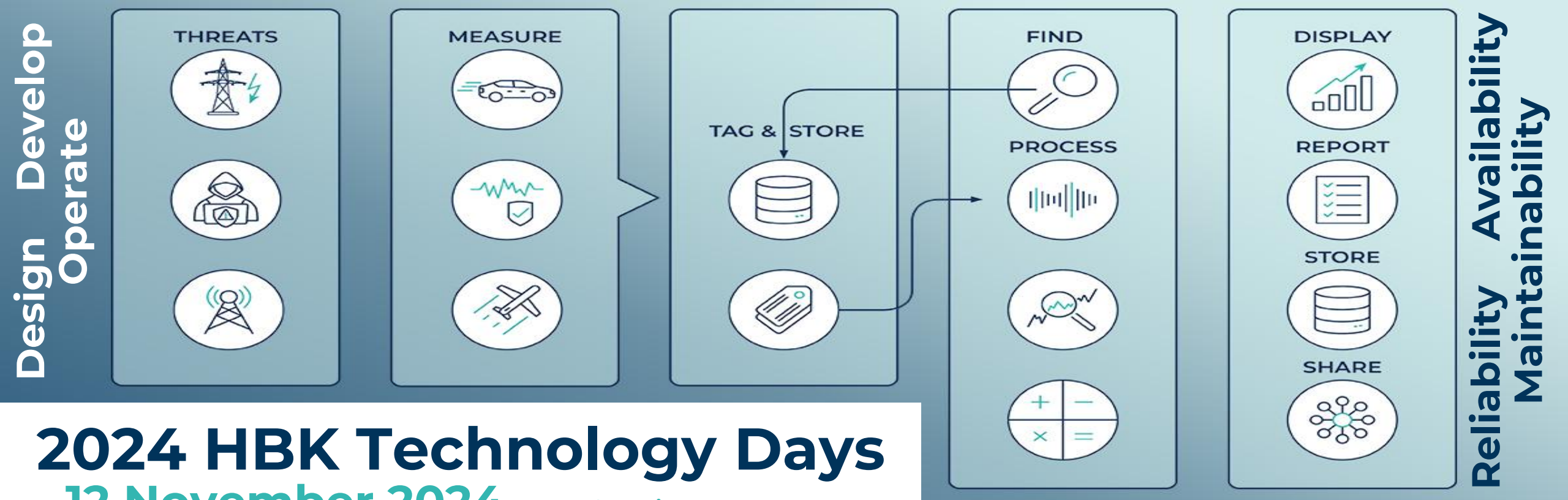
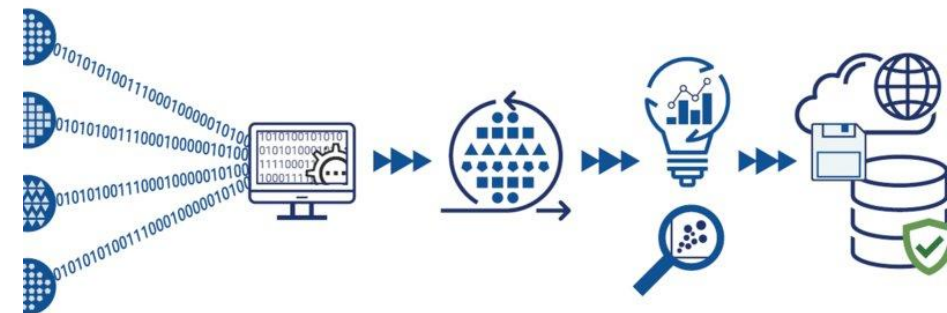




2024 HBK Technology Days

12 November 2024

Session 4: Vehicle Network Analytics



- Extracting insights from streams and data lakes of vehicle network measurement data

Frédéric Kihm,
Product Manager – Analytics and Signal Processing, HBK

Extracting insights from streams and data lakes of vehicle network measurement data

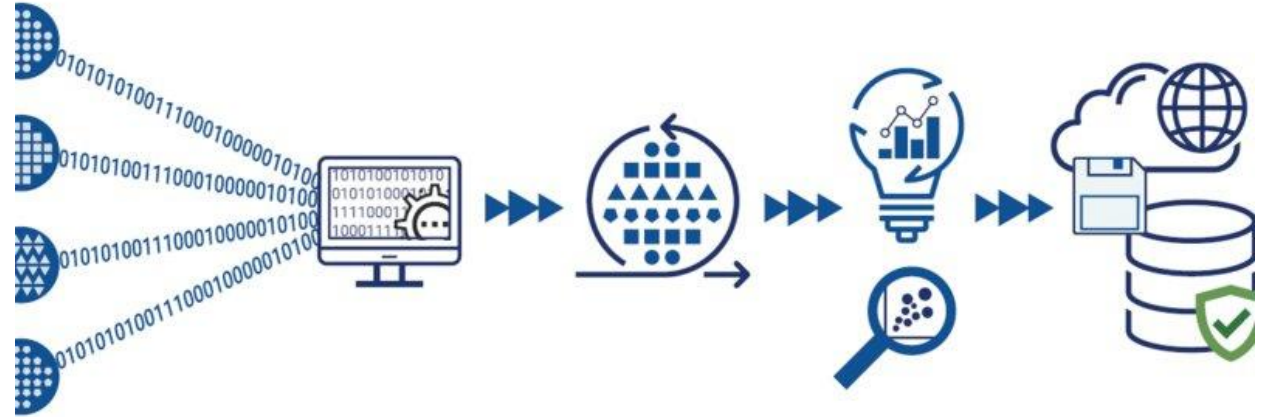
Abstract

- It is increasingly easy to acquire very large quantities of vehicle network measurement data, from data within a single vehicle network, streaming from multiple vehicles in near real time, and/or previously acquired and stored in an expanding data lake. Conversely, these increasing quantities of vehicle data make it more difficult to extract important insights of individual vehicle behaviour and performance, and how they compare with a distribution of many vehicles. This presentation will describe effective ways to:
 - search these data
 - augment these data with calculated tags, indicators and metrics
 - reduce these data to meaningful histograms
 - show trends and statistics
- The objectives being to understand product usage, validate new designs, compare program iterations, identify and extract scenarios for simulation and simulators, create inputs for virtual simulation models, create warranty reliability models, create predictive maintenance models, and more. To be able to retrospectively query the data lake and run analyses on the results to study a newly identified scenario.

Extracting insights from streams and data lakes of vehicle network measurement data

- ▲ About the presenter, Frédéric Kihm,
Product Manager – Analytics and Signal Processing, HBM
 - Frédéric Kihm is Product Manager at HBM, responsible for the signal processing related software products, which includes GlyphWorks, VibeSys and nCodeDS. Frédéric previously worked as an Engineering Consultant for nCode and then HBM, involved with signal processing, durability and vibration analyses in the automotive, aerospace, and defence industries. Frédéric holds a MS in Mechanical Engineering from IFMA University in France and a PhD from the Institute of Sound & Vibration Research (ISVR) in Southampton, UK.

Extracting insights from data lakes of vehicle network measurement data



Fred.Kihm@hbkworl.com

Data, data everywhere...

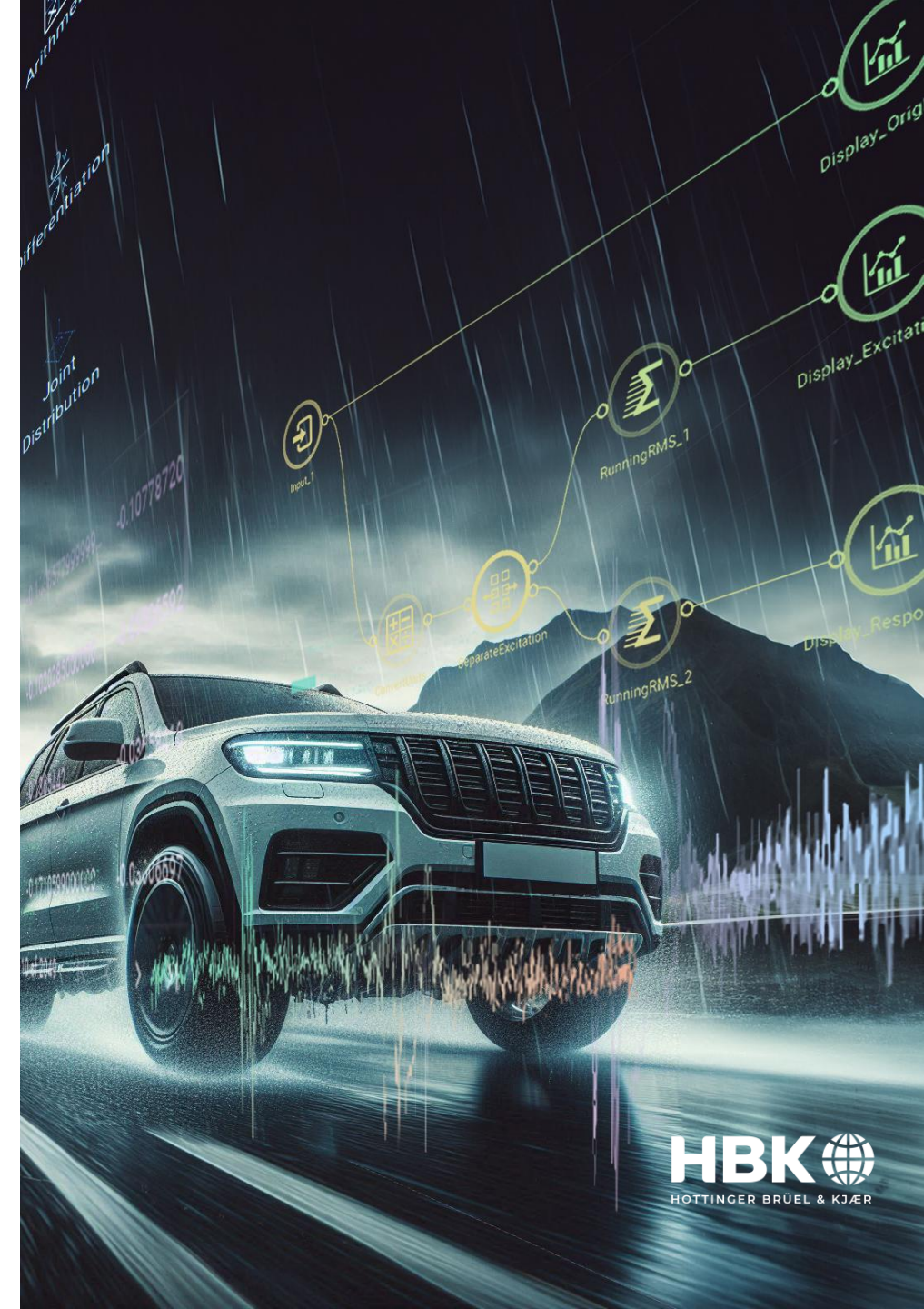
Large quantities of:



Increased volume, velocity, and variety of data

- High quality data acquisition recordings
- Time stamped CAN bus data
- Data lakes of connected sensor measurements

Collecting data is relatively easy but getting the value from all this data is much harder.



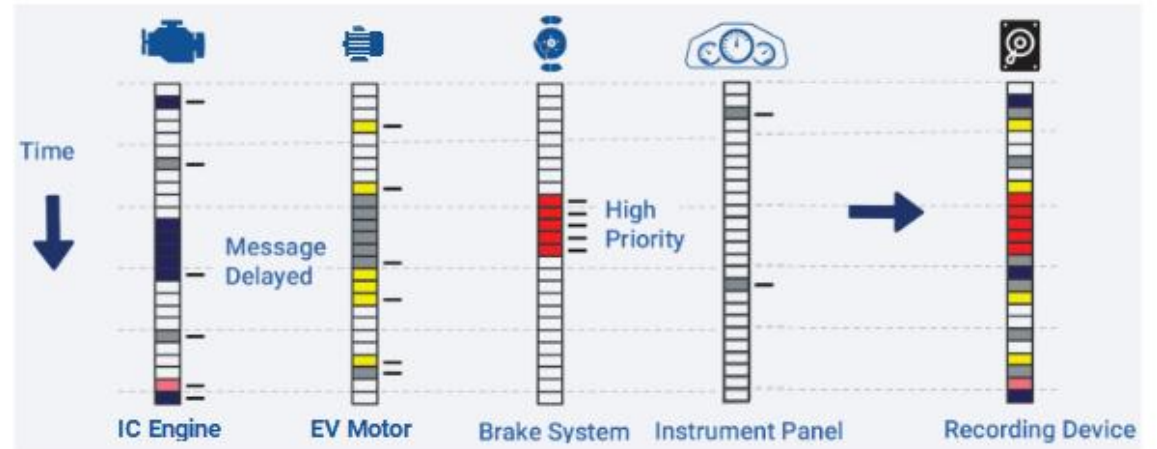
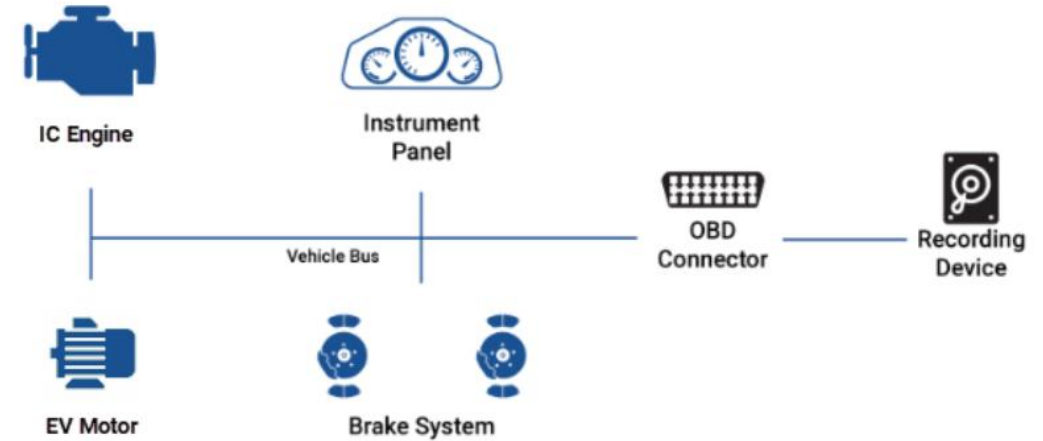
Data from DAQ systems

- ▲ DAQ systems are used to collect sensor data at high quality, high precision.
- ▲ The sensor data is filtered and usually sampled at a fixed rate.
- ▲ But each DAQ system has its own binary format(s), which:
 - may be very difficult to decode,
 - can be proprietary,
 - typically evolves with time,

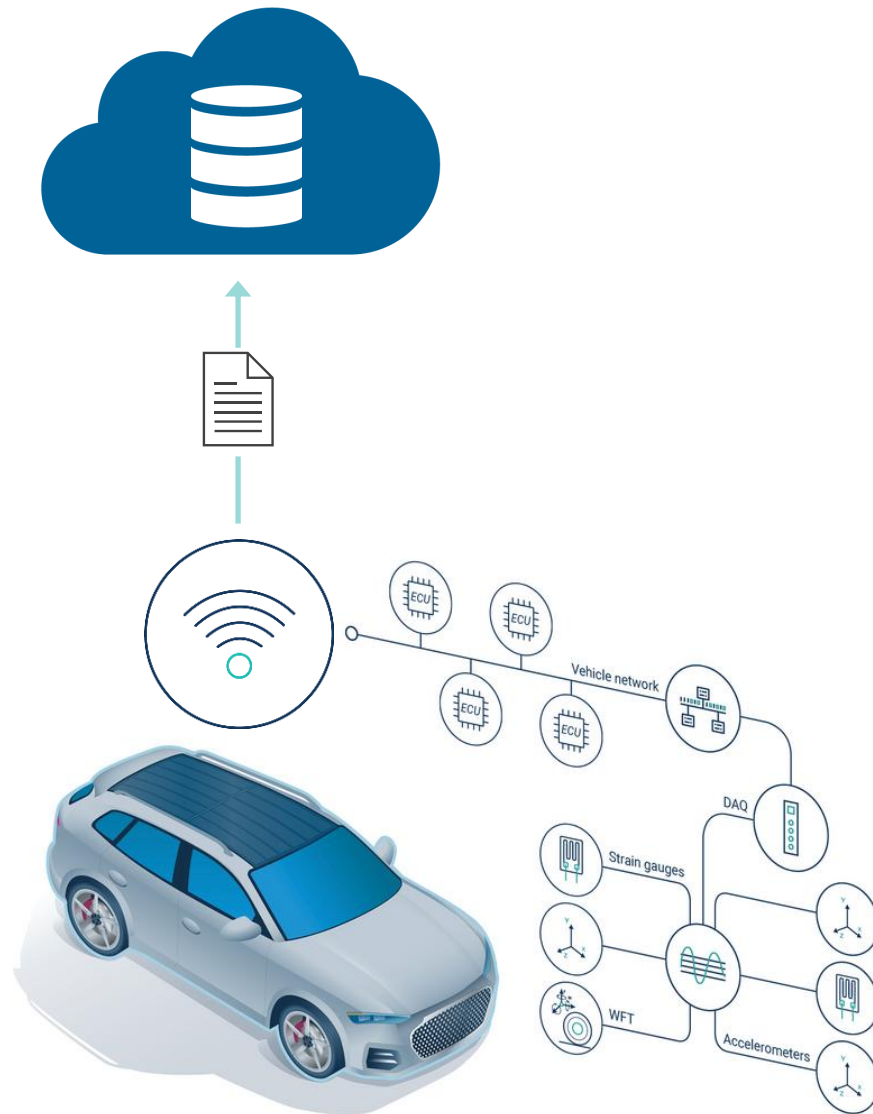


CAN bus data

- ▲ CAN data loggers enable to save messages and / or signals in files (MDF4, CSV, etc.).
- ▲ Conversion to signals with engineering units is often required.
- ▲ But the data is :
 - generally unequally spaced, time stamped,
 - often made of a mixture of numeric values and string, Enum, binary, etc. or just missing (n/a)!



Data lakes

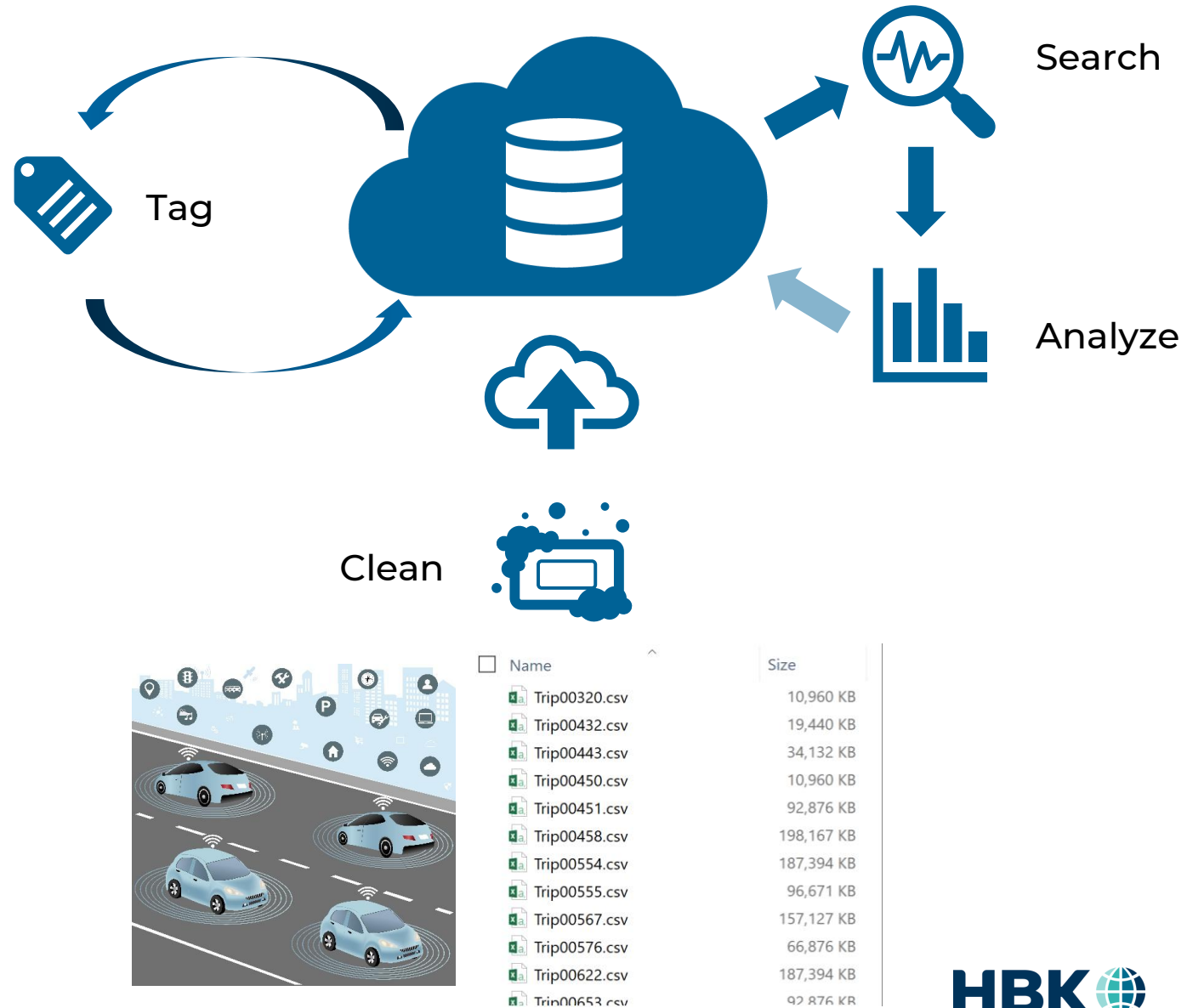


- Automotive OEMs have millions of connected customer vehicles.
- Instrumented test vehicles can also be connected.
- Easy to store large volumes of data in the cloud.
- Increasingly common is storage as Parquet files in data lakes.
- But the volume is huge...

How do you get insights from all this data and avoid a data lake becoming a data landfill?

The overall process

- Make sure your data is useful/usable
 - Ensure titles follow a convention
 - Use consistent units
 - Look for anomalies
- Make your data searchable
 - Tag it with attributes, metrics, indicators
 - More tags can be added at any time
- Make sure you are equipped with the right tools to exploit the data:
 - Perform engineering analyses
 - Find insights
 - Extract scenarios
 - Create predictive models
 - ... at high performance !



Clean the data

Look for:

- Not A Number
- Unrealistic values: either full scale or not physical
- Inconsistent units,
- Lack of naming convention

Unequally
spaced time
values

Correct it:

- replace with numerical value when possible
- Drop the whole row if it is all corrupt
- Apply units conversion and naming convention

TimeCode	VEH_SPEED	LONG_ACCEL	BRAKE_PEDAL	PRESSURE_
#UNITS				
s	km/h	g	binary	bar
#DATA				
6.00E-05	256	-0.2	1	23.4
1.23	256	-0.2	1	23.4
1.35	256	-0.2	1	23.4
3.3	0	-0.2	1	23.4
3.43	2	-0.2	1	23.4
3.47	NaN	-0.2	NaN	NaN
3.52	NaN	-0.2	NaN	NaN
3.56	4	-0.2	1	23.4
3.57	4	-0.2	1	23.4
3.61	4	-0.2	1	23.4

or psi?

Unrealistic
values

Missing values
or “Not a
Number”

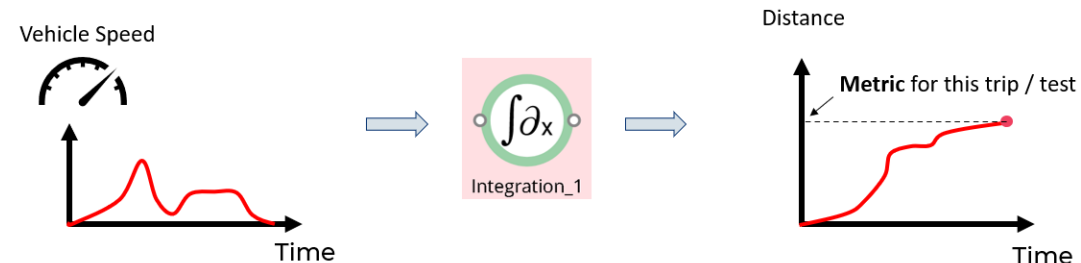
Tag the data

▲ The usual stuff first:

- Calculate stats: maximums, minimums, averages, etc. on each column of data
- Associate the vehicle and driver attributes to the data

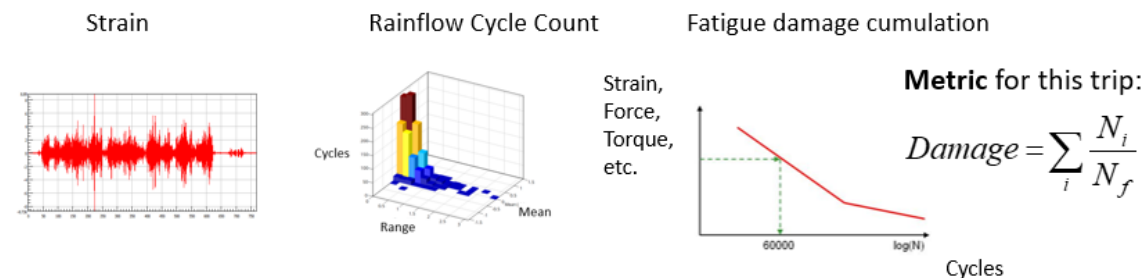
▲ And more advanced metrics: example 1

- Integrate the vehicle speed to obtain distance driven
- Keep the last value as the total distance driven during this trip



▲ Even more advanced metrics: example 2

- Rainflow cycle count signal
- Compute (relative) fatigue damage as a metric



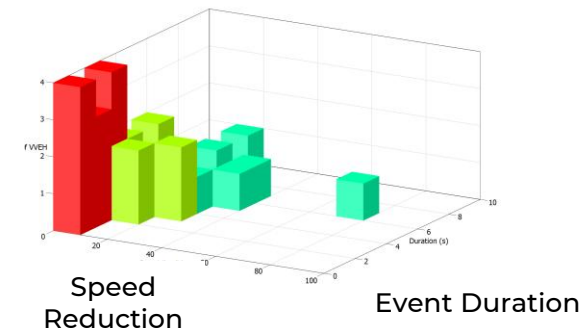
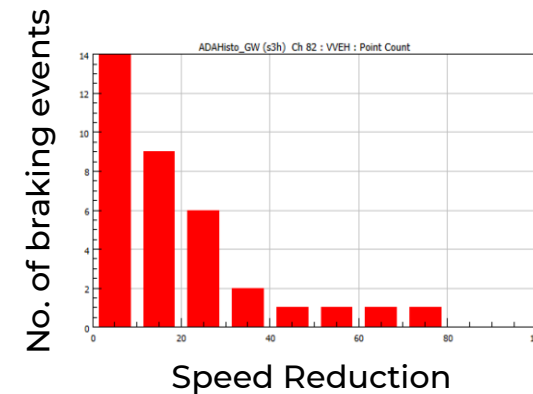
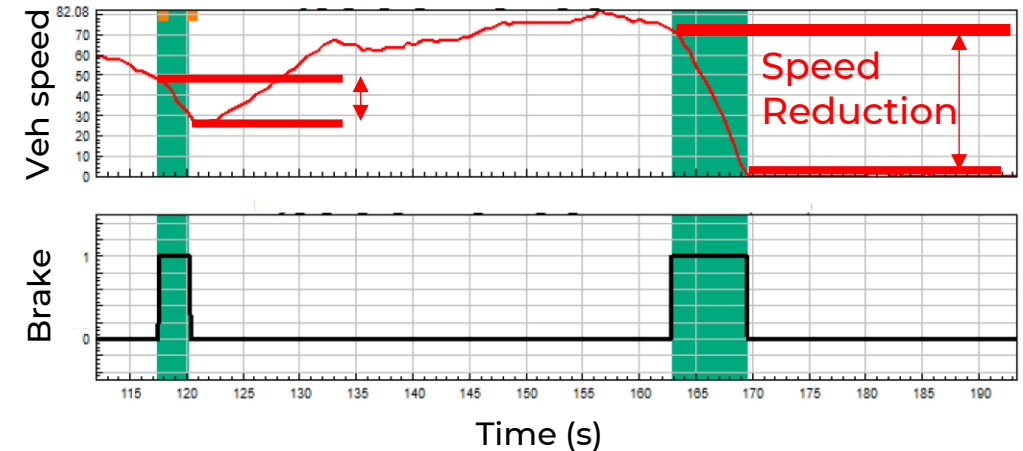
Search and analyze

Search the data for a given vehicle type, date range, user profile, etc.

Example analysis:

- Find every time the brake pedal is actuated
- Extract the corresponding vehicle speed
- Calculate drop in vehicle speed for all braking events
- Build a histogram showing the number of braking events by vehicle speed reduction

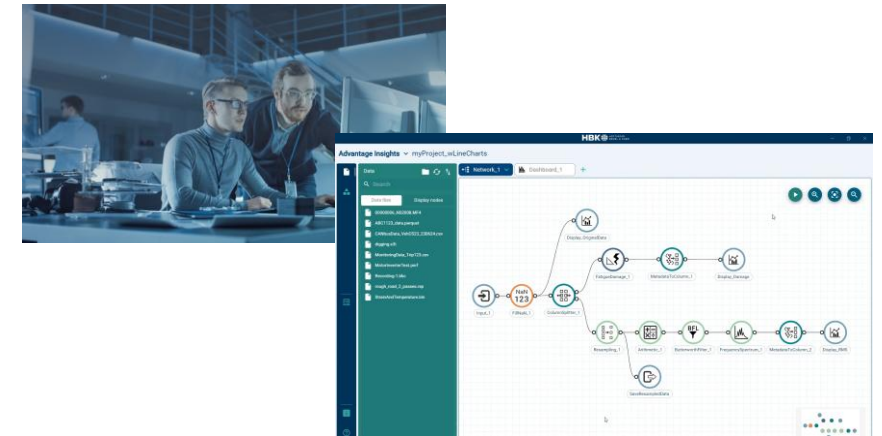
Combine histograms to give an overall result



Conclusions

- ▲ Collecting field data is a good first step but *just* data will not help.
- ▲ Make sure the data is meaningful and tagged with attributes and metrics in a data lake.

- ▲ Search for specific data and run analyses in order to:
 - understand product usage,
 - identify and extract scenarios for simulation and simulators,
 - create predictive maintenance models, etc.

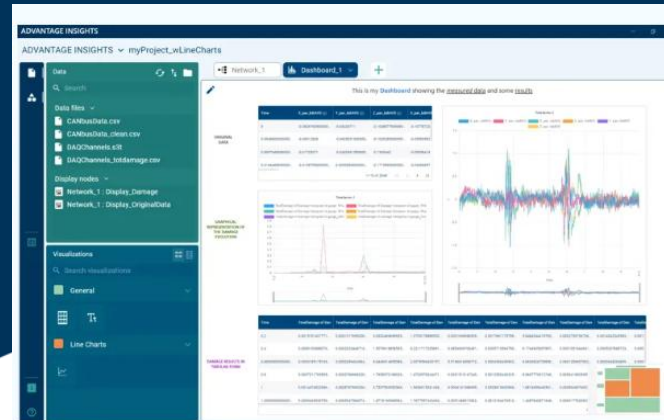


- ▲ HBK can help with:
 - **Tools** to augment the data with calculated engineering metrics, and to manage/share/deploy.
 - **Advantage Insights** lets engineers easily create, edit, and run powerful sensor data analytics, esp. for large time series of measured data.
 - **Aqira** is a web-based system, enables remote work by creating, sharing, and running engineering apps and provides robust test data management.
 - **Expertise** in engineering/physics, data acquisition and data analysis.

Thank you

hbkworld.com/advantageinsights

Fred.Kihm@hbkworld.com



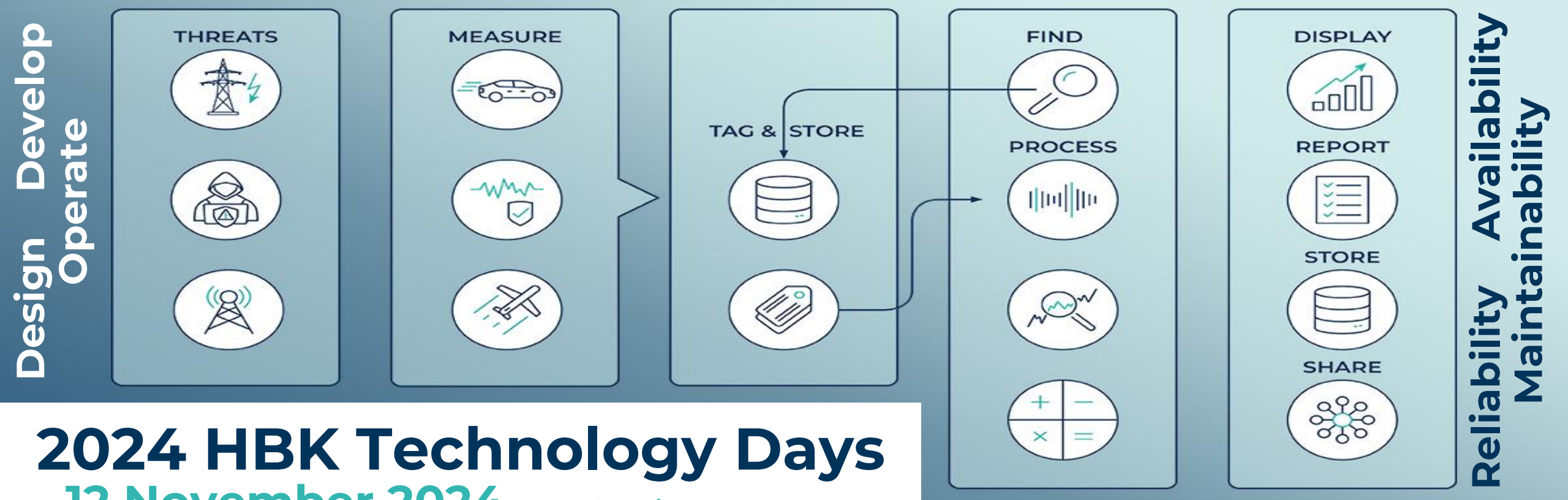
Turn Data Overload into Actionable Insights - Introducing Advantage Insights empowered by nCode

Join this exclusive webinar for an introduction to *Advantage Insights* – HBK's groundbreaking solution designed specifically for engineers that need to cut through the noise of massive data streams. In today's fast-paced, data-driven world, having access to quality information is key to making the best decisions.

REGISTER FOR FREE >

Questions?

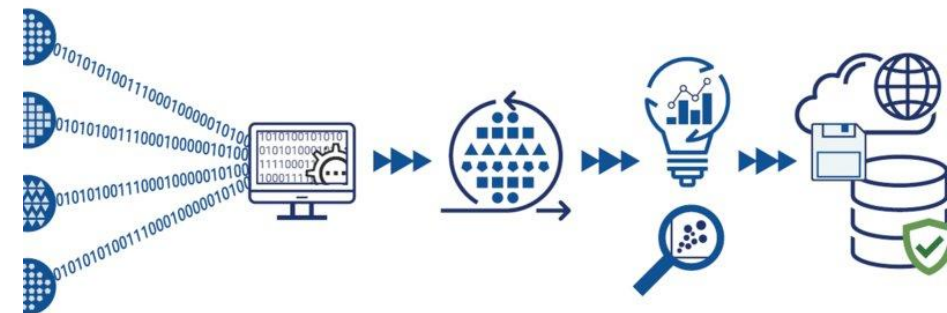
Fred Kihm – Product manager
Fred.Kihm@hbkworl.com



2024 HBK Technology Days

12 November 2024

Session 4: Vehicle Network Analytics



- Extracting insights from streams and data lakes of vehicle network measurement data

Frédéric Kihm,
Product Manager – Analytics and Signal Processing, HBK