

# THALES

Building a future we can all trust



## How to find a needle in the haystack

Improved Log Analysis & Big Data

Aymeric HOLL AUS



# Unique Diversity of Expertise

This document may not be reproduced, modified, adapted, published, translated, in any way, in whole or in part or disclosed to a third party without the prior written consent of Thales - ©Thales 2023. All rights reserved.



**420**  
Employees 

**26**   
Nationalities

**52%**  
Export rate

R&D Centre in Ground  
Transportation within  
Thales Group

**80%**  
Engineers

**€91.4** million  
Sales in 2021

## Key Facts

# Portfolio for Integrated Solutions

Field of expertise: Main Line Rail  
Signaling Solutions

Local R&D, Engineering and Services



## Route Control Systems

Electronic Interlocking



## Train Control Systems

European Train Control System (ETCS) – trackside & onboard – all levels



## Traffic Management Systems

Operation Management Centre (OMC), Aramis TMIS



## Complementary Solutions

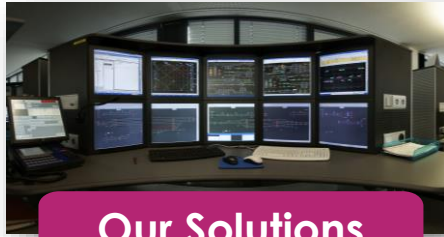
Axle counter, LED Signals, SCWS, Decentralised Element Controller, CheckPoint



## Transverse Digitalisation

Services, Cybersecurity, Security

# Introduction



Our Solutions



```
112740 X25-STATUS: SCU.-1 (1/1)
112741 Receive Line 3 Layer 2
112742 Zeit: 0 sec Layer2: up: 0 down
112743 Good Frames: 7638 Bad Frames: 0
112744 IGN 0 RR 3470 RMR 0 REJ
112745 DISK 0 UA 0 FRMR 0 BADN
112
```

Log Data



Information



Create improved log analysis environment

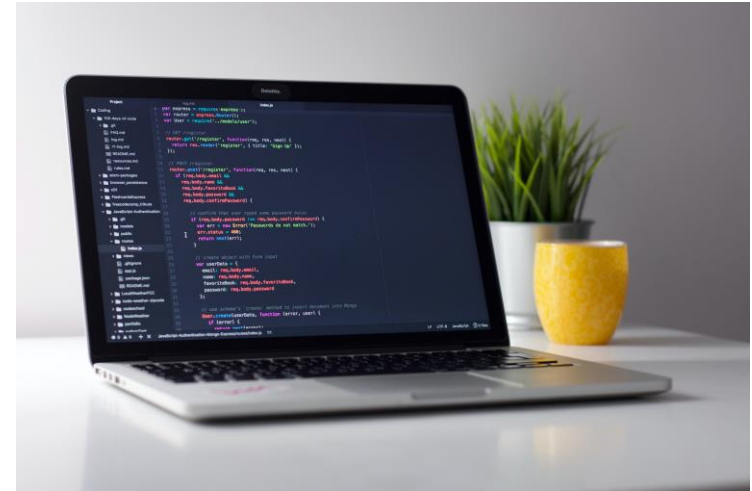


Helps to efficiently **monitor and analyze log data**  
Visualizes & identifies **issues** more quickly  
Provides insights into **system behavior & patterns**

**Make complex information & data  
easy to understand for everyone**

**Facilitate analysis of systems for  
customers and engineers**

**Digitalization is key for knowledge  
sharing**



## How to find the needle in the haystack

- No need to search log files manually
- Easy-to-access web interface to data
- Available examples as reference to do advanced analysis
- Possibility of access to data, independent of country or product
- Refined data base for general queries



## DGP Diagnostics Processor

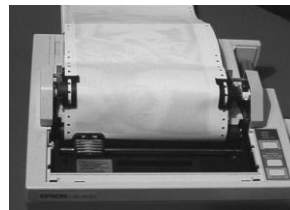
Due to historic reasons, original format is based on a terminal output

One index pattern but each country has its own configuration in our stack

- Indices generated based on station country – year – month
- Sorted based on timestamps

Data are being generated daily and stored permanently

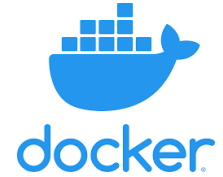
```
112740 X25-STATUS: SCU.-1 (1/1) 22-01-17 23:58:33 (461)
112741 Receive Line 3 Layer 2
112742 Zeit: 0 sec Layer2: up: 0 down 0 Layer1: down 0
112743 Good Frames: 7638 Bad Frames: 0
112744 IGN 0 RR 3470 RNR 0 REJ 0 INFO 4168 SABM 0 DM 0
112745 DISK 0 UA 0 FRMR 0 BADN 0 UNK 0 XID 0 BADL 0
112746 TO 0 L2UP 0 L2DN 0 L1DN 0 CON 0 T4 0 T4N2 0
```



# Our architecture & stack

**Kubernetes** allows for simple deployment and scaling

---



**Logstash** handles data ingest and field/pattern definition

**Elastic** is our distributed search engine

---



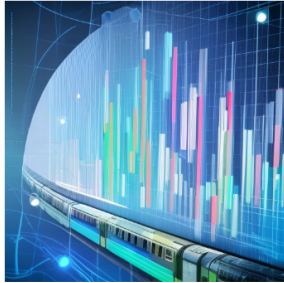
**Kibana** and **Jupyter** allow us to visualize and analyze our data





# Dashboards 1

This document may not be reproduced, modified, adapted, published, translated, in any way, in whole or in part or disclosed to a third party without the prior written consent of Thales - © Thales 2023. All rights reserved.

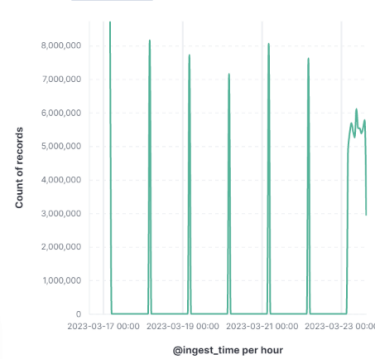


## Information on our Data

Welcome to the **Information on our Data** Dashboard on Kibana! This dashboard provides a high-level view of our search-engine based database of *DGP* messages. You can expect to see visualizations that give you a better understanding of the data. With this dashboard, you can easily spot trends and patterns in the data quality, giving you valuable insights into the performance of our tool *ILONA*. Thank you for visiting!

[Feature Request? Issues? Feel free to contact us by clicking here!](#)

## Ingest Time

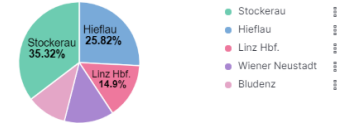


## Total count of records

2,771,623,629

Count of records

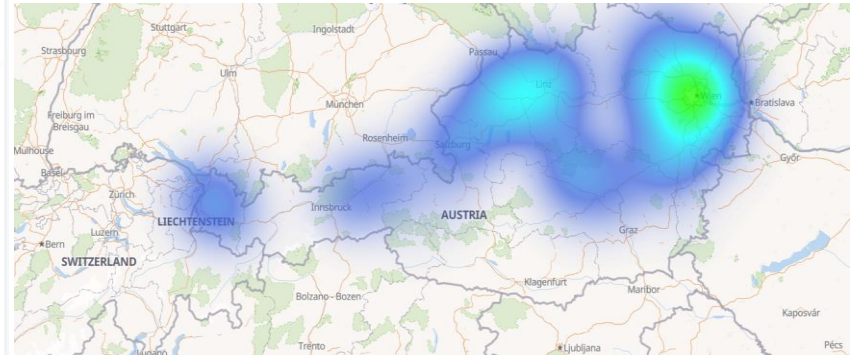
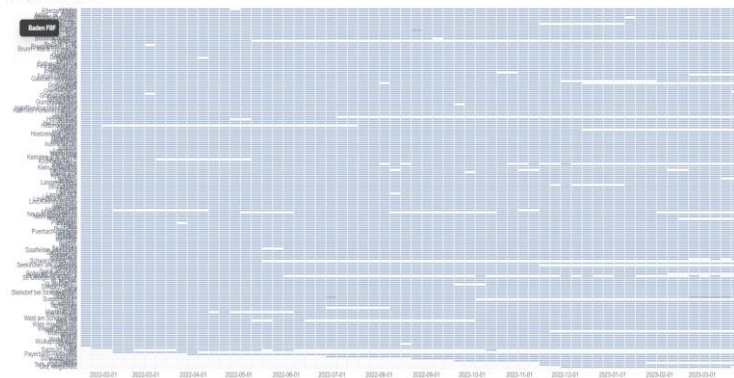
## Top 5 Stations with most data in Austria



48.57%

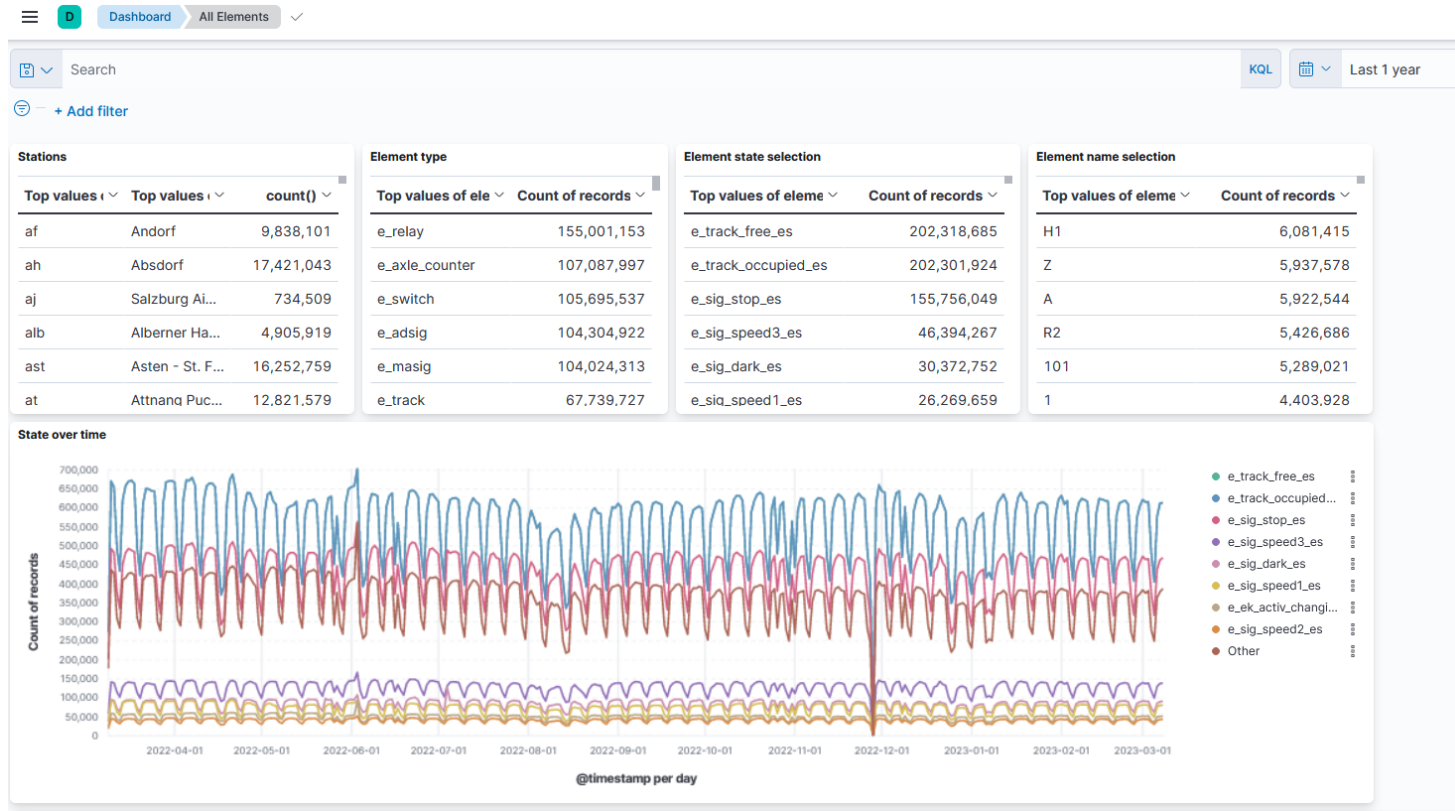
Approximate Pattern Coverage

## Heatmap of Stations data over time



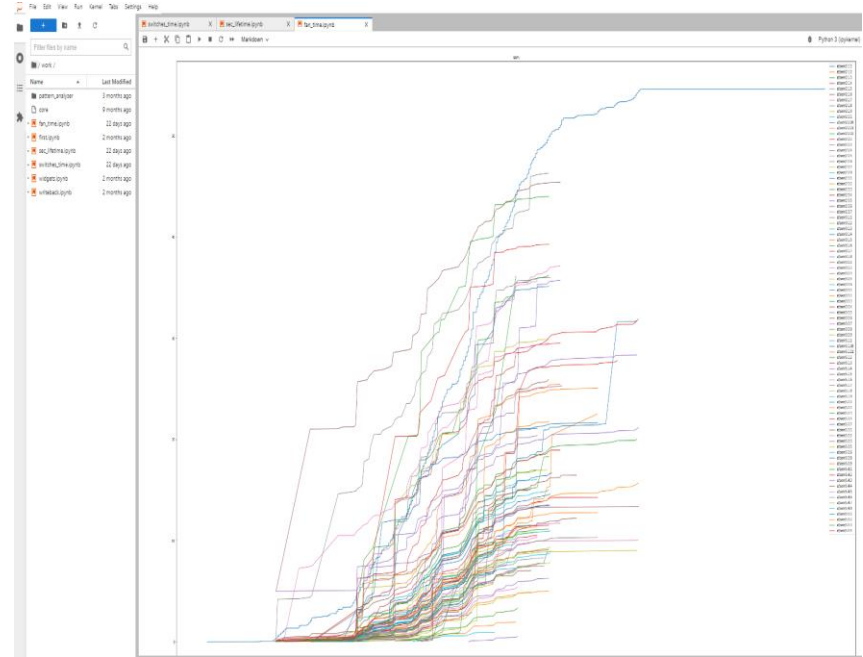
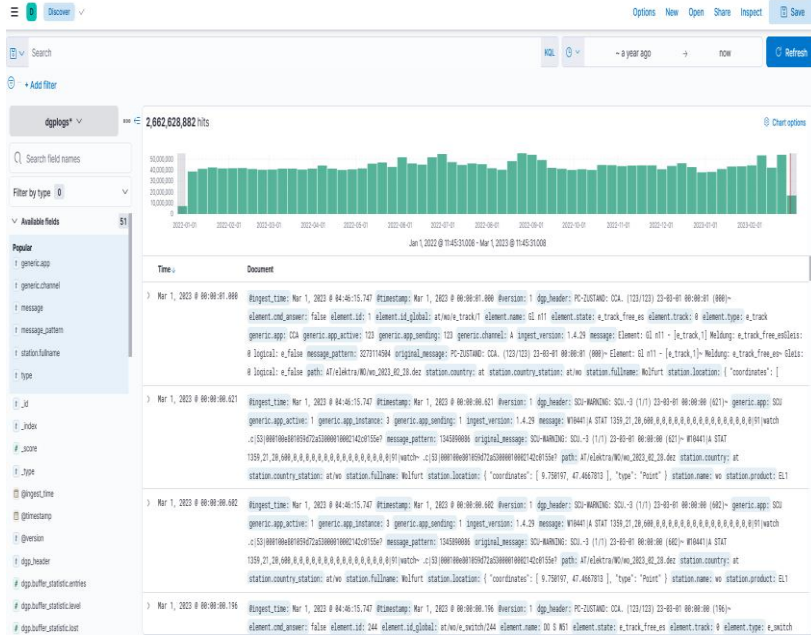
# Dashboards 2

This document may not be reproduced, modified, adapted, published, translated, in any way, in whole or in part or disclosed to a third party without the prior written consent of Thales - ©Thales 2023. All rights reserved.



# Queries & Data Manipulation

This document may not be reproduced, modified, adapted, published, translated, in any way, in whole or in part or disclosed to a third party without the prior written consent of Thales - © Thales 2023. All rights reserved.



# Target User Base

- Employees that work within RAMS
- Service Center and Field Operations
- Product Managers
- Customers
- Engineers that wants to tinker around 😊



# Achievements & Prospected Results

**Improved system reliability** through early detection of issues

**Increased visibility** into performance and usage patterns

**Faster troubleshooting** and resolution of field issues

**Improved collaboration and knowledge sharing**

**Streamlined log analysis workflow**

**Customizable dashboards** and visualizations

**Utilize excellent synergy with other technologies such as asset management**

---

**Make use of advances in machine learning for data analysis automation**

---

**Integrate additional data sources and analysis tools to enable more powerful insights**

**Goal for the future** = ensuring the reliability, availability,  
and maintainability of our solutions & drive  
ongoing optimization and improvement initiatives



OPEN