The Paranal DataLab: first steps towards advanced system monitoring and maintenance schemes

EIROFORUM Workshop Big Data, 28.10.2020

Christian Stephan for the ESO DataLab Team

Current Paranal Configuration

VLT (UT, VISTA & VST):

- 6 telescopes
- 14 instruments
- AO and AOF

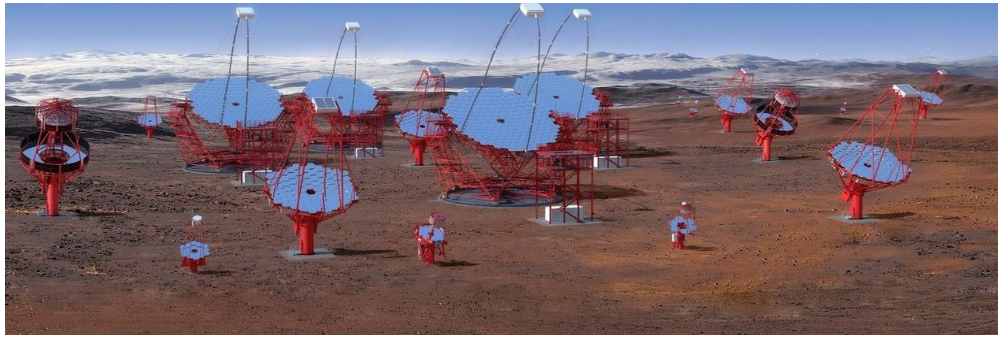
VLT (VLTI):

- 4 telescopes
- 8 AO Systems (UT & AT)
- 3 instruments

Supporting Systems:

- Power Supply
- Cooling Supply
- Coating Units

Evolution of Paranal Configuration

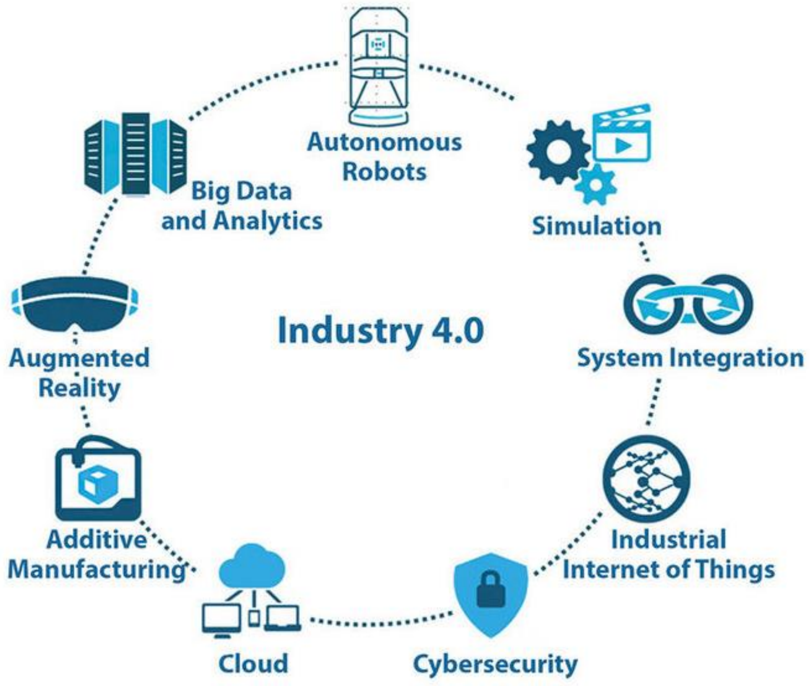


Credit: Gabriel Pérez Diaz, IAC / Marc-André Besel, CTAO

Integrated Operations Programme

Prepare Paranal for multi-side observatory operations:

- **Remote** (be able to control and monitor systems without being physically present at system location):
  - Minimize on-site activities (inspections and corrective maintenance)
  - Make better use of support from Vitacura and Garching
- **Lean** (use available resources efficiently):
  - Optimize and automize processes (operations, maintenance and logistics) in order to make better use of available manpower.
  - Provide necessary infrastructure for efficient multi-side operations
- **High Performance** (enhance/maintain performance & availability):
  - Strengthen Data & System Analysis to operate systems at the best possible performance for given conditions
  - Ensure that the necessary competences for the operation of multiple complex systems are available



DATA & SYSTEM ANALYSIS FOR PARANAL OPERATIONS

Operations in the Middle of the Atacama

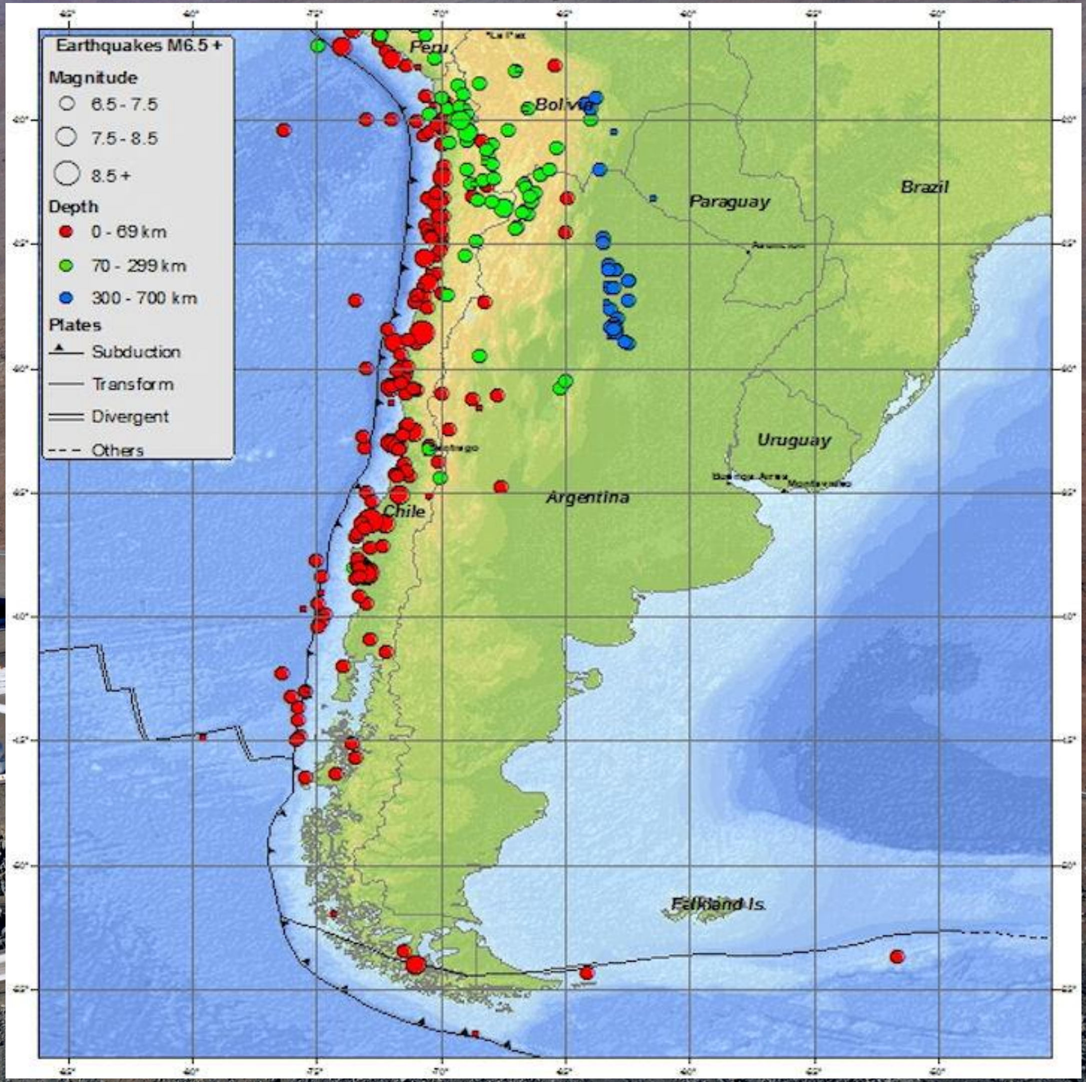
Operations in the Middle of the Atacama

- 24/7/365 Operations in remote location → staff working in 8/6 shifts.
- Telescope lifetime > 50 years, instruments > 15 years
- Infrastructure can host only a limited number of staff on site, only limited remote support from ESO HQ and instrument consortia.
- Living Environment, no isolated, well controlled lab environment.
- Science drives and constraints maintenance activities.

Operations in the Middle of the Atacama

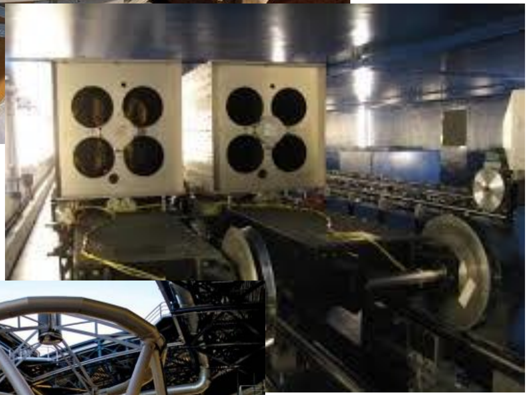
Operations in the Middle of the Atacama

Operations in the Middle of the Atacama



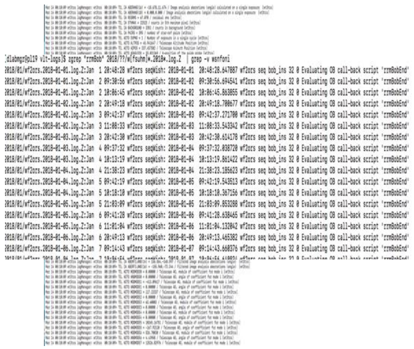
General Challenges

- Wide range of Systems:
  - Unit Telescopes and technical infrastructure were designed and build in the 90s
  - Wide range of scientific instruments with specific needs and capabilities
  - Active/Adaptive Optics with fast control loops and hundreds/thousands of actuators
- Operational needs sometimes underestimated in the requirements → missing/insufficient monitoring means
- Quality and availability of data.
- Many stakeholders with different knowlege's, experiences and expectations.
- Institutional bias that "data" means "science data" → missing infrastructure for operational and engineering data storage and processing.



A bit of History –
Different Tools for Data Analysis and Maintenance

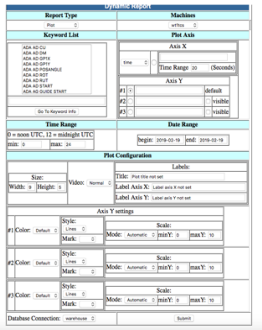
LogMonitor, ('90)
VLT LOGS (text)



- Sufficient for simple analysis
- Hard to conduct more complex tasks

Retention:
- Weeks (local)
- All (in Garching Archive)

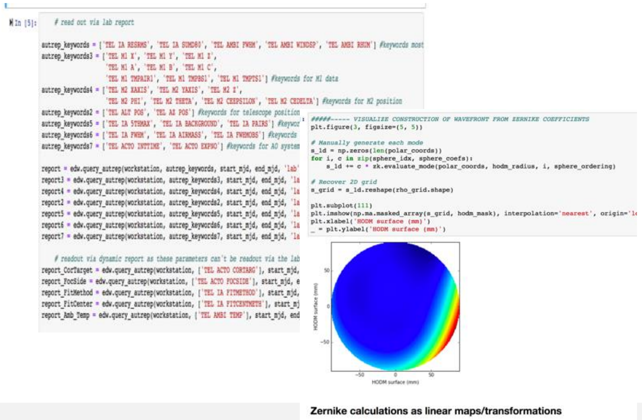
Autrep (~2005)



- Graphical (web) user interface
- Scheduled reports Do not store text logs
- Updated daily
- Complex analysis require an external tool

Retention: 1 Year in Paranal

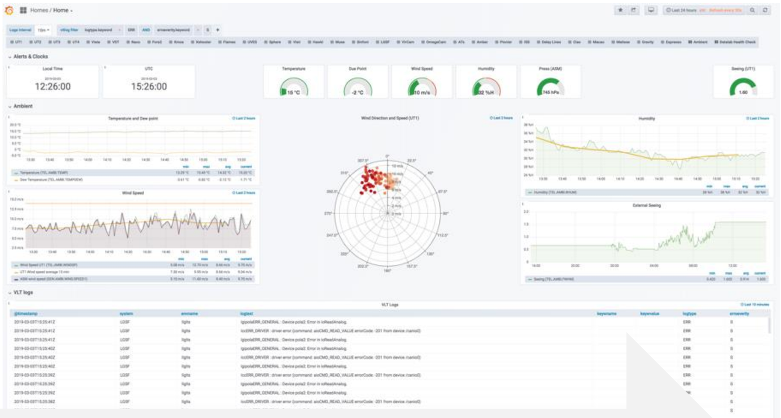
Sysmon (~2014)



- Uses Python / Jupyter for data analysis
- Permits to share the analysis
- Hard to maintain and use
- Use of HDF5 file

Retention: 1 Year in Paranal

DataLab (~2019)

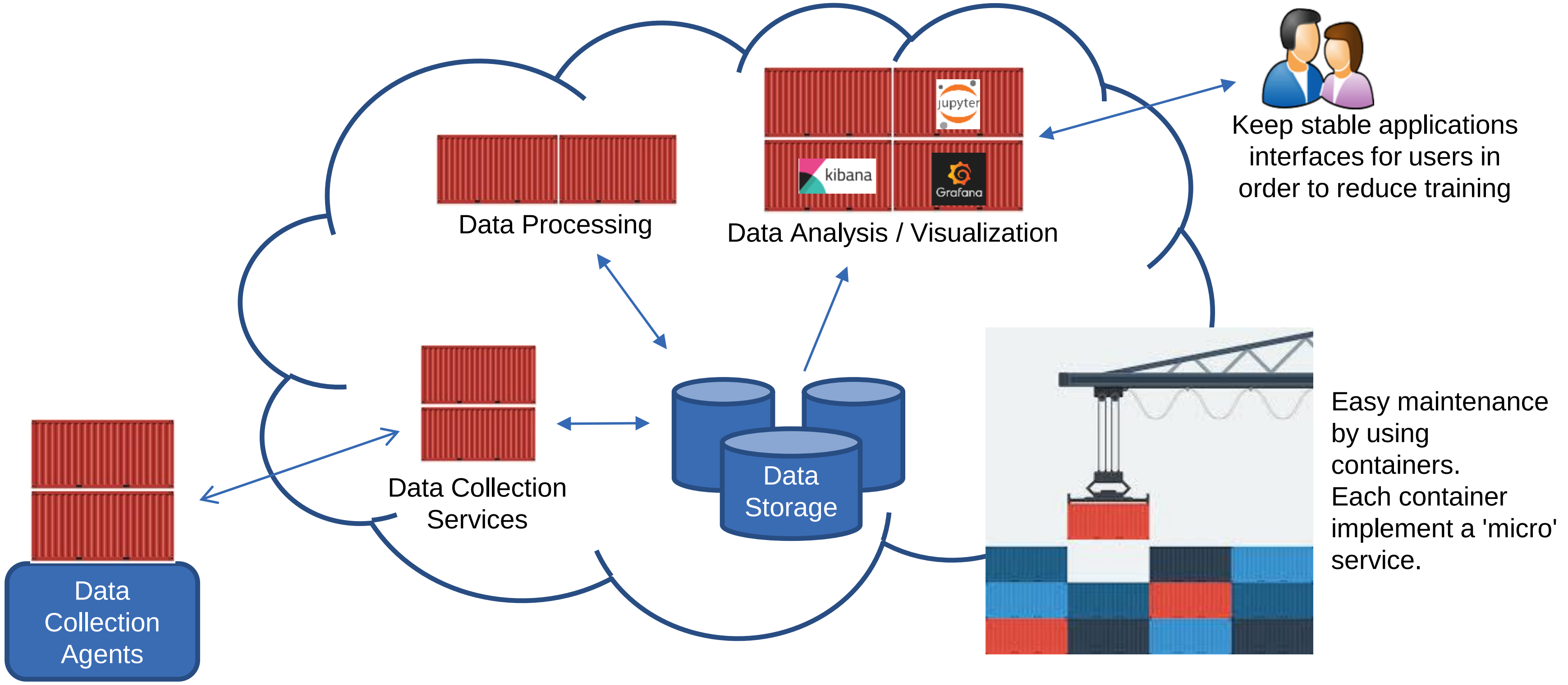


- Can use more data sources
- Including online data
- Schedule complex reports
- Uses a combination of several specialized tools

Retention: system lifetime

DataLab in a Nutshell

- Flexible and scalable infrastructure for data storage and processing.
- Allow to support all aspects of observatory operations.
- Gives room for staff to experiment with new techniques in a safe environment.
- Provides tools and services for user with different knowledge in data analysis.



Available tools for Data Analysis

AutRep
Automatic Reports

kibana

Grafana
Playground

jupyter

lab
JupyterLab

Reports Gallery

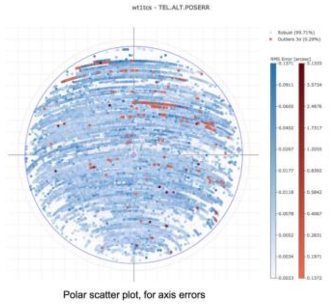
voilà
Interactive Reports

QoW
Quality of Work

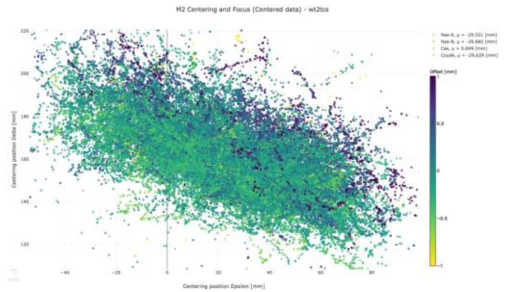
KairosDB
Time Series DB UI

Current Performance Analysis & Troubleshooting

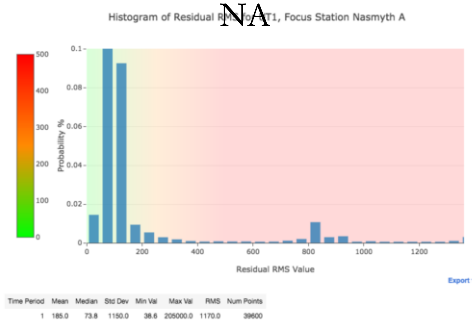
Axis error polar (alt/az)



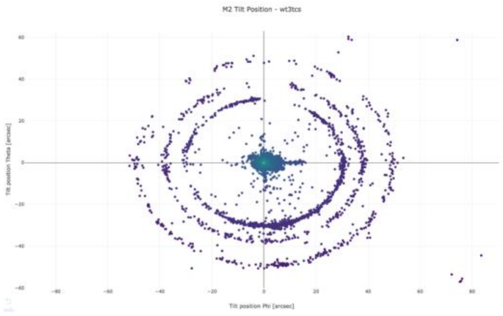
M2 Centering and Focus



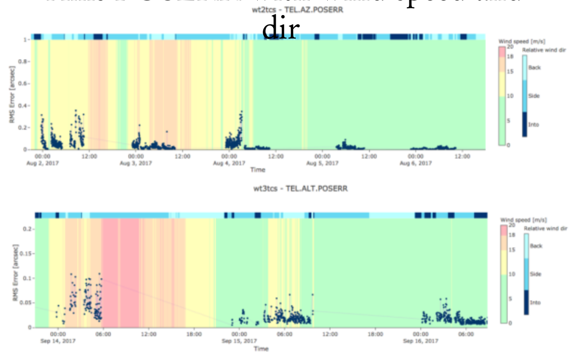
Histogram of Residuals RMS UT1, NA



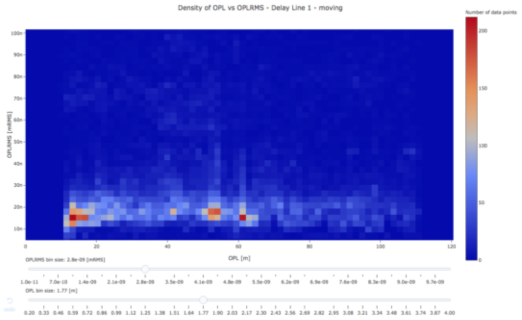
M2 tilt position (Phi, Theta)



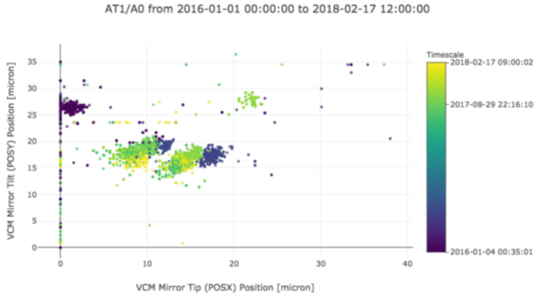
Axis POSERR with wind speed and dir



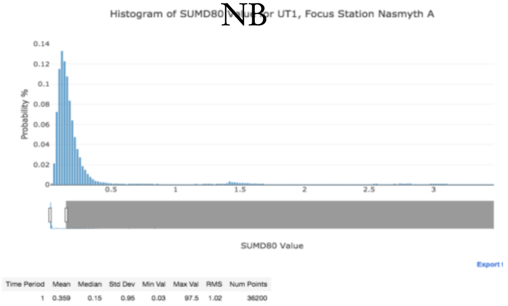
Density of OPL vs OPLRMS



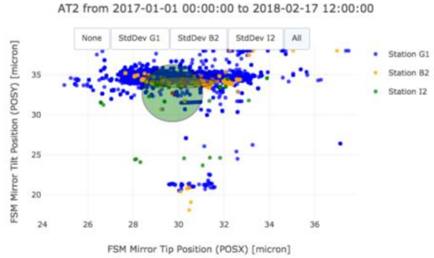
AT VCM Pointing, Long term



Histogram of SUMD80 UT1, NB

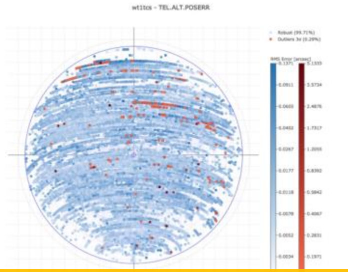


AT FSM Pointing, Long term, with user selectable StdDev region



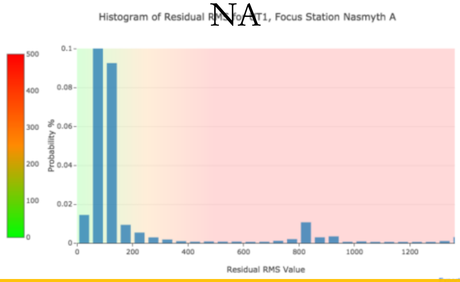
Current Performance Analysis & Troubleshooting

Axis error polar (alt/az)



M2 Centering and Focus

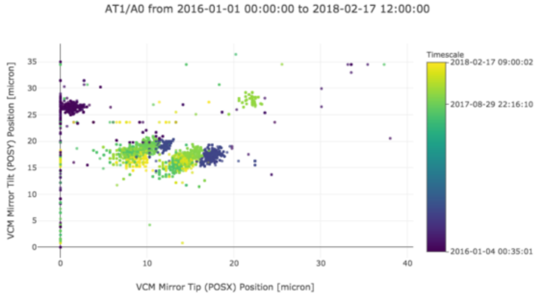
Histogram of Residuals RMS UT1, NA

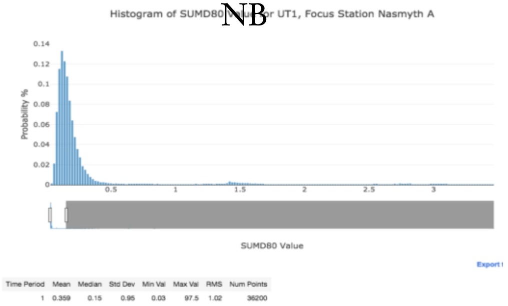


Tools are only used to visualize data, interpretation done by people!

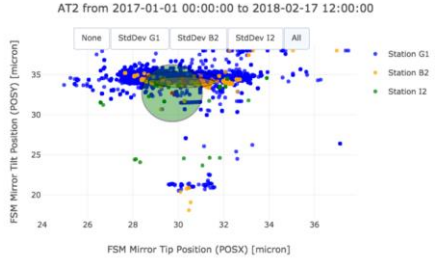
How to use the date for predictive maintenance?
How can we automatize anomaly detection?

Can we enhance from predictive maintenance to predictive operations?

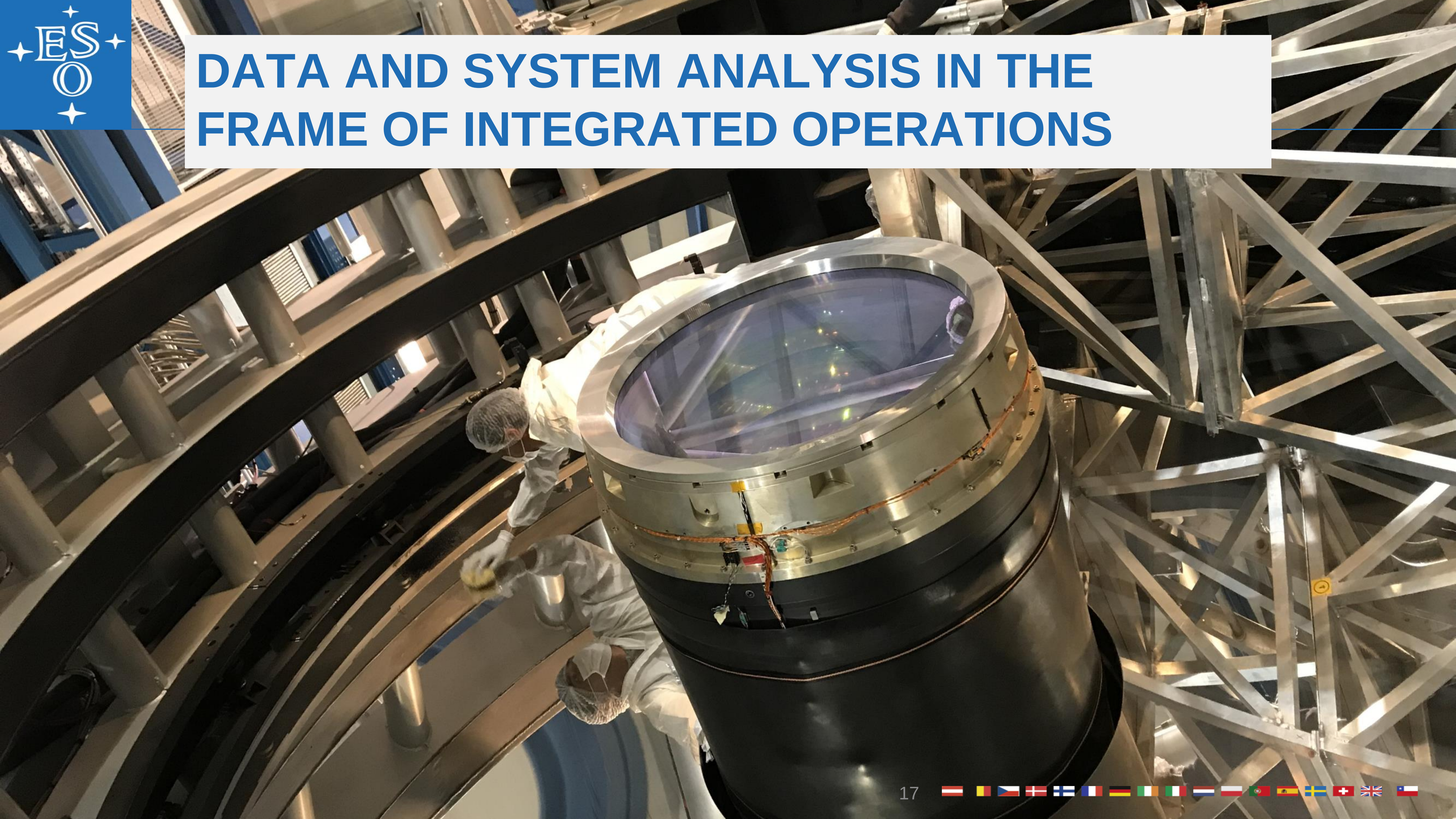




user selectable StdDev region



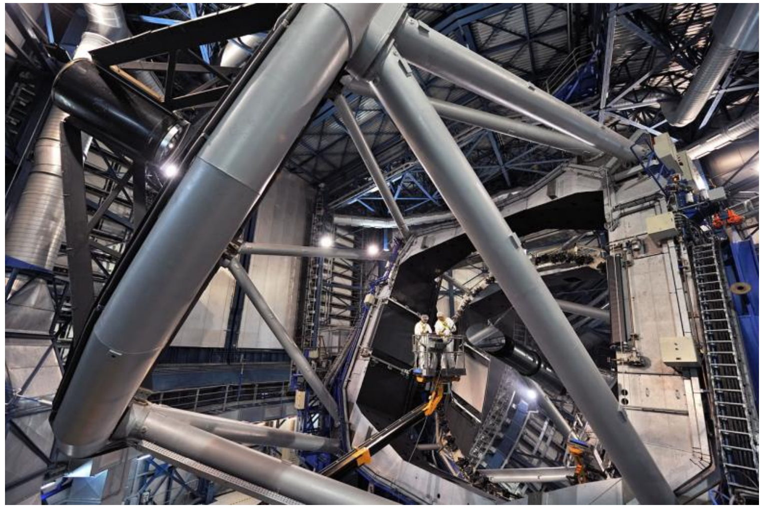
DATA AND SYSTEM ANALYSIS IN THE FRAME OF INTEGRATED OPERATIONS



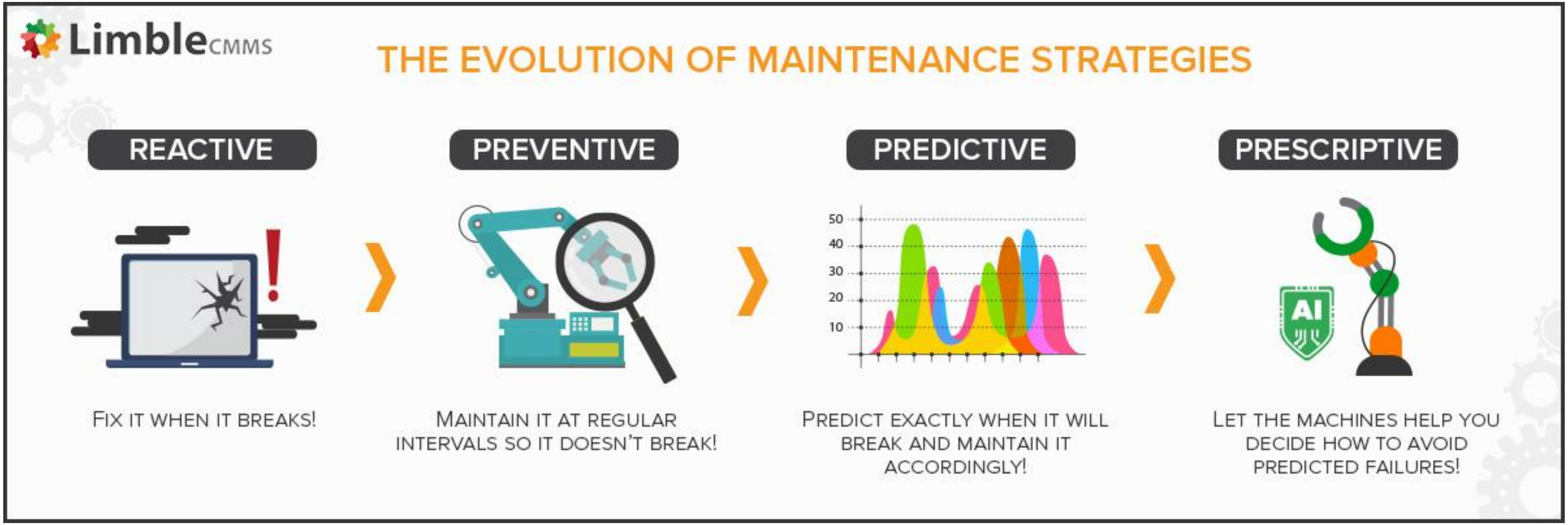
Paranal Data & System Analysis Approach

Improve operational performance and efficiency:

- Understand and enhance observatory performance:
  - Be aware of baseline performance for given conditions.
  - Understand impact of system degradations on global performance in order to prioritize actions
  - Being able to predict operational performance in advance and online during operations.
- Plan and prioritize daily necessary maintenance actions during night based on operational needs and KPI.
- Build up and maintain core competences in relevant disciplines in a sustainable manner

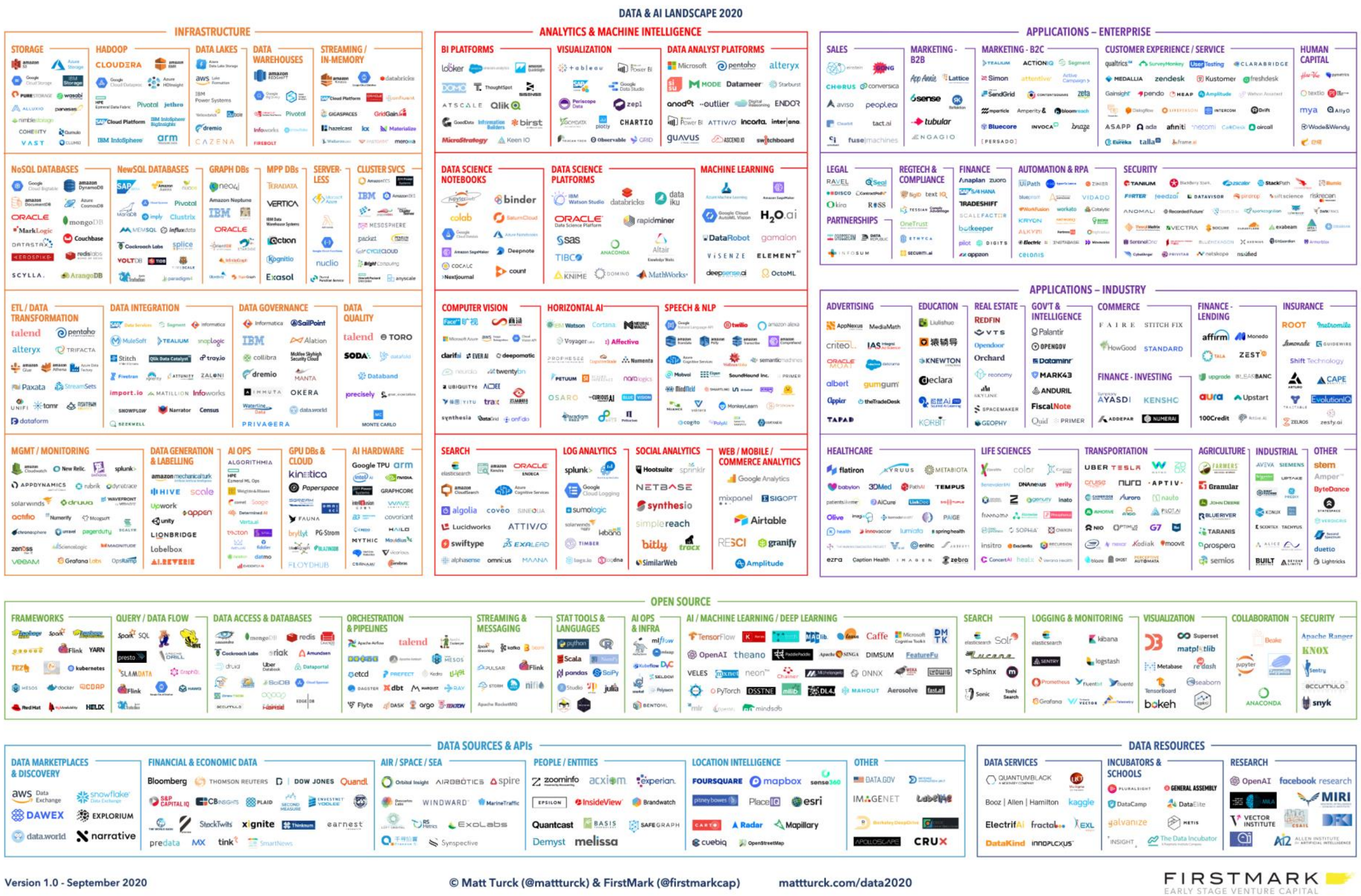


System & Data Analysis in Maintenance

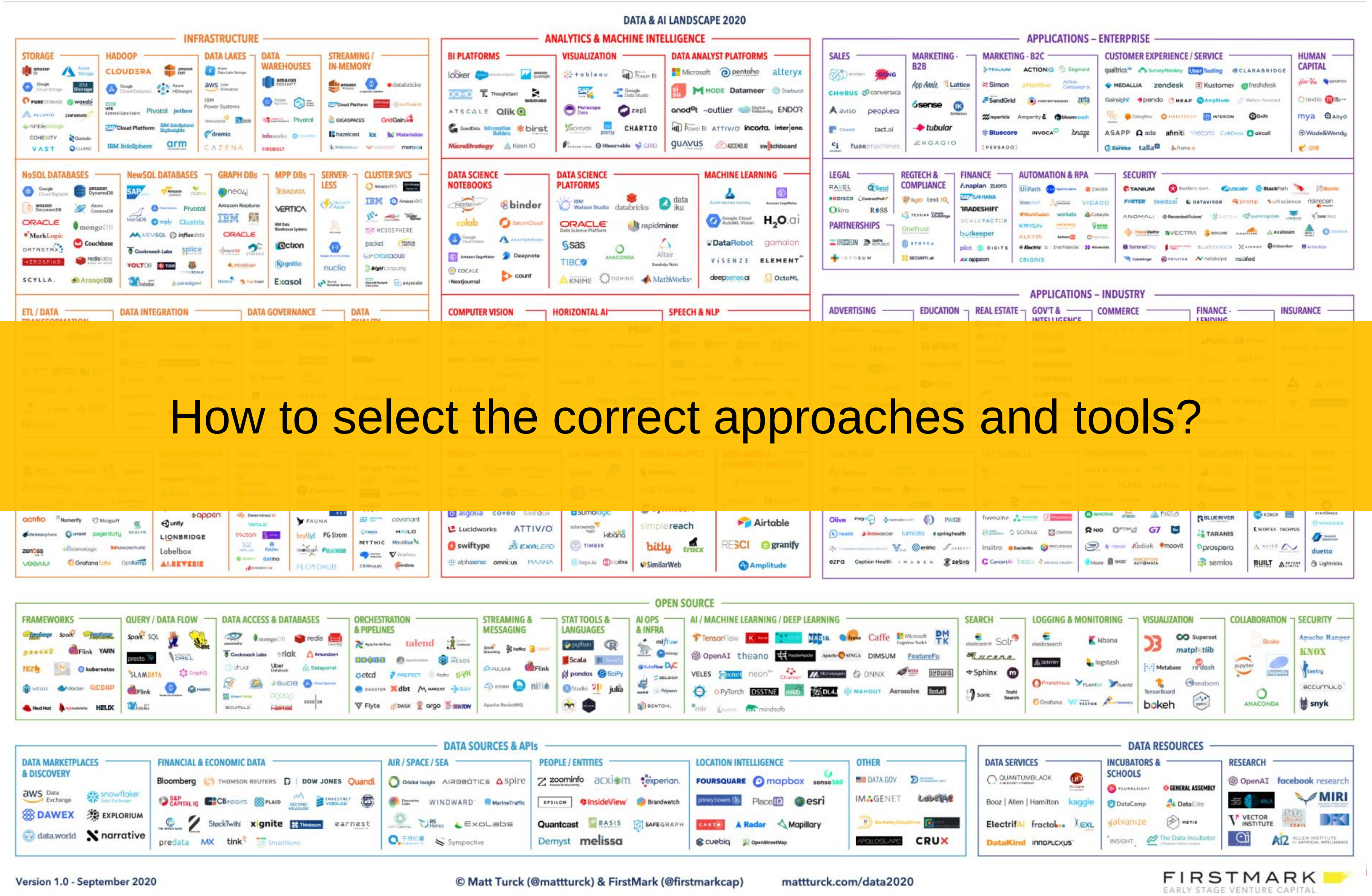


Source: https://www.omnisci.com/technical-glossary/predictive-maintenance

Data & AI Landscape 2020



Data & AI Landscape 2020



How to select the correct approaches and tools?

How to Increase Knowledge & Experience?

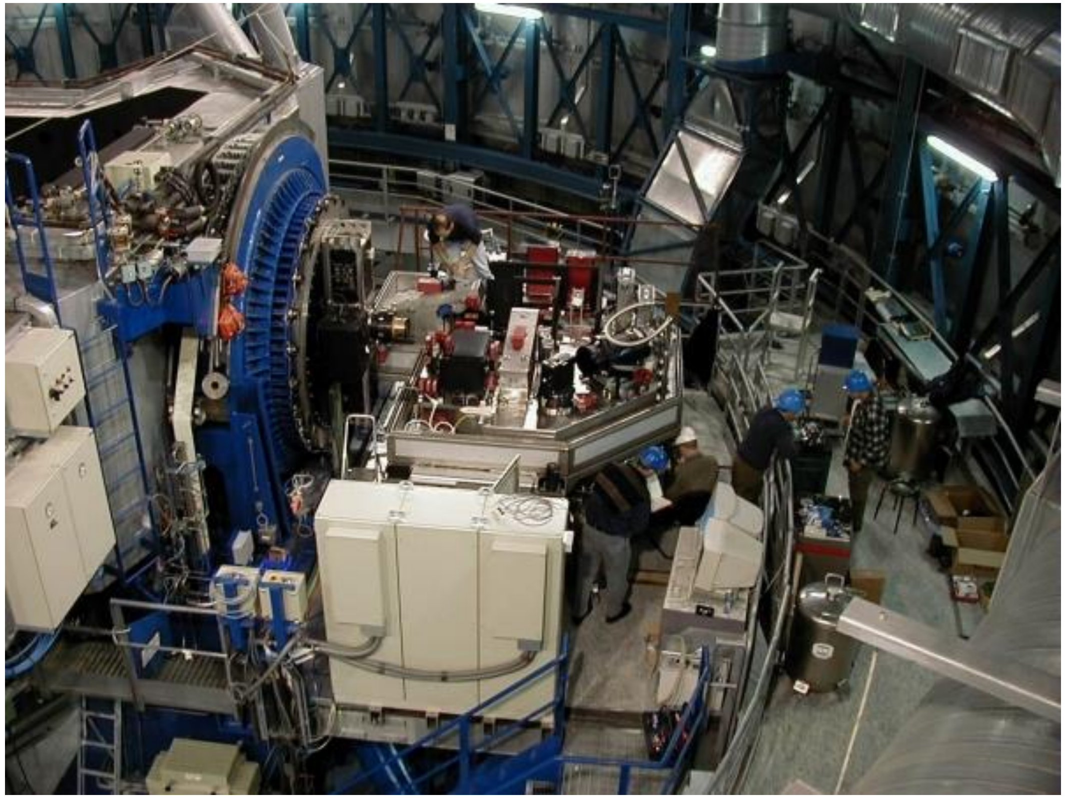
- Clarify operational needs and perform trade-offs:
  - Identify critical systems and KPI
  - Statistical methods vs classical machine learning vs deep learning…
- Hire experienced staff to coordinate and guide → slow process
- Form a dedicated team to bundle experiences to provide services to the observatory.
- Provide general and tailored training
  - Use online platforms like edX
  - Collaborations with universities for specific and applied training
- Give staff room to experiment and share experiences
  - DataLab Coffee
  - DataLab Contest
- Seek partners for collaborations don’t invent the wheel new

ESO – MICROSOFT COLLABORATION

UVES Anomaly Detection on Calibration Data

Ultraviolet and **V**isual **E**chelle **S**pectrograph

- High spectral resolution
- Main spectroscopic instrument



Instrument calibrations:

Day-time calibration frames

- Used in the generation of final science data

Data production may suffer from defects

- Detector problems (electronic noise)
- Optical problems (mis-alignments)
- Instrument internal problems (lamps)

Calibration images are processed automatically, and spot checked by a human for possible problems.

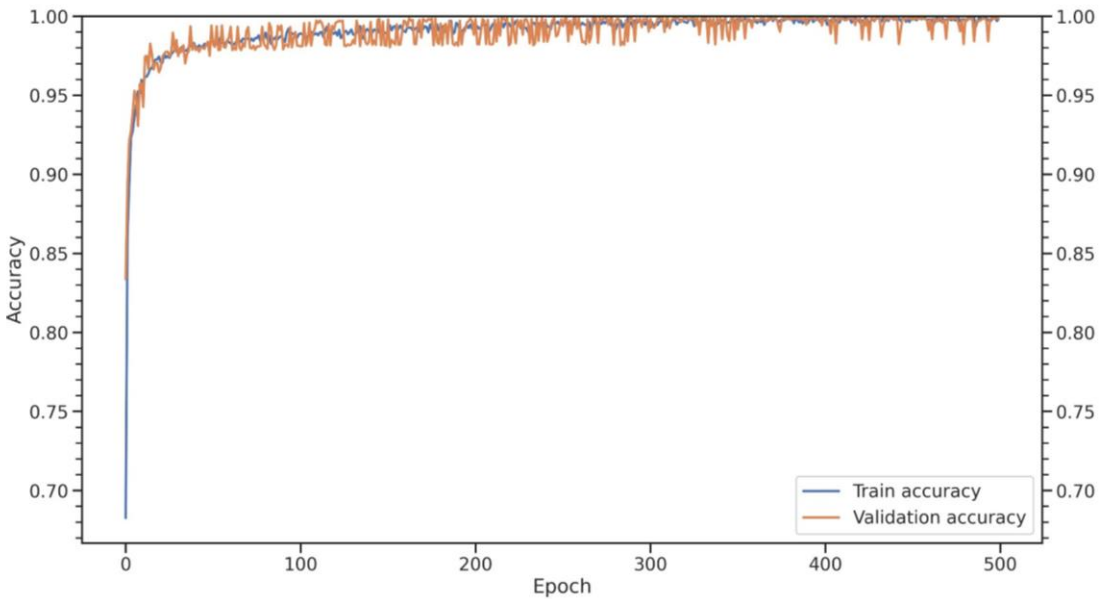
Challenges:

Big format images (2k x 4k)

11 different types of images

Few anomalous images in the full data set

UVES PoC – Used Approach & Results



- Instead of having a single model to identify anomalous images several sequential steps and create classification models for each step. The first step consist of identifying what calibration type the images belongs, classification model trained using thousands of images and deep learning techniques.
- Instead of training a classification model from scratch, we decided to apply transfer learning from a classification model trained using terrestrial images. We used a Resnet50 architecture trained on the Imagenet dataset and removed the last classification layer, so the output of this model is a descriptor of length 2048.
- Model for 500 epochs trained and reached an accuracy of 0.99 → proof on concept, now implementation into operation.

PARANAL COLLABORATIONS WITH UNIVERSITIES

From Pilot Projects to Integrated Project Teams

UT Vibration mitigation

- Define a nominal vibration spectrum per UT from one year of data
- An algorithm able to trigger an alarm when an anomalous vibration spectrum is measured on a telescope
- An algorithm able to learn and build up a database of frequent vibrations sources

UVES spectral resolution

- An algorithm that can discriminate between normal resolution variations (due to temperature, pressure) from variations due to malfunction of the instrument (like wrong slit settings).

UT axis behavior

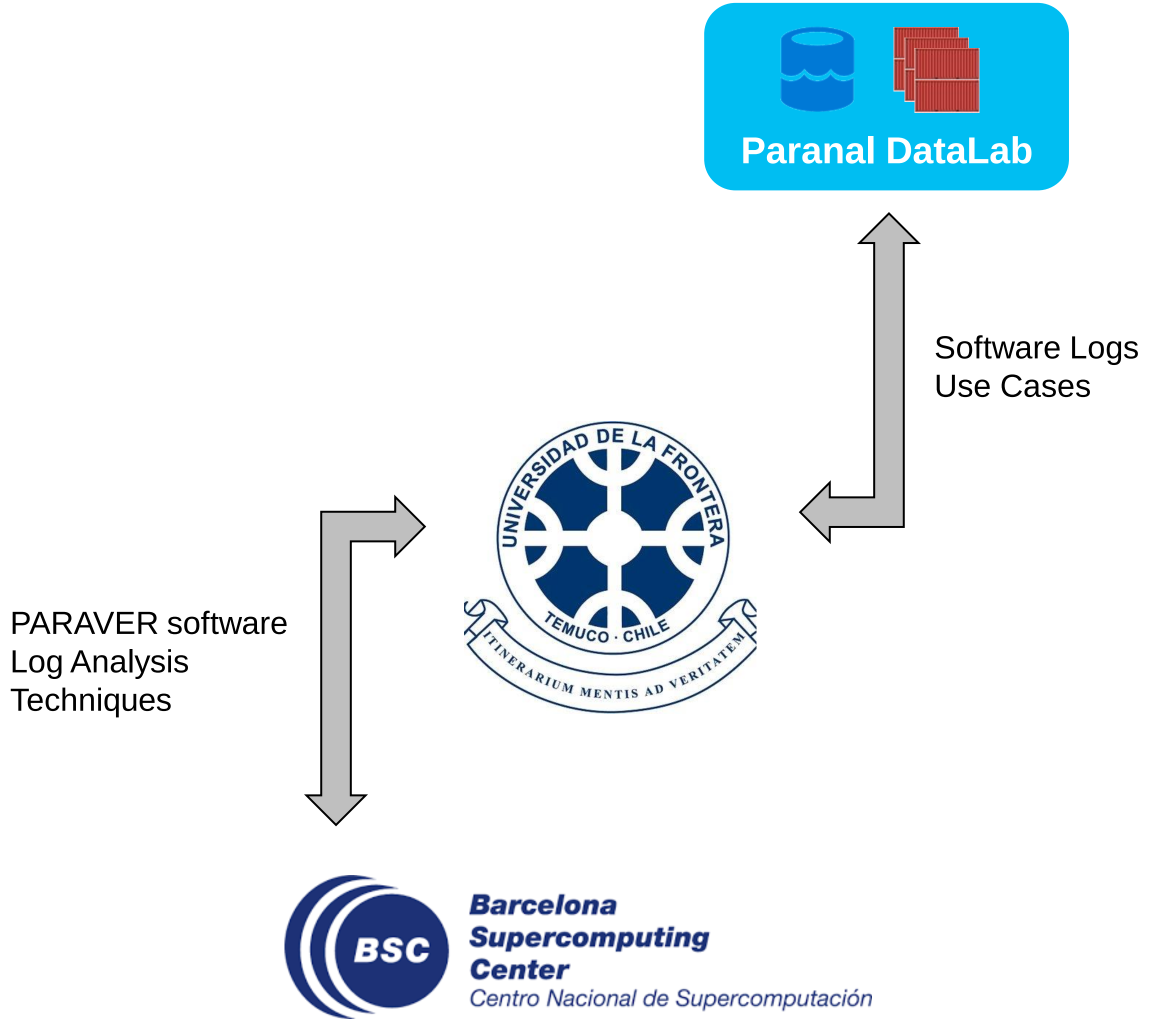
- An algorithm that can analyze the axis torque behavior in order to allow a better monitoring of the Hydrostatic Bearing Systems.

→ Promising first results → Move away from isolated pilot projects towards an integrated approach to establish a real operational use case → Form joint project team with strong ESO involvement to avoid ”customer-client” relationship → Use projects as hand-on training activites

Paranal – UFRO – BSC Collaboration

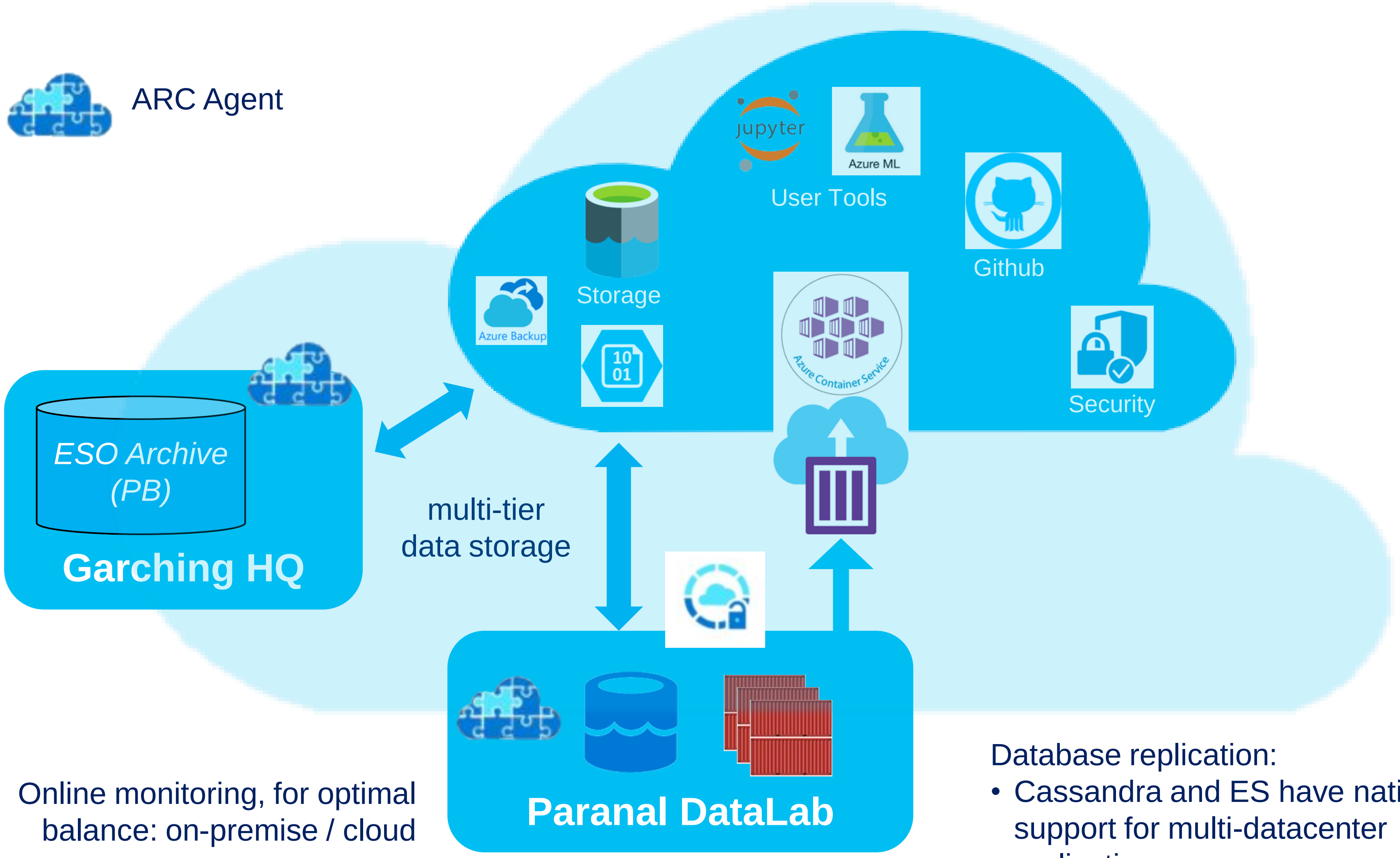
Use Paraver to analyze Paranal logs

- Paraver is a flexible data browser part of the CEPBA-Tools toolkit, from Barcelona Super Computer Center (BSC). It analyses logs from their MARE NOSTRUM cluster.
- Instruments at Paranal observatory generates software logs that are used daily to solve operational problems.
- Universidad de La Frontera (UFRO) act as a broker by transforming Paranal logs into Paraver format, and by collecting BSC best log analysis practices to be applied back in Paranal.



DATALAB IN THE CLOUD

DataLab + Azure



Online monitoring, for optimal balance: on-premise / cloud

Database replication:

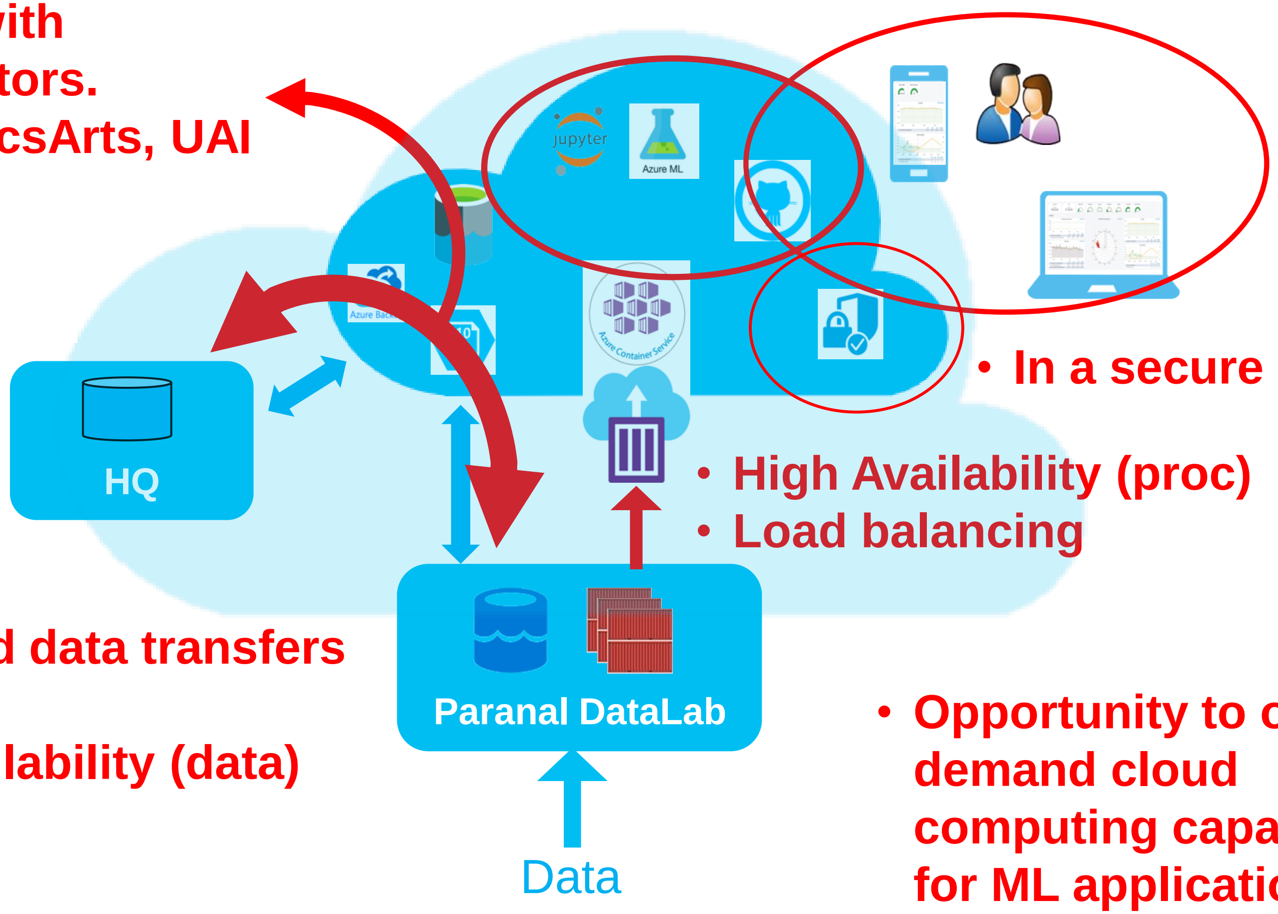
- Cassandra and ES have native support for multi-datacenter replication.

DataLab + Azure: Expected Benefits I

- Simplify data sharing with collaborators. e.g. MetricsArts, UAI

- Pre-Installed tools
- With access to data

- Global (secure) Access
- Collaboration (github)



- In a secure manner

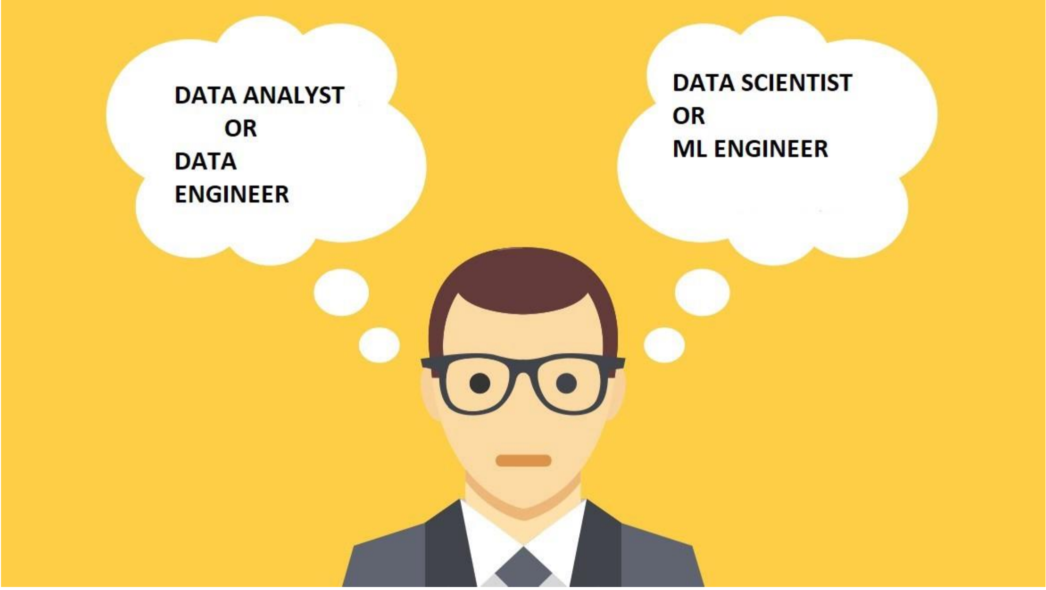
- High Availability (proc)
- Load balancing

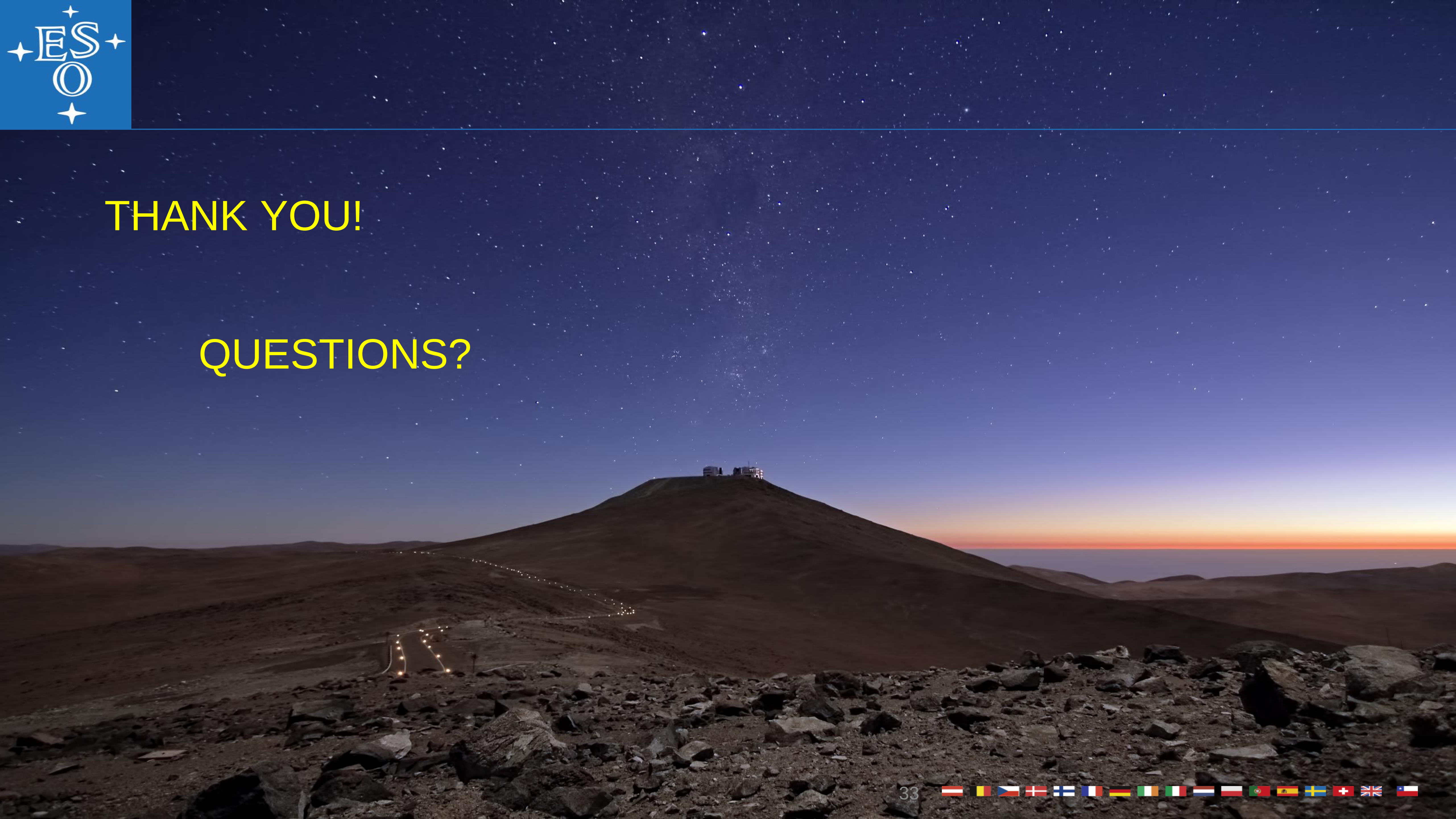
- Simplified data transfers
- Backups
- High Availability (data)

- Opportunity to on-demand cloud computing capabilities for ML applications

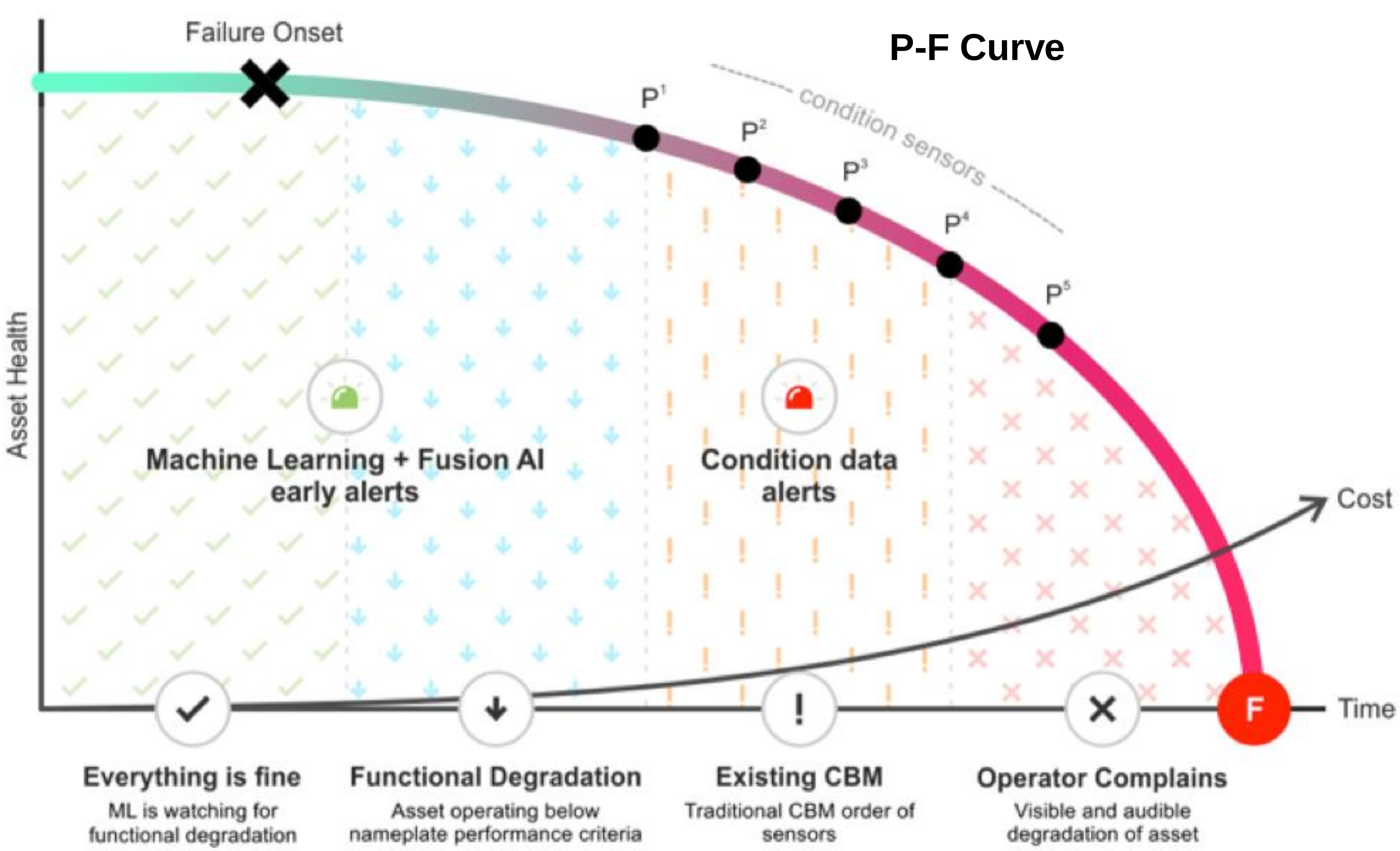
Paranal Data & System Analysis – Next Steps

- Intensify collaborations to enhance experiences
- Clarify necessary competences needed for operational data & system analysis.
- Introduce a quality and user support scheme for DataLab
- Implement a light integrated operations pilot to evaluate usefulness of advanced data & system analysis schemes
- Play & Experiment!!!




THANK YOU!
QUESTIONS?

Why Invest into Machine Learning for Maintenance?



Provide the proper tools to enable enhanced Conditions Based Maintenance capabilities

→ better control and optimization of maintenance effort.

Source: https://www.quartic.ai/smart-industries/predictive-maintenance