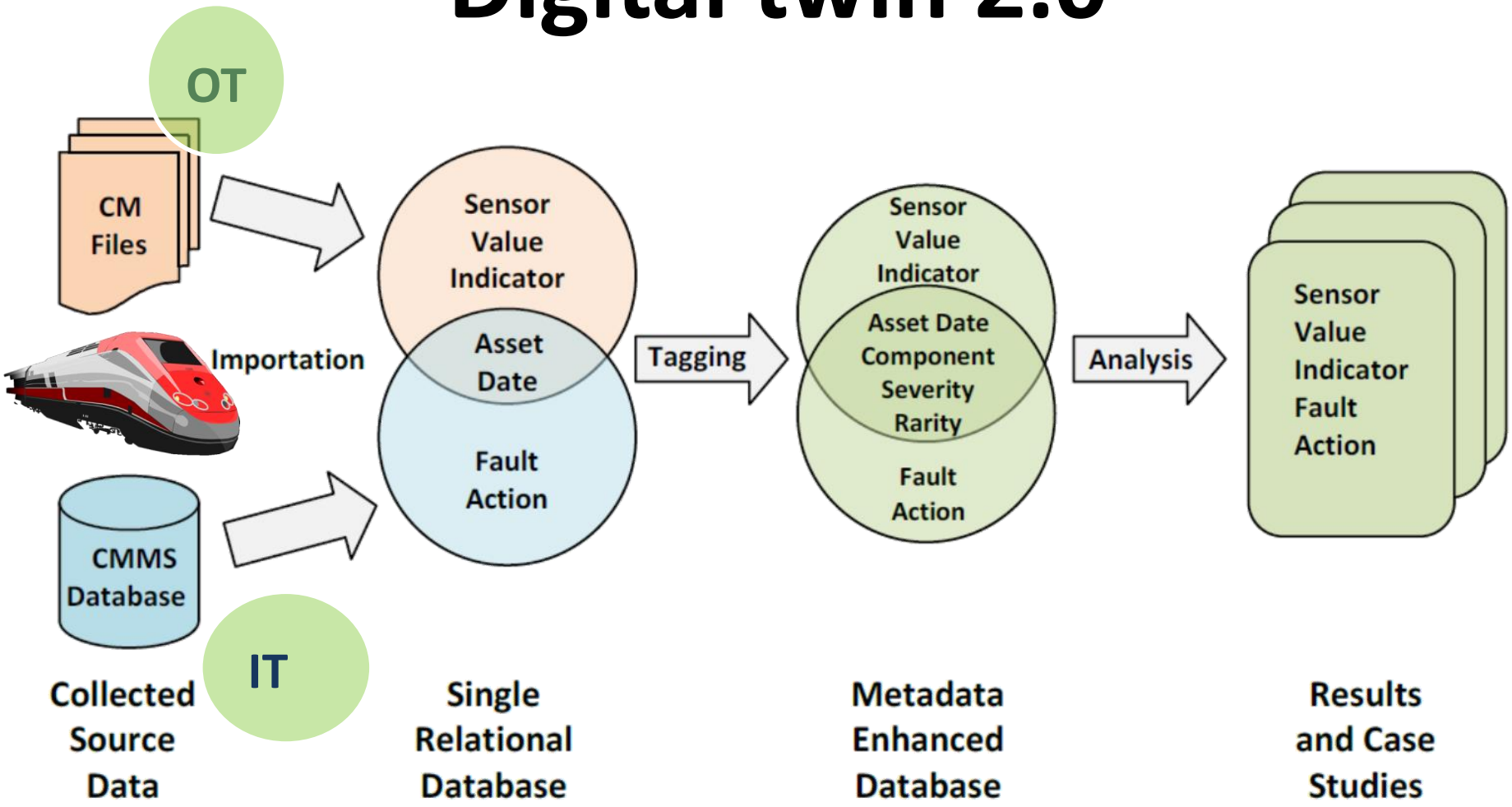
2

TRANSFORMATIVE MAINTENANCE SOLUTIONS

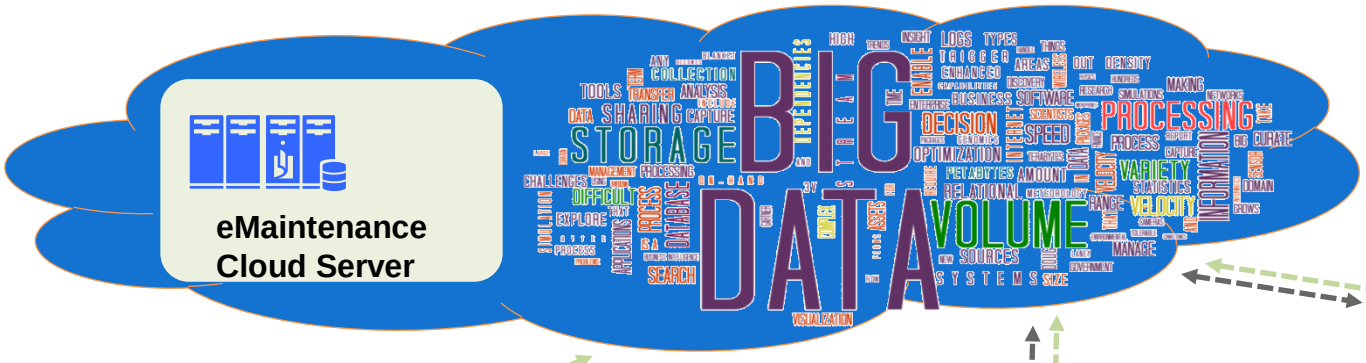
Integration & Application of Technologies



Digital twin 2.0

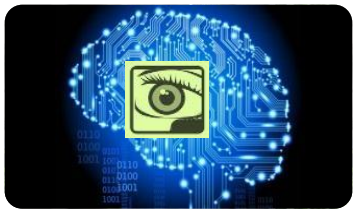


Digital twin 2.0



Technical services

On board Wireless System



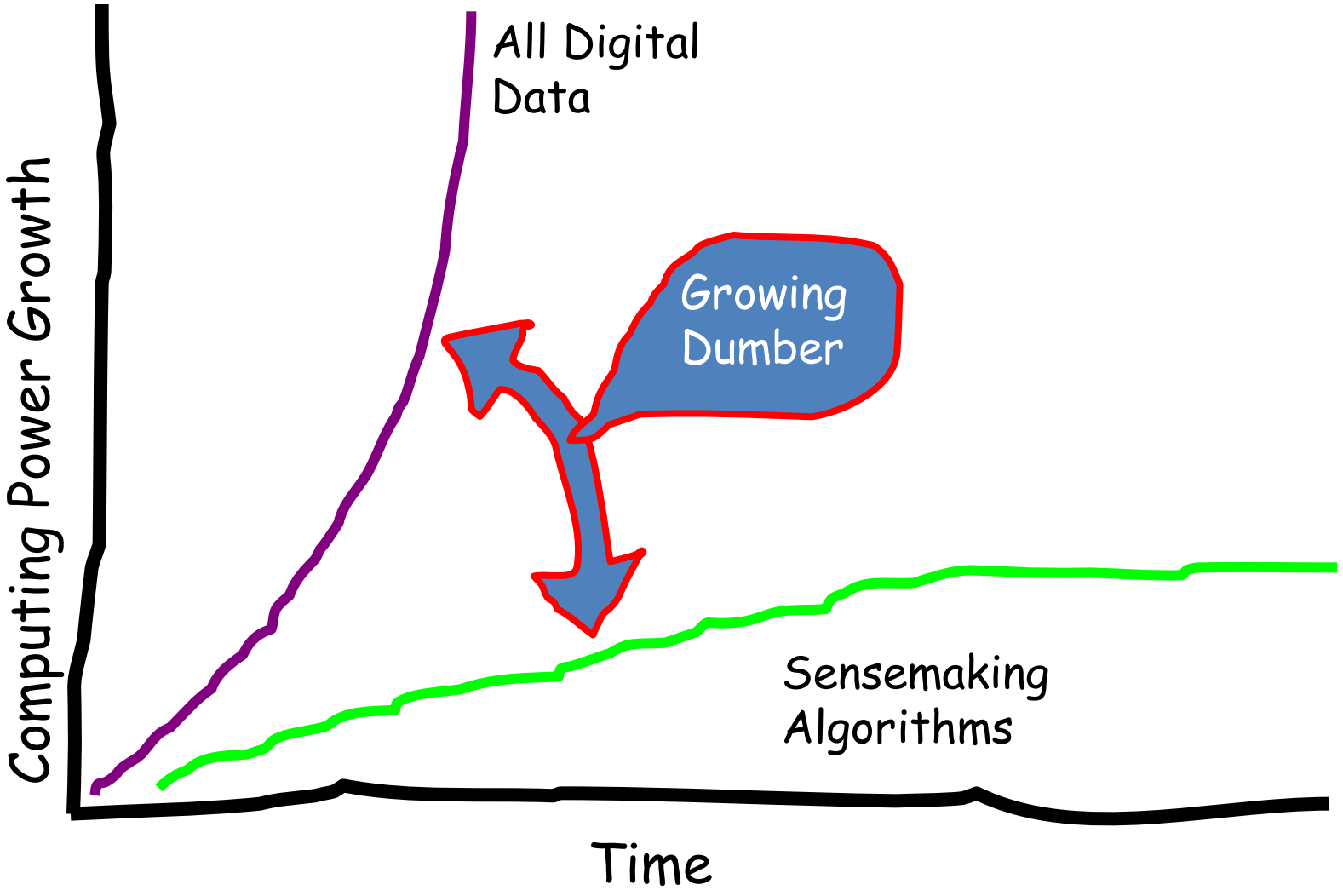
Machine Maintenance Analytics

Truck scheduling

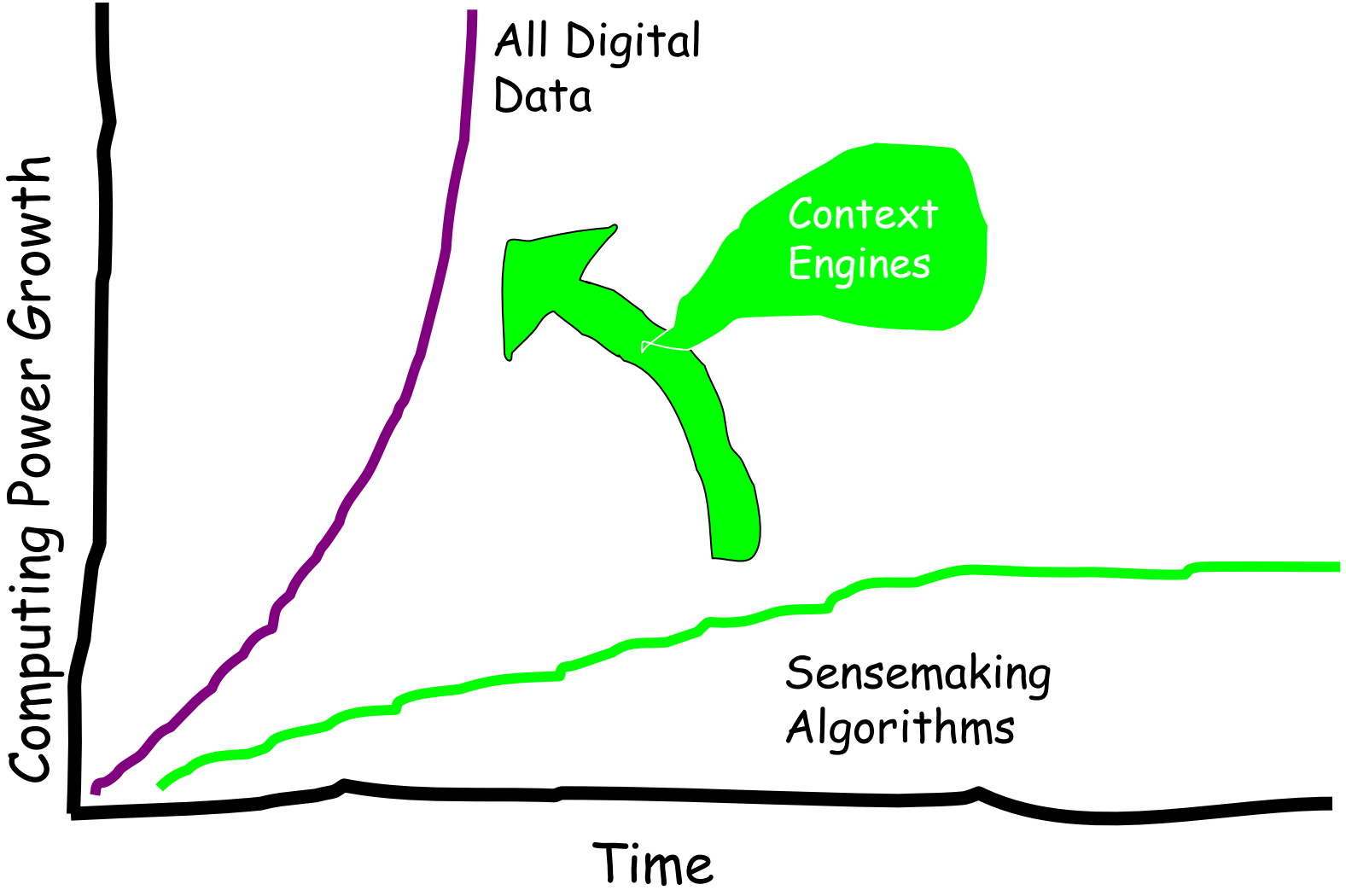
Data → Information → Knowledge

The need for sensemaking Maintenance Analytics

Trend: Organizations are Getting Dumber



The Way Forward

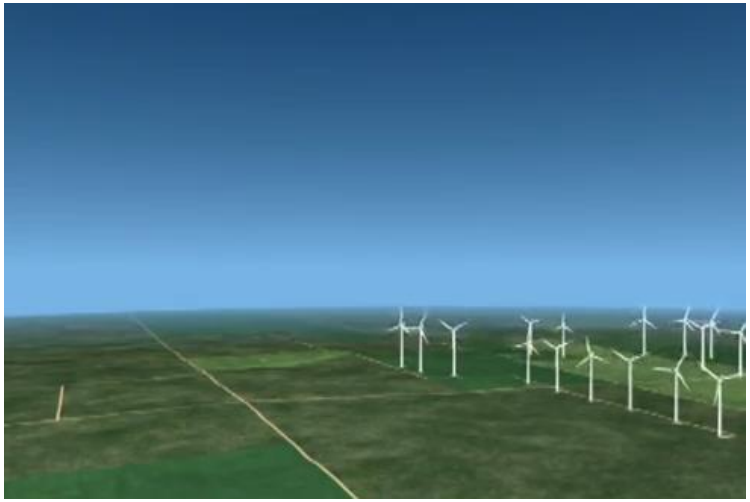


What is context?

"Any information that can be used to characterize the **situation of entities that are considered relevant** to the interaction between a user and an application"

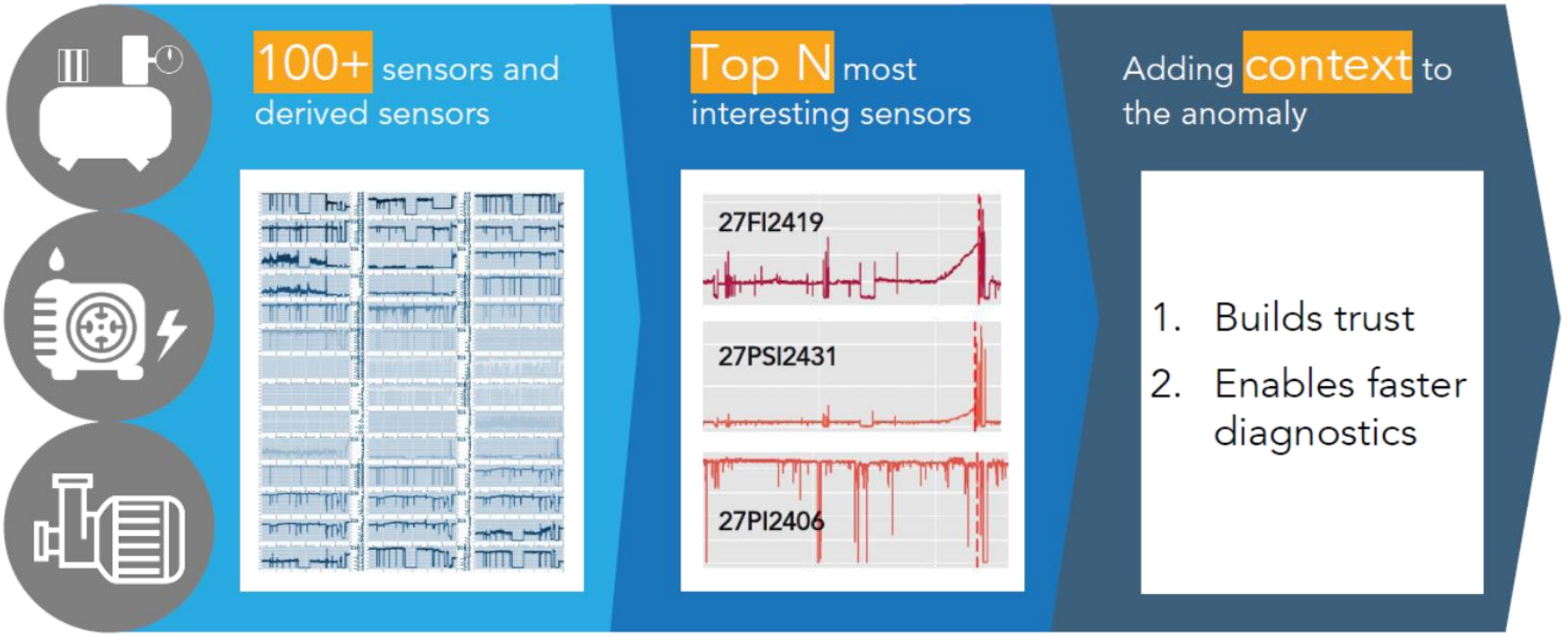
Dey et al.

"A pattern of behavior or relations among variables that are outside of the subjects of design manipulation and potentially affect user behavior and system performance"



Sato

Even Smarter: Machine learning diagnostics



Context-aware Maintenance Decision Support Solution

Digital twin 2.1

Condition Data

Maintenance Data

Business Data

Information models

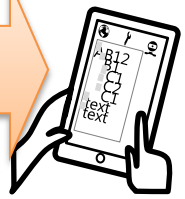
Knowledge models

Context models

Data Fusion & Integration

Big Data Modelling & Analysis

Context sensing & adaptation



From data to information, Train/Ground communication

Train events

Train status

Train position

What can I see in my data?

Now casting

1) What has happened

2) What is happening

Forecasting

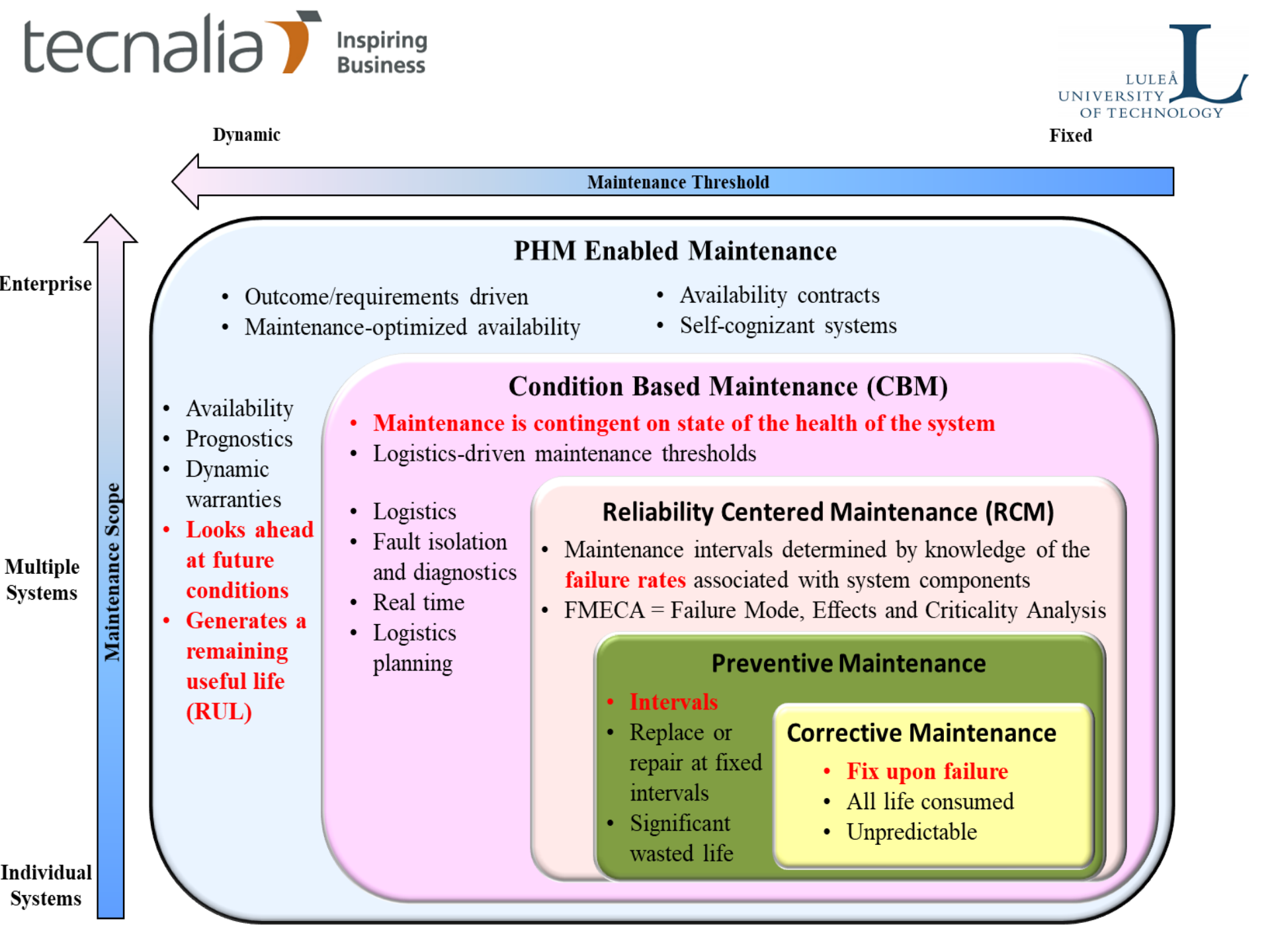
3) What will happen in the future

4) When will it happen

Health index calculation

Remaining useful life prediction

Maintenance, when needed


Dynamic
Fixed
Maintenance Threshold
Maintenance Scope
Enterprise
Multiple Systems
Individual Systems
PHM Enabled Maintenance
• Outcome/requirements driven
• Maintenance-optimized availability
• Availability contracts
• Self-cognizant systems
• Availability
• Prognostics
• Dynamic warranties
• Looks ahead at future conditions
• Generates a remaining useful life (RUL)
Condition Based Maintenance (CBM)
• Maintenance is contingent on state of the health of the system
• Logistics-driven maintenance thresholds
• Logistics
• Fault isolation and diagnostics
• Real time
• Logistics planning
Reliability Centered Maintenance (RCM)
• Maintenance intervals determined by knowledge of the failure rates associated with system components
• FMECA = Failure Mode, Effects and Criticality Analysis
Preventive Maintenance
• Intervals
• Replace or repair at fixed intervals
• Significant wasted life
Corrective Maintenance
• Fix upon failure
• All life consumed
• Unpredictable

DETECTION, ISOLATION & PROGNOSIS

Detection

Through sensors, Models etc

Isolation

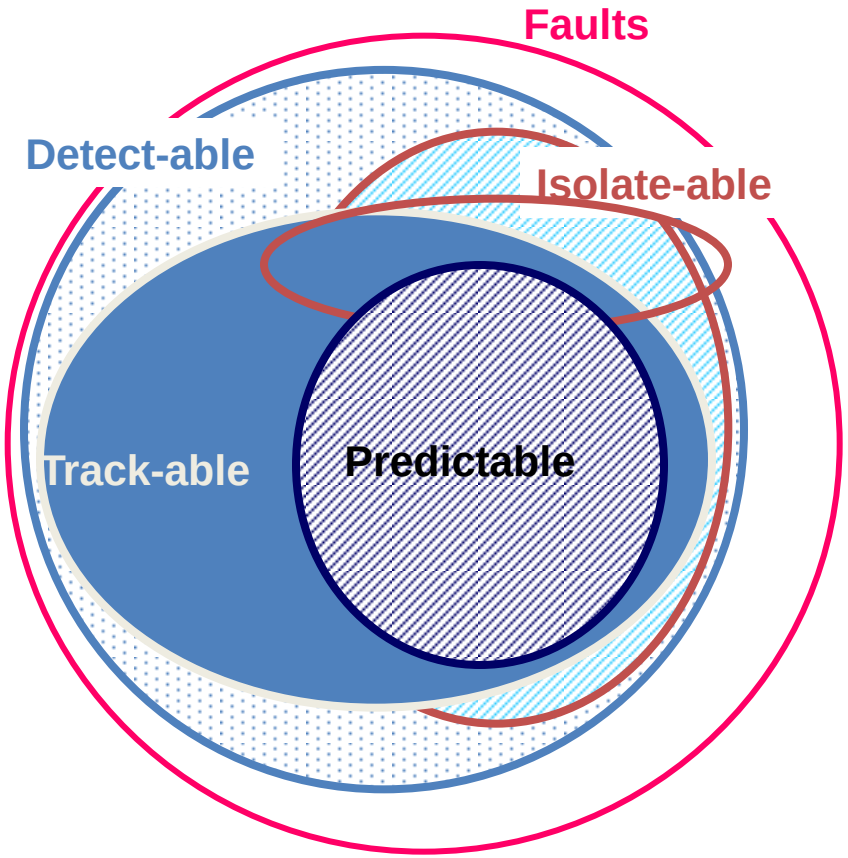
Information fusion from sensors, Models etc.

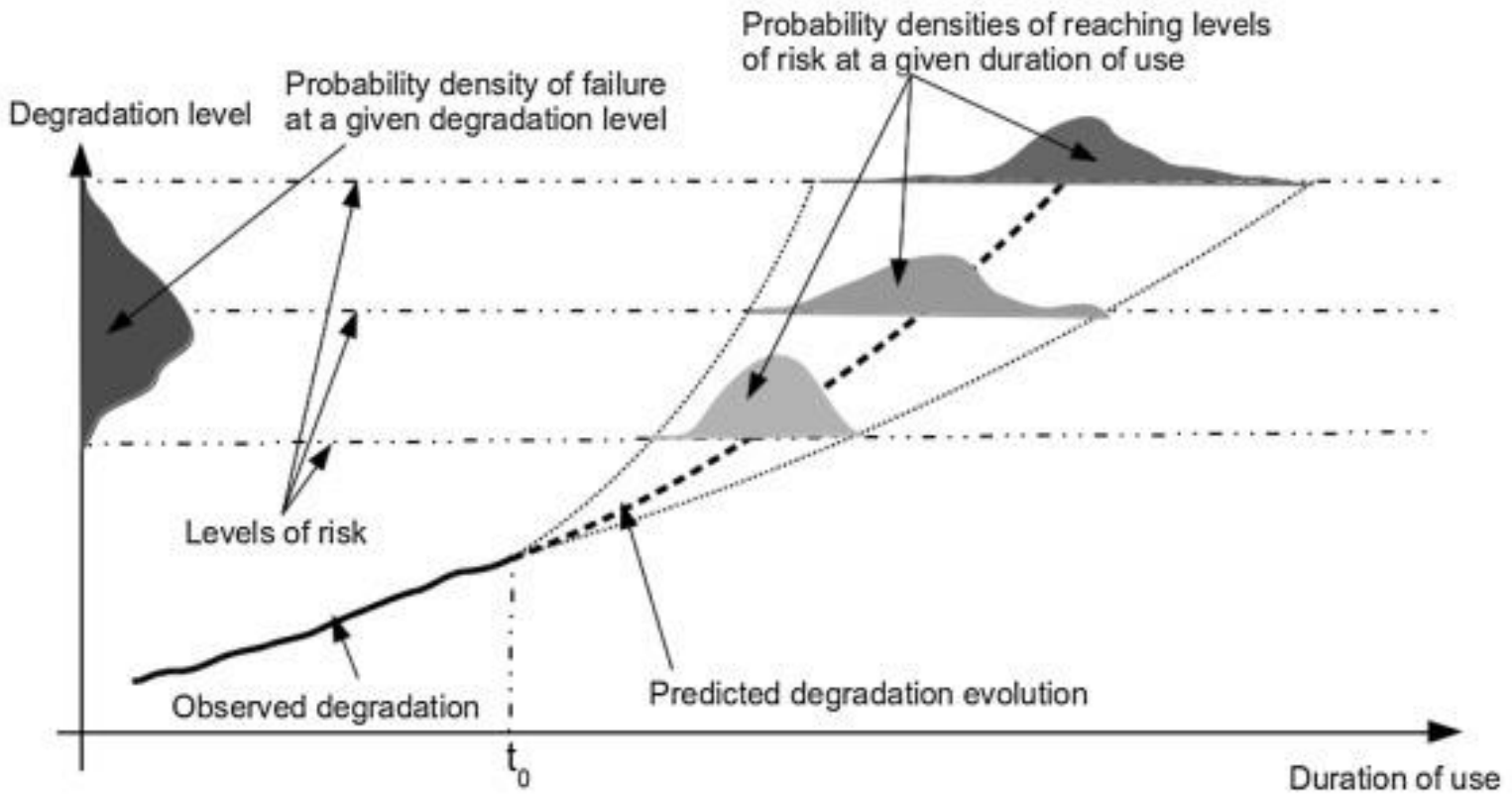
Tracking/Trending

Processed PHM data

Prediction/Prognosis

Based on tracking/trending, & lifing models





1. In the absence of direct "stressors, loading meters" how can we infer the best (descriptors/features) to capture future damage dimensions?
2. How can we accurately predict the progression of a specific failure mode? Considering that multiple failure modes may occur at any time in a complex equipment, system?
3. Given the numerous sources of uncertainty, how do we assign confidence associated with the predictions?

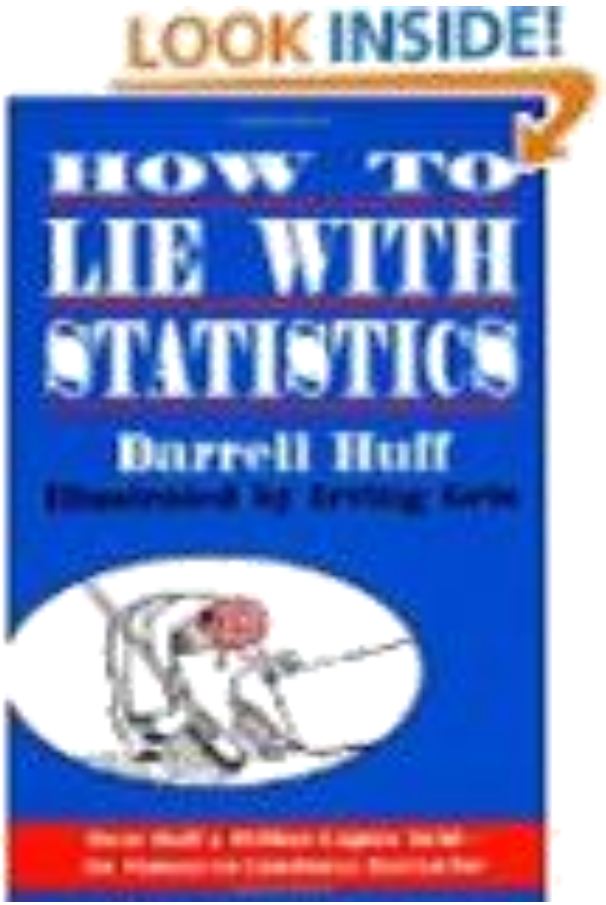
Data science… Narrow vision and mistakes

tecnalia Inspiring Business

LULEÅ UNIVERSITY OF TECHNOLOGY

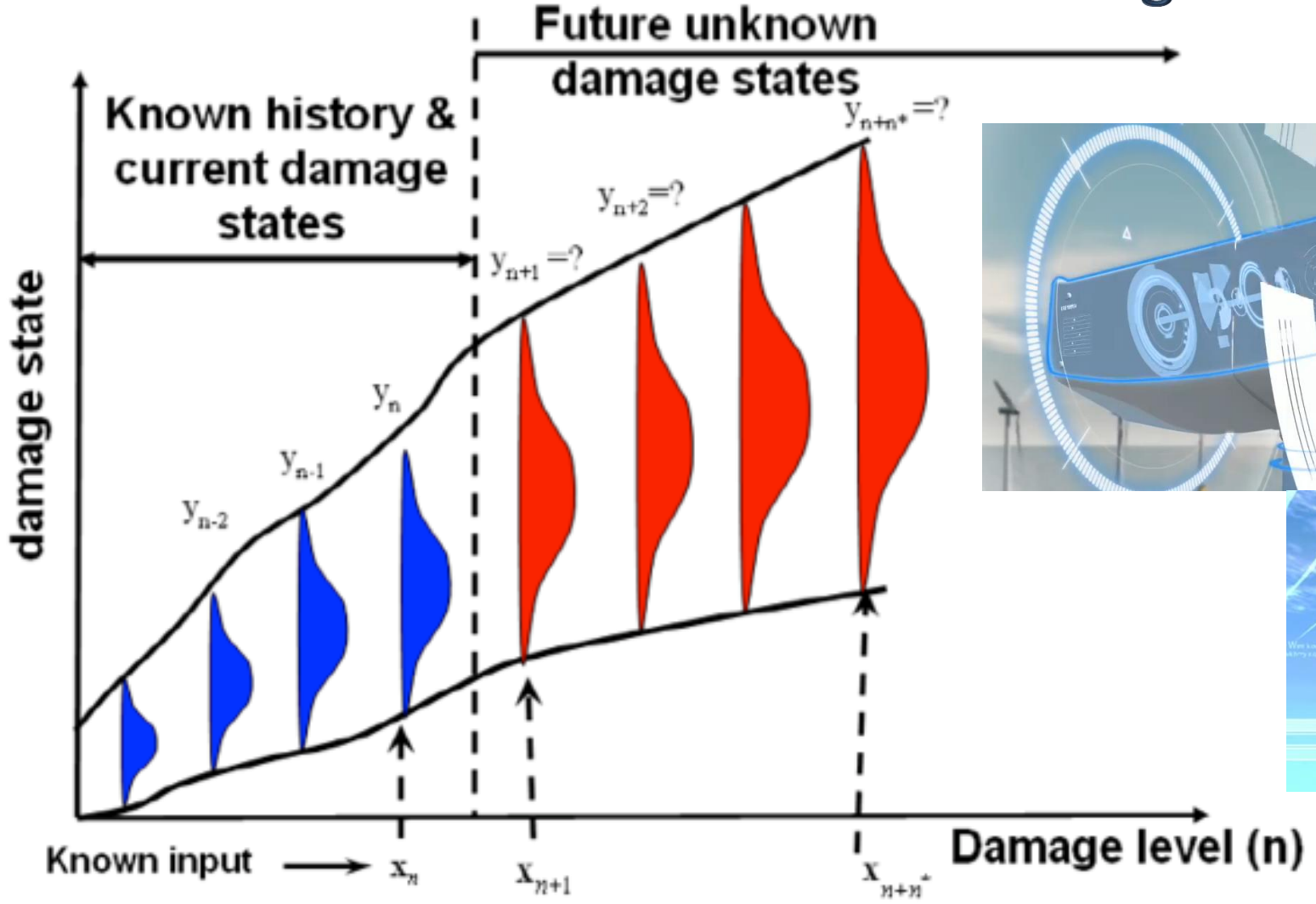
Let us be careful bigger = smarter?

- tolerate errors?
- **discover the long tail and corner cases?**
- more data, more error (e.g., semantic heterogeneity)
- still need humans to ask right questions, lack of analytics

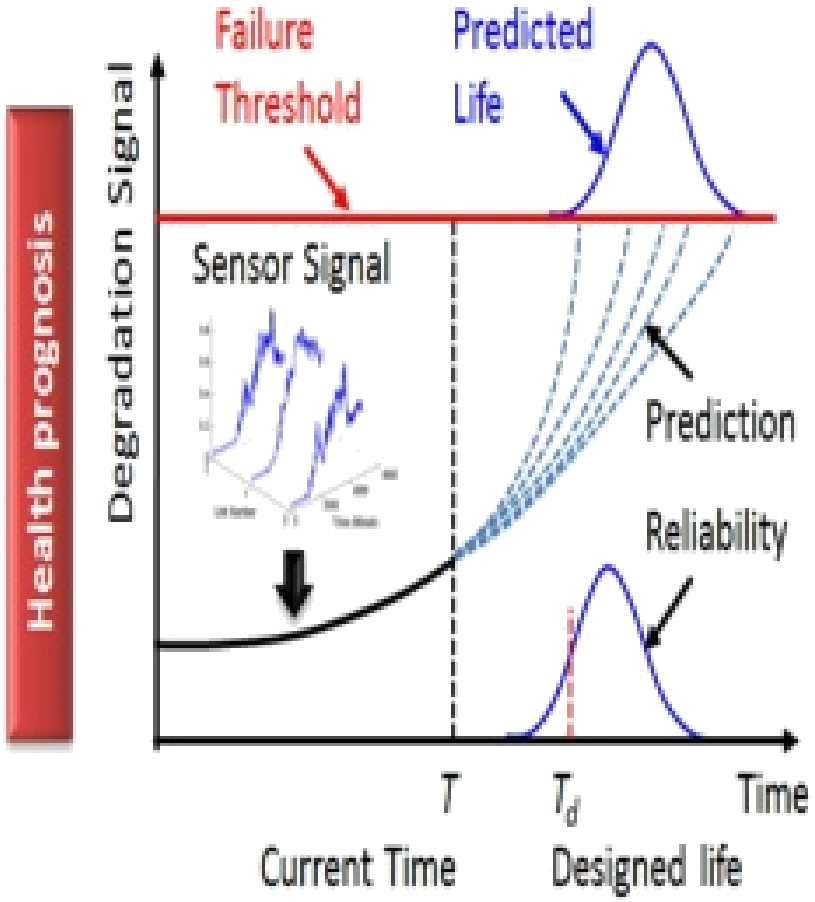


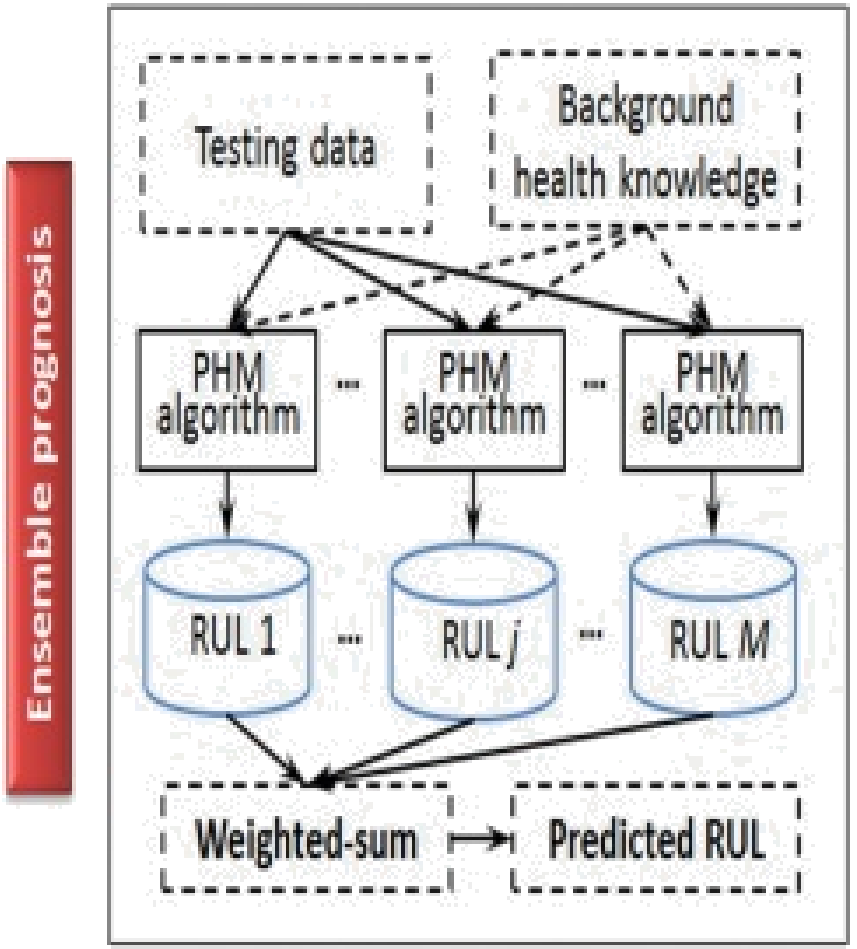
Prognosis when ignorant…..?

Remember the unknown stages



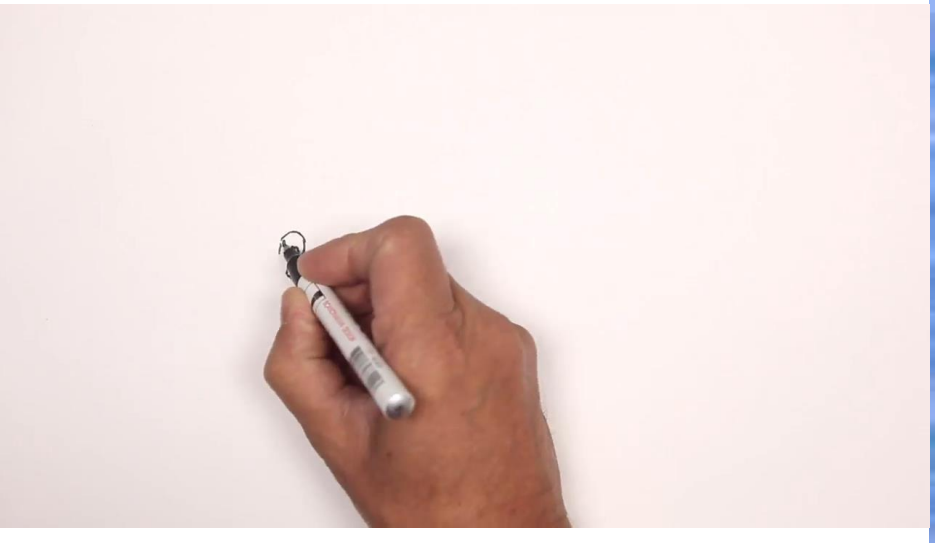
And the Uncertainty in RUL minimized with physics, maximized with data





Black Swan Losses

- Loss Distribution
  - Tail events are rare – very little data
  - Typically strong model assumptions

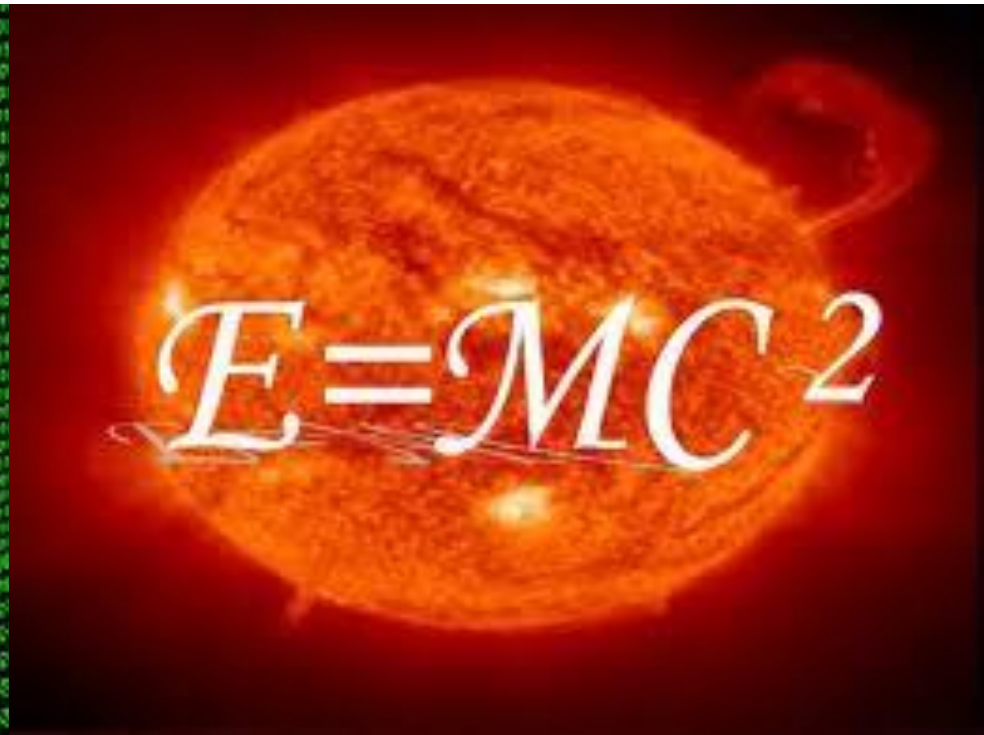


tecnalia Inspiring Business

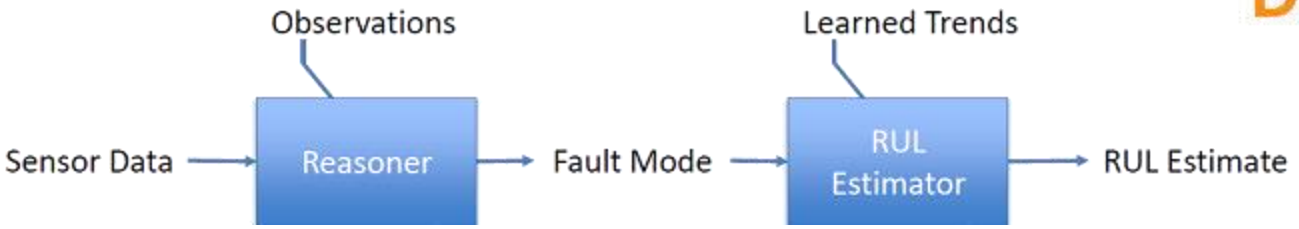
LULEÅ UNIVERSITY OF TECHNOLOGY

Data driven or model based?

Data-Based or Physics-Based Models? – That is the question!



Data driven methods

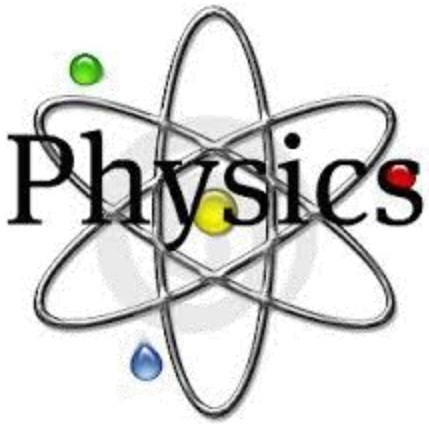


- Fit mathematical model to observations (trending)
  - No guaranty that extrapolation will be meaningful
- Collect statistics of failures as a function of current state
  - Requires volumes of data and is difficult to know when you have enough

tecnalia Inspiring Business

Physical based methods

- Physics of Failure Model Driven
  - Capture physical basis of failure in model that relates the forces that cause damage to their effect
    - Requires a detailed understanding of the problem
- Many Implementations Are a Combination of Both



Hybrid models

- Combine knowledge about the physical process and information from sensor readings to enhance prognostics capabilities.
- Integration of measured data and physics can lead to a reduction of uncertainty (e.g. adjust predictions from model using observed data).
- Integration can be implemented at different levels of the PHM process:
  - Online model parameters updating.
  - Model predictions correction based on observed data.
  - Measure current damage level and propagate.
  - Build empirical degradation models from data.

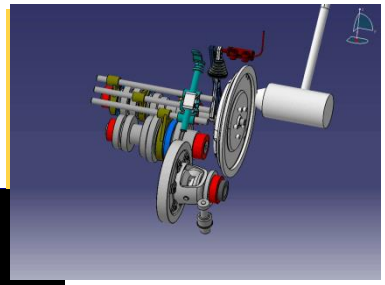
Evolution of the Process

Technological Advance

Design & Validation of products

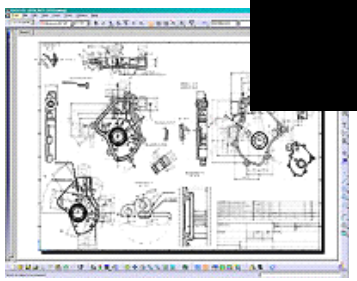
Knowledge Capture

Digital Mockup



Digital twin

2D



3D

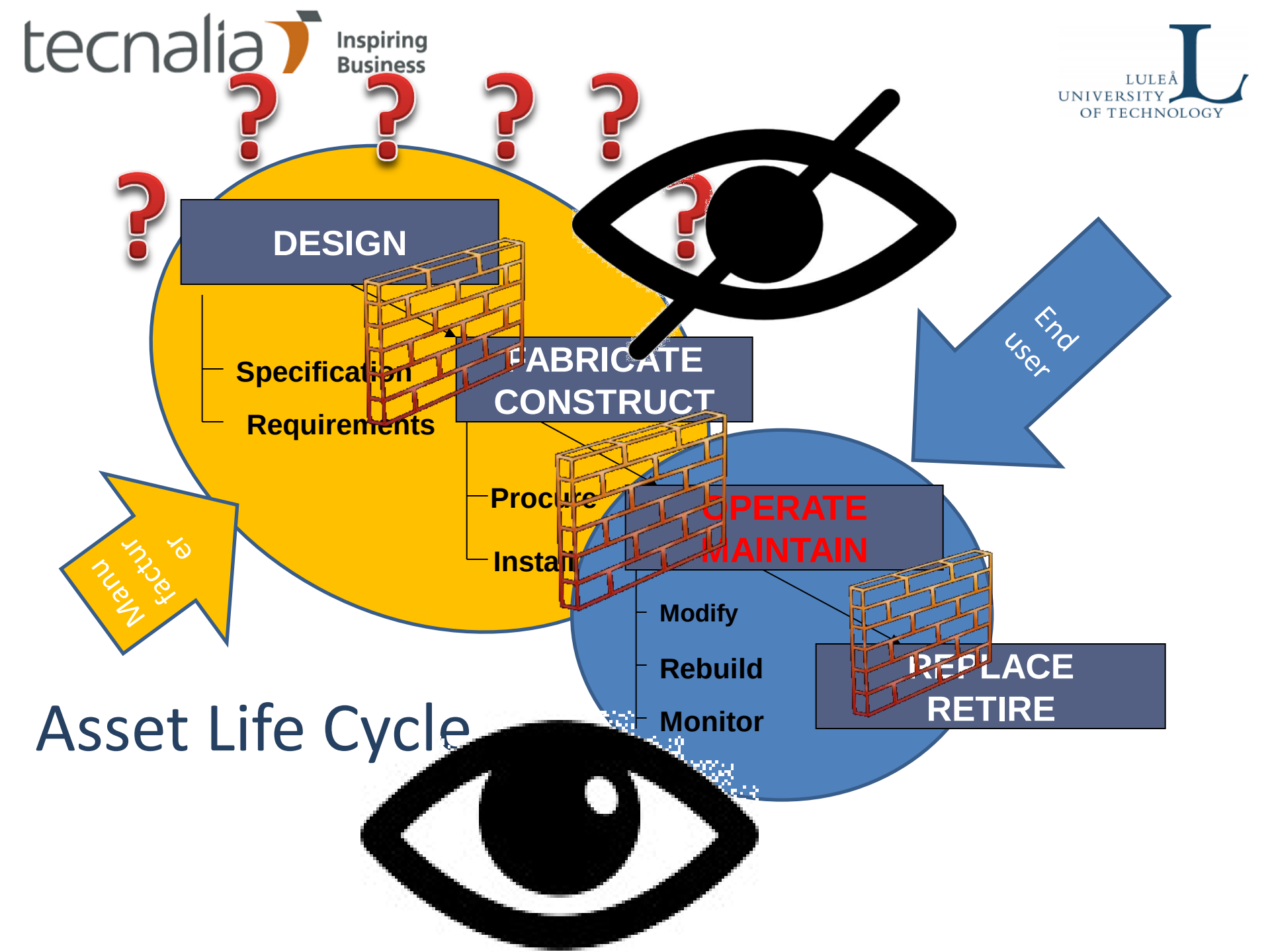
Integration of Product Design and O&M

80s

90s

2000

2016.....


tecnalia
Inspiring Business
LULEÅ UNIVERSITY OF TECHNOLOGY
DESIGN
Specification
Requirements
FABRICATE CONSTRUCT
Procure
Install
OPERATE MAINTAIN
Modify
Rebuild
Monitor
REPLACE RETIRE
End user
Manufacturer
Asset Life Cycle

PLM and digital twins

Product Lifecycle Processes

Design → Build → Operate → Maintain

O&M Processes

Digital Wind Solutions

Create Processes → Plan Processes → Simulate & Validate Processes → Execute Processes

Engineering Design ← Validate ← As – Built Records

Hybrid & Context Driven Services

Physics of failure based

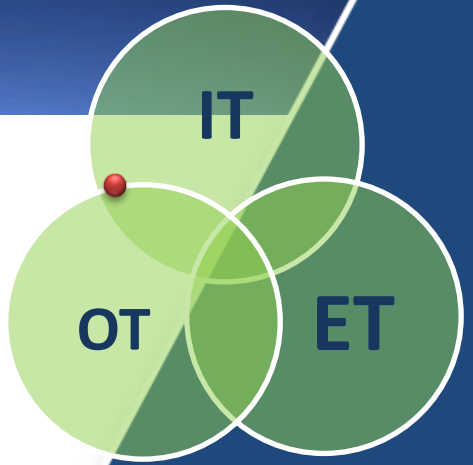
Hybrid models

Data driven

Context Driven Services

Context Awareness

Digital twin 3.0



DISRUPTIVE TECHNOLOGIES

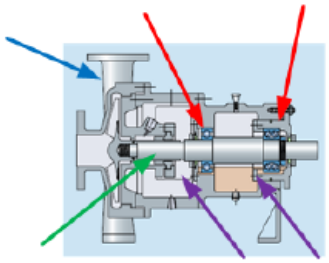
Transformative Technologies
(predictive and prescriptive technologies)

Optimizing Technologies
(predictive technologies)

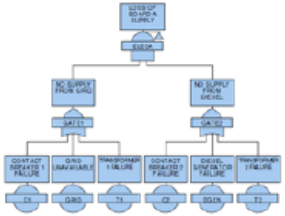
Supportive Technologies
(prescriptive technologies)

The process of twin 3.0 building

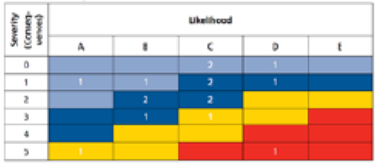
The asset (machine, equipment, electronics, system, structure, etc.



FMECA identifying monitored failure modes and parts taxonomy



Defining taxonomy of parts within the asset



Articulation of Failure Physics

Methodology

Rolling stock

HVAC

Identify failure modes of Equipment units, Subunits, Components and Parts based on the Taxonomy

Analyze failure effects / causes

Classify failure effects by severity

Perform criticality calculations

Rank failure mode criticality

Determine critical items

Identify means of failure detection, isolation and compensating provisions

Document the analysis. Summarize uncorrectable design areas, identify special controls necessary to mitigate risk.

Make recommendations

Follow up on corrective action implementation / effectiveness

Taxonomy

FMECA Analysis (failure mode identification)

Relate available variable with failure modes

Identification of possible modeling methods for failure identification

Critical system, sub-system, component identification

Diganosis/Progonosis

Condition Monitoring

Maintenance Planning

Identification of the **system** for the analysis

Survey for the functionality of the system

Identification of the Equipment units, Subunits, Components and Parts

Complete Taxonomy creation

Identify the measurements for the detection of failures

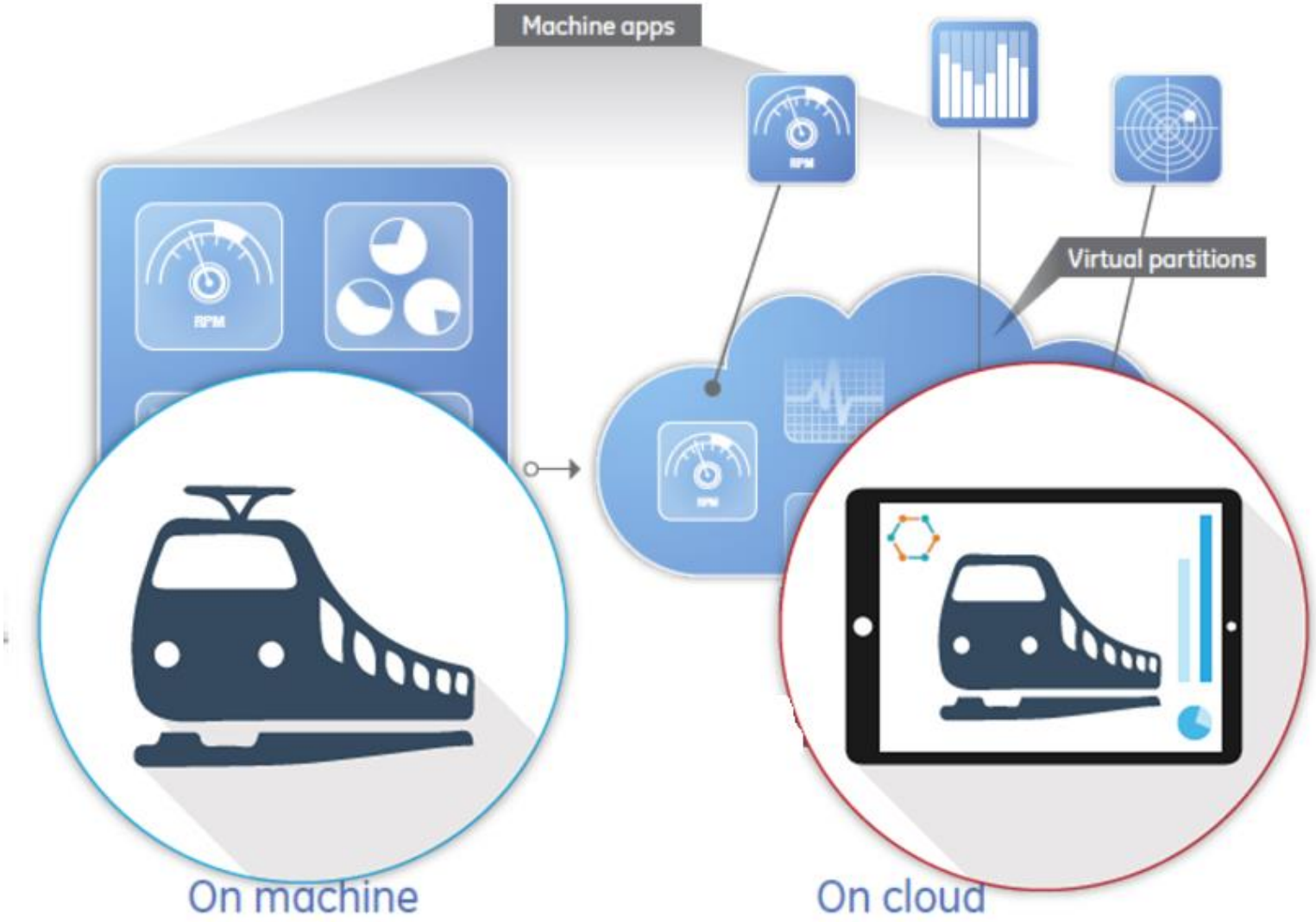
Identify the available variables and parameters in the system

Exploration of possible data extraction form variables

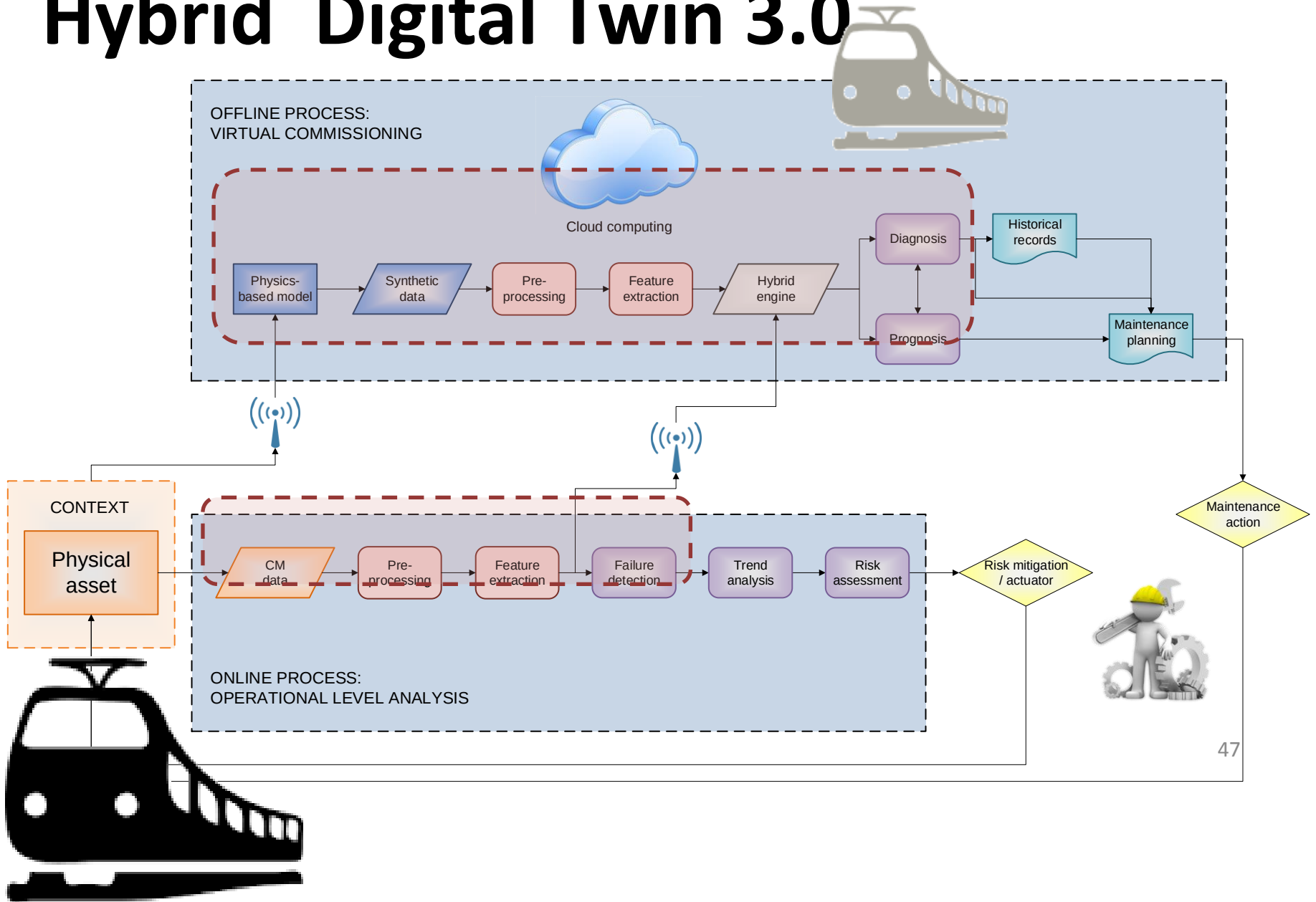
Sensitivity analysis for the detection of failure modes from available variables

Analysis of indirect monitoring of failure modes through variables

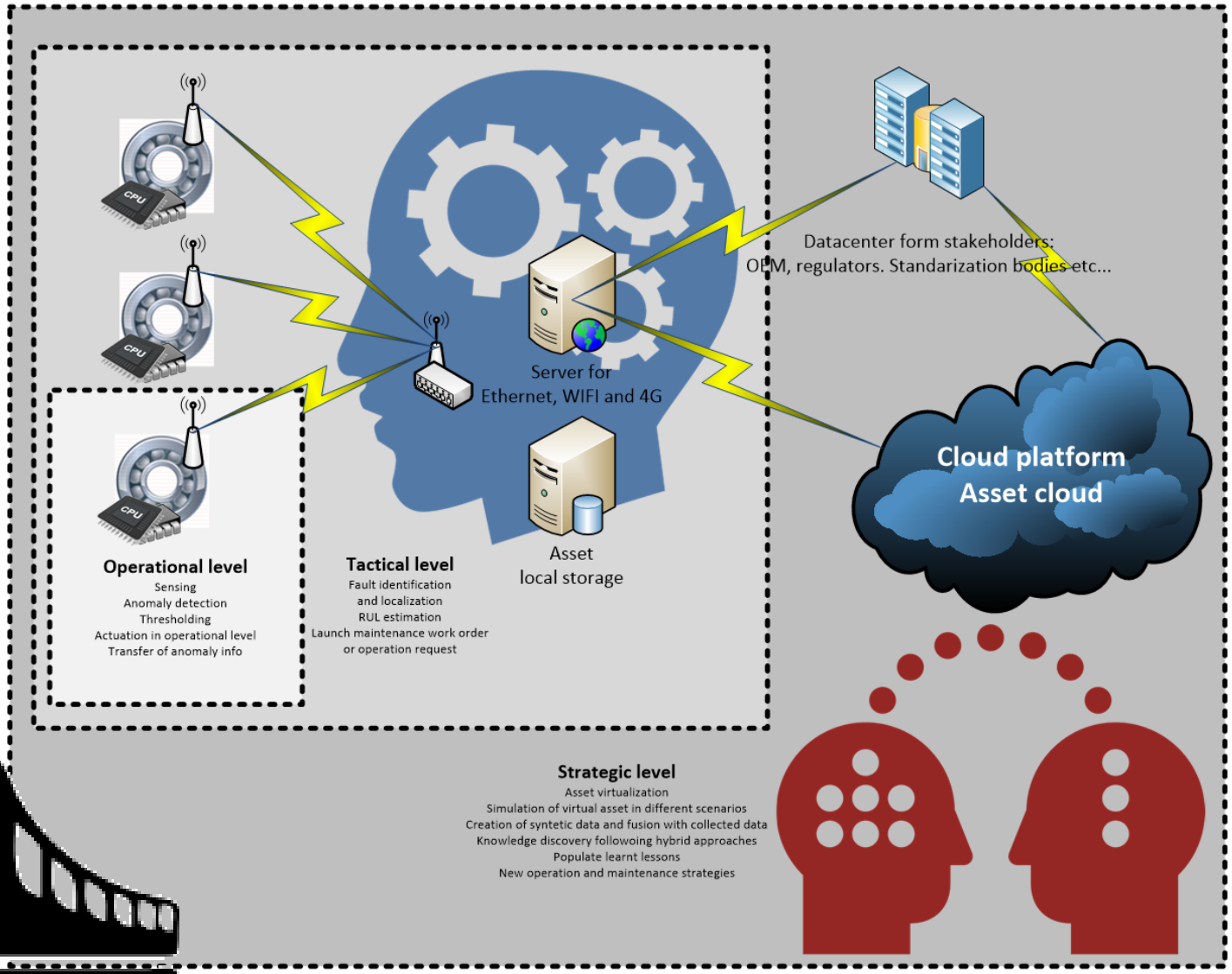
Virtual railway assets

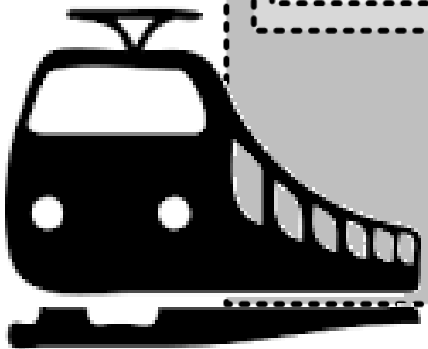


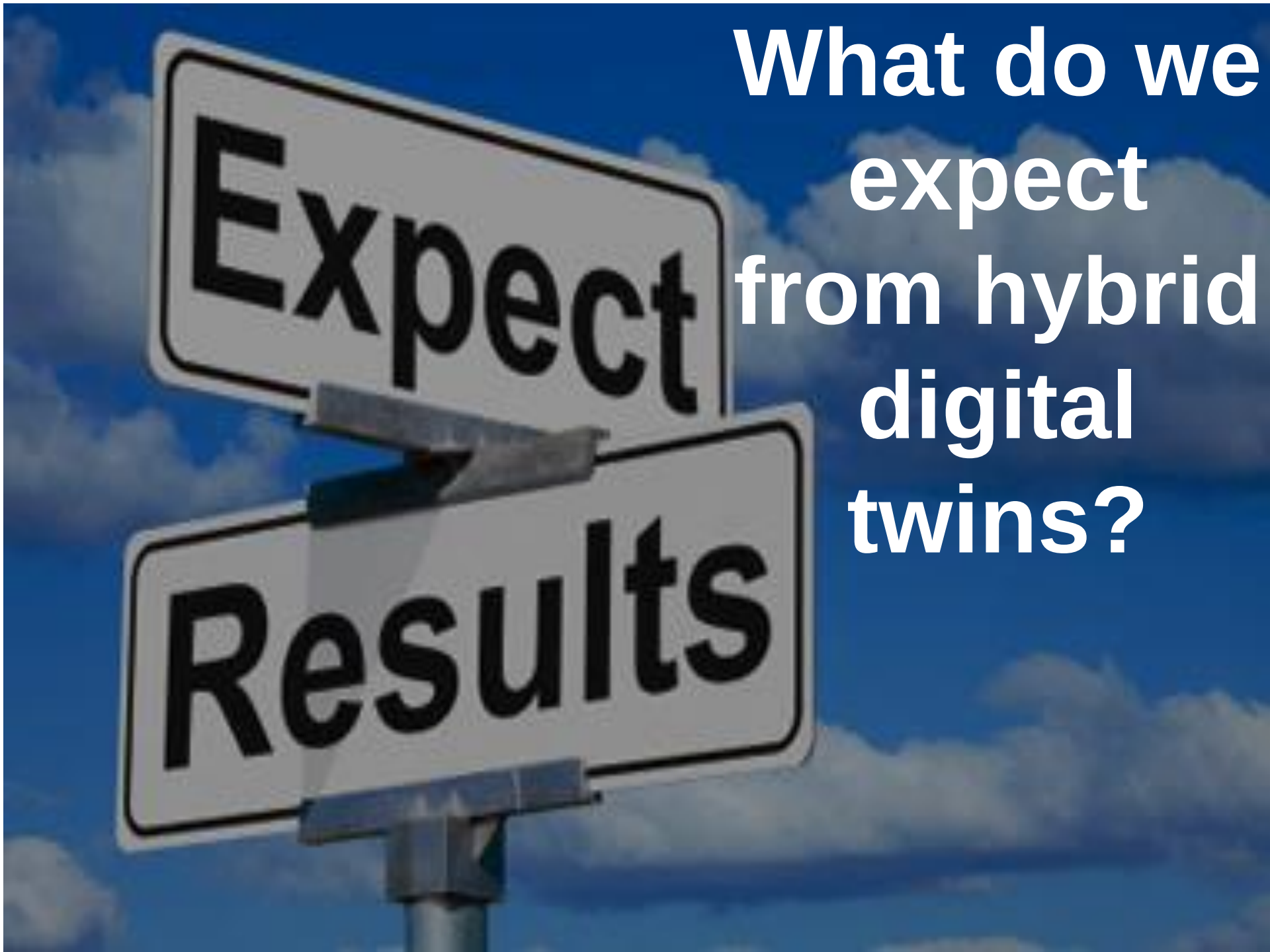
Hybrid Digital Twin 3.0



Are digital twins 1.0 and 2.0 useless????



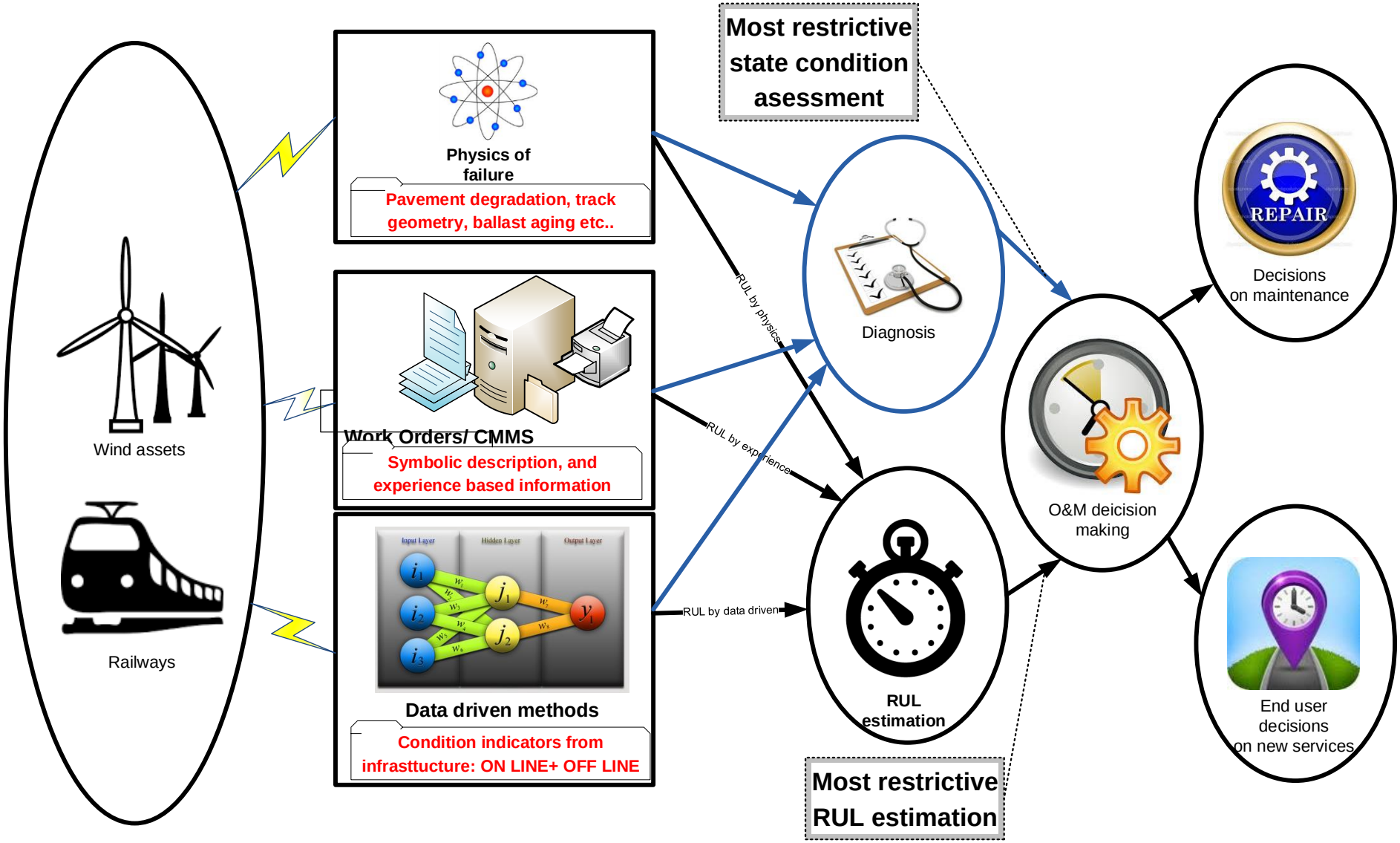




What do we expect from hybrid digital twins?

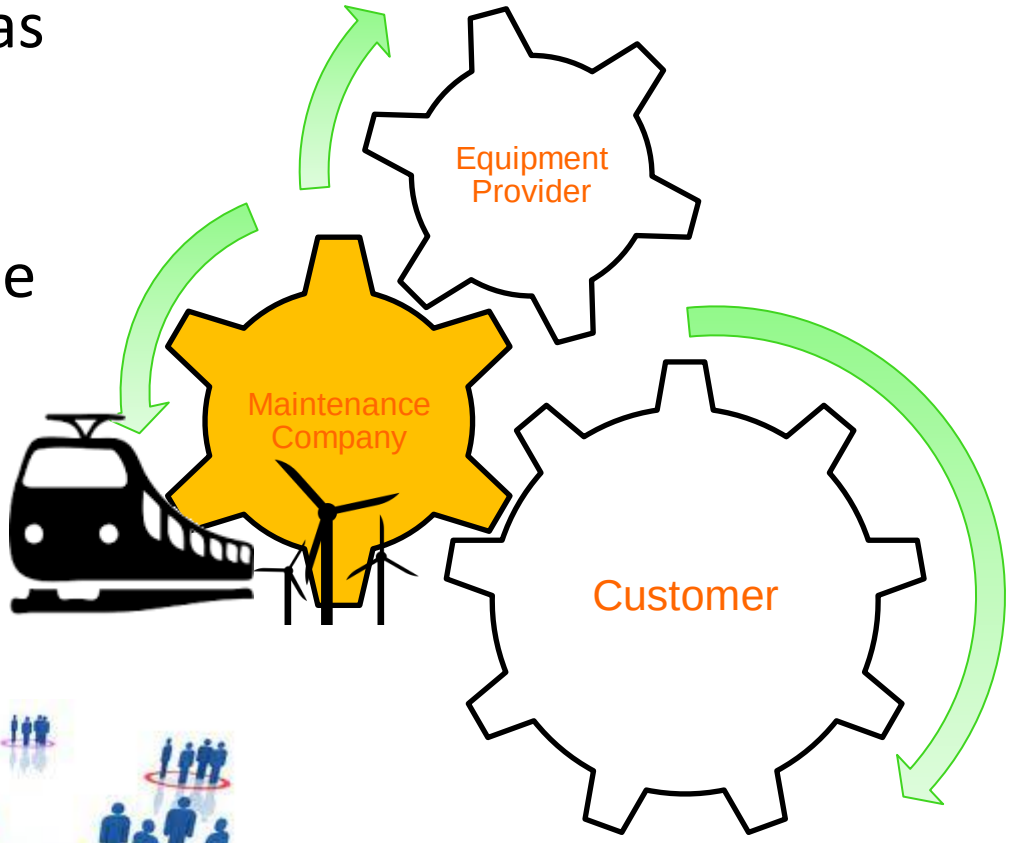
Application of railway virtual commisioning

Diagnosis and Prognosis, enablers for new business models

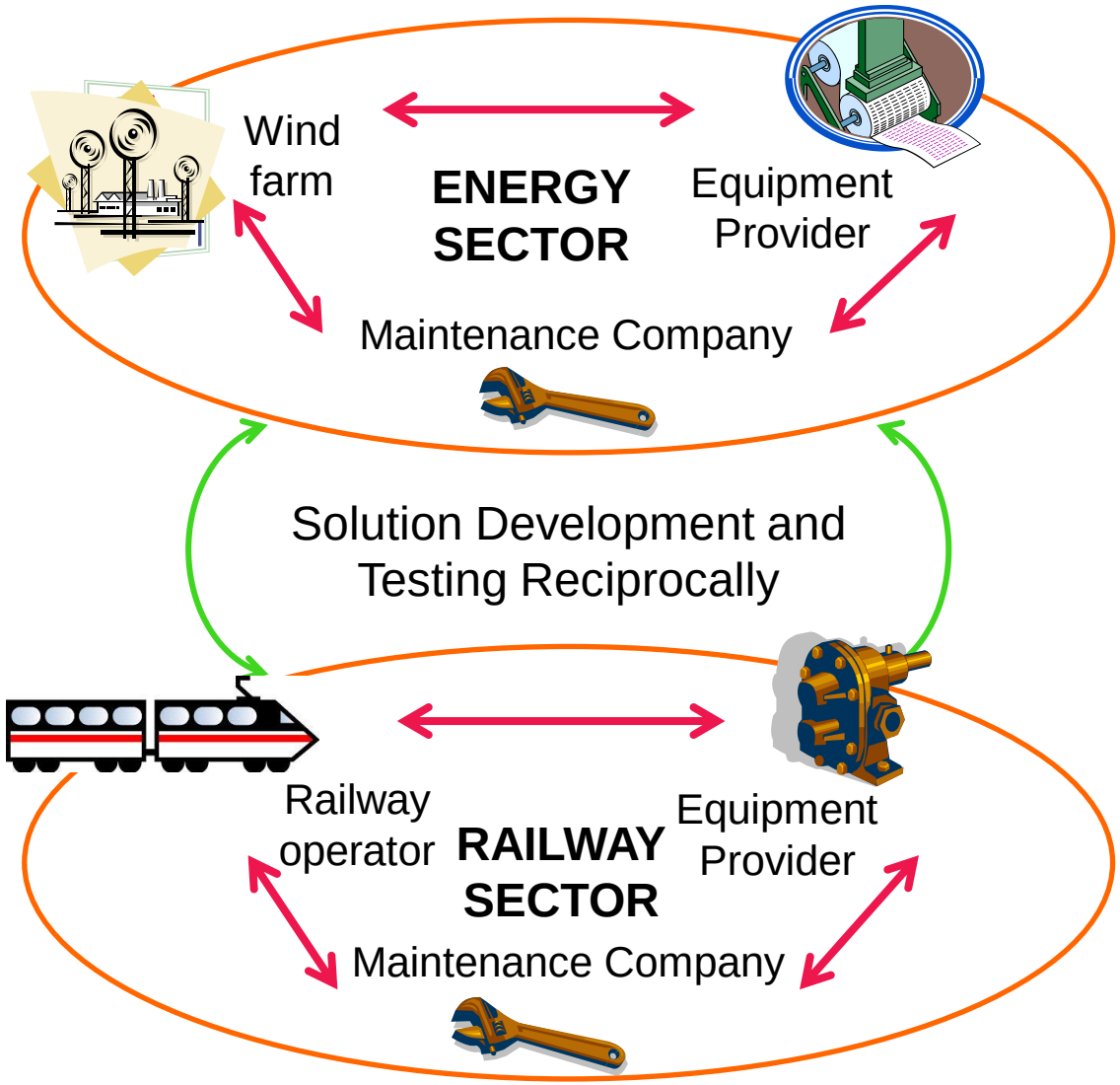


The Renewing Business Network

The **increased outsourcing** of maintenance services has created a need to make both the value and profitability of maintenance **more transparent** at the business network level.



Maintenance as a service



The additional value of the maintenance services achieved by the separate members of the business network must be **measured**. Thus, the **benefits** can be increased and the **risks** controlled to augment the competitiveness of the whole business network.

Concluding remarks

- **Digital twins and Hybrid models** are needed for virtual commisioning to deliver O&M services
- O&M based on Data driven solutions can lead to **catastrophic failures**
- **Life extension is not possible** with big data analytics
- **Manufacturers must provide the integration of systems and data**
- **Digital twin 4.0 will consider evoltionary models and normality dynamics**

tecnalia Inspiring Business

LULEÅ TEKNISKA UNIVERSITET

diego.galar@tecnalia.com

diego.galar@ltu.se