An Integrated MBSE Analytics Platform (IMAP) -

Tracing AI based analysis of AIT/AIV telemetry data results to MBSE models

Armin Müller, Michael Benkel ScopeSET GmbH Im Rothmoos 5 83730 Fischbachau, Germany amueller@scopeset.de	Andreas Lindner 3DSE Management Consultants GmbH Seidlstr. 18a 80335 Munich, Germany a.lindner@3dse.de	Rosalinde Borrek, Florian Schummer Chair of Astronautics Technical University of Munich Boltzmannstr. 15 85748 Garching, Germany rosalinde.borrek@tum.de

Paper-Submission for the MBSE2022 Workshop – Addressing the Objectives[1] O-3, O-4, T-3 and S-2.

Abstract

MBSE and telemetry/operations data are separated areas today and the tracing of telemetry data back to MBSE models is time consuming and requires a lot of manual effort. Therefore System models (MBSE models) are often not used in assembly/integration/test (AIT) phases and become outdated.

This paper introduces the Integrated MBSE Analytics Platform (IMAP) which has been developed in the course of the OSIP[2] MBSE campaign's MBSE2DataLake project. It demonstrates how MBSE data can be downstreamed from early phases and integrated with Telemetry data or system operations data. Applying MachineLearning analytics and customized visualization on these sets of integrated data provides extended analytics capabilities and seamless data browsing experience for the user.

The approach enables linking of the former separated domains of System Design and AIT/Operations. This in turn leverages the MBSE contained knowledge across later phases of system development lifecycle, as well as providing a single point of access for cross-domain users to the combined data.

The platform is implemented as an operational technology demonstrator (TRL 3), with publicly available data from the MOVE-II project[3], including a SysML model and an operational Telemetry database from the project's AIT phase.

State of the Art and Challenges

At the beginning of the OSIP activity interviews with stakeholders and space engineers have been conducted. Based on the interview results and the experiences from former projects of partners in this activity, a clear picture of the state of the art has been perceived.

MBSE and telemetry/operations data are still silos today, both in terms of the people involved as well as in terms of the data and models generated or used in daily business. Although partial "point to point interfaces" of data exist, a holistic integration has not been realised. In the early space project phases serious amounts of time and effort are invested to create valid MBSE system models. These Models are often used for documentation purposes only and in later project phases mostly not used at all. Not

[1] https://indico.esa.int/event/407/page/602-objectives
[2] https://www.esa.int/Enabling_Support/Preparing_for_the_Future/Discovery_and_Preparation/The_Open_Space_Innovation_Platform_OSIP
[3] https://www.move2space.de/MOVE-II/

using the same models causes inefficiencies due to duplication of work, additional efforts in communication, increase in complexity and inconsistencies in data and documentation. The knowledge transfer to successor projects is interpersonal rather than model based. Time and money for (new) model designs are invested repeatedly in each project. Using a model-based approach along the whole lifecycle is hard to judge and evaluate from a business perspective, since business/ management is too far away from core systems engineering.

Our working hypothesis: By integrating the engineering system model (MBSE system model) with later phases, e.g. systems integration/ test /operations data, the identification and elimination of (functional) failures in the context of the AIT process, but also in operations, could be significantly accelerated, simplified and optimised. This is additionally supported by suitable visualisations and manifests itself in less system failure, faster troubleshooting, and improved communication between the project participants.

Solution to the challenges

In order to address the state of the art challenges and to provide a proof of concept for the working hypothesis, an Integrated MBSE Analytics Platform (IMAP) demonstrator was developed. Figure 1 gives an overview of the platform and the supported workflow.

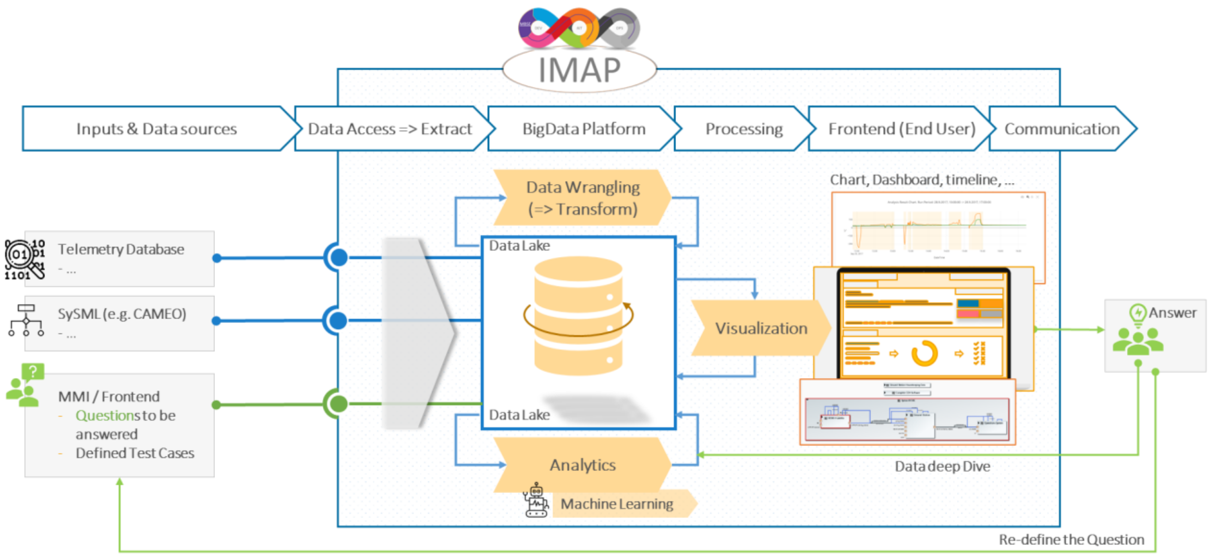

Figure 1 IMAP - workflow from Question to Answer

SysML models and AIT/telemetry data can be imported and stored in a Data Lake. Analytics based on MachineLearning can then be executed on the data and the results are shown in a web based frontend.

With the MOVE-II data as the main application case, it was shown that the gap between systems design models and test/operation data can be closed in an integrated manner. Holistic analyses for fault identification, referencing the system model in the background, are possible and can be well supported by Machine Learning. After initial training, analyses become increasingly better/meaningful due to learning effects with MachineLearning.

Furthermore, the platform can be scaled to larger missions and datasets. This is facilitated by a BigData capable base infrastructure as well as a separation of hard-coded core functions and flexible analytics functions via Jupyter Notebooks, which drive end-user compliant web visualizations.

Platform characteristics and implementation

IMAP consists of a set of integrated tools, services and web front ends to ingest both MBSE data (in the form of SysML models) as well as Telemetry definition and sensor data. The imported SysML model is mapped[4] to an RDF (Resource Description Framework) representation and stored in the platform's DataLake/Hadoop file system. This BigData scalable file system also stores CSV files as created by the Telemetry Importer. Next, a semi-automated tracing is performed which algorithmically connects TM/TC sensors to SysML flow items by relying on naming conventions and product structure aspects being reflected in the TM/TC database. An additional manual tracing capability is provided in order trace arbitrary sensors.

Figure 2 shows the Trace UI with all traces between Telemetry Sensors and SysML Flow Items. Missing traces can be added by drag&drop between the upper two tables. All traces are also persisted in the platform's DataLake in RDF representation.

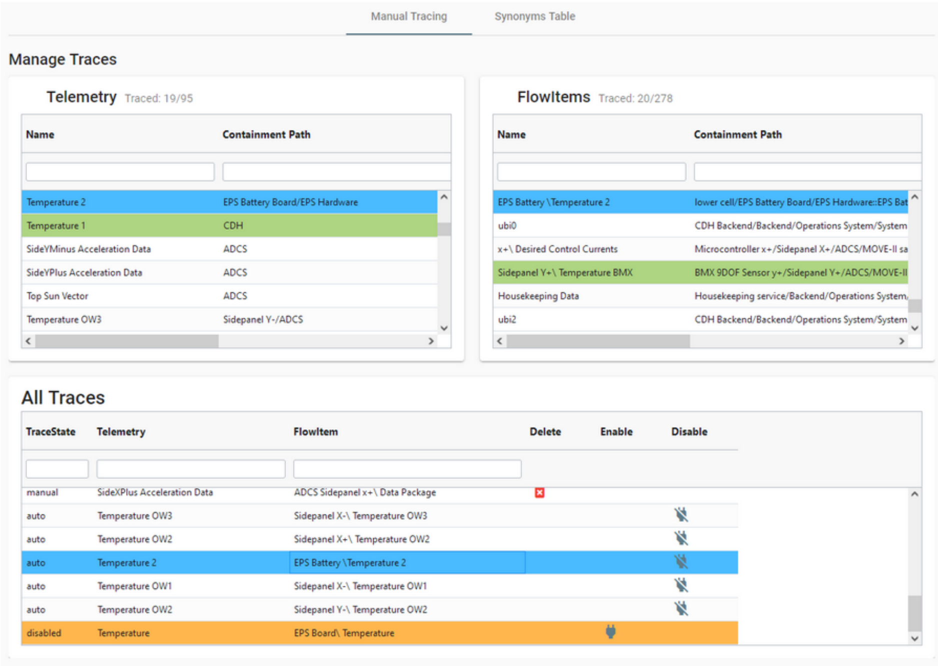

Figure 2 Trace UI

Once the import and tracing is completed, expert users can create a set of Analysis scripts to detect anomalies in the imported sensor data CSVs without having to deal with TM/TC databases and protocols. Although these Analysis scripts can implement a certain pattern or MachineLearning approach (e.g. LSTM[5] for the platform's demonstration case), every mission, model and kind of telemetry data will have their own unique characteristics metamodel. In line with a general DataLake approach ("Collect and store all data first, take care about use cases and data processing later"), these Analysis scripts are therefore not hardcoded as platform services, but are implemented in Python in Jupyter[6] notebooks (which is the de-facto standard for data analytics).

[4] IMAP uses an OWL based Ontology derived from the CIP datamodel, eventually to be aligned with the SpaceOntology

[5] https://en.wikipedia.org/wiki/Long_short-term_memory

[6] https://jupyter.org/

Obviously, dealing with Jupyter Notebooks requires a kind-of expert level understanding, typically provided by a domain expert/data analyst. To allow easy access to non-expert end users, the platform provides a set of front-end components which retrieve their data either directly from the output of a Jupyter based Analysis script or the integrated and linked data in the DataLake. That way, the complexity of data provision and preparation remains hidden in the background. A non-expert user can thus concentrate on the analysis task/error elimination exclusively.

Figure 3 shows a consolidated view by which Analysis results can be seen in the context of a selected requirement. Clicking on these results opens a chart view (Figure 4), which shows the detected anomalies on the measured sensors values which have been traced back to the SysML model as part of the platform's import function.

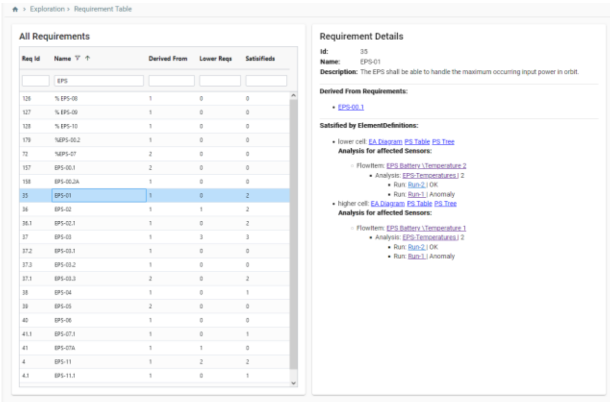

Figure 3 - Requirements to Analysis Tracing

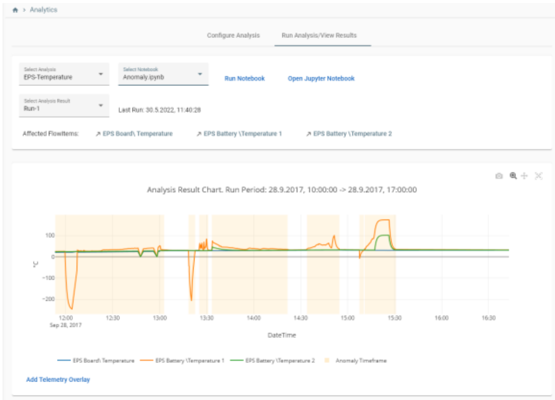

Figure 4 - MachineLearning based anomaly detection results

Ultimately, this approach provides the capability to link and overlay MachineLearning based Analysis Results to well-established SysML block diagrams as shown in Figure 5.

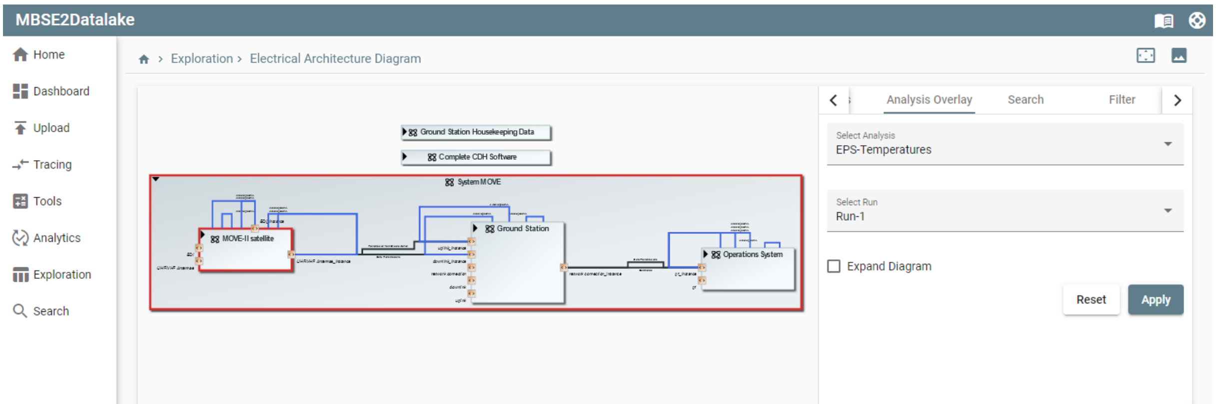

Figure 5 - SysML Analysis Overlay

Those block diagrams are generated and autolayouted from the imported MBSE data automatically. In case of a detected anomaly by an Analysis script, the affected blocks and parent blocks are rendered in red. Parent blocks can optionally collapsed and in this case, they are rendered in red as well. These collapsed blocks can be expanded incrementally, in order to identify the root cause of a problem from a system model perspective.

That way, the time-consuming process of manually referencing from a sensor anomaly to the bigger system picture is fully automated and significantly more efficient. Last but not least, this approach exposes and leverages the MBSE system models view into the AIT and possibly operational phase, where it can continue to serve as a central communication means, thus protecting/returning the original investment of creating these models.

In addition, the platform provides a search interface for all artefacts stored in the DataLake and an integration with Neo4J, which enables further graph-based analysis capabilities. Following the above approach, these graph analyses are also maintained in Jupyter Notebooks, including navigations into the Neo4J Graph browser.

The overall architecture of the platform is influenced by the BigData Europe project[7], it is distributed as a set of Docker images and can basically be deployed from a single docker compose file. In the future, this stack can be scaled up into a Kubernetes cluster where different services are deployed on different hardware resources.

In a bigger context, we see the platform as an (initially) read-only consumer of an MBSE Hub, providing a broader picture and analysis capability. Eventually, results being produced by the platform could be streamed back to a MBSE Hub however.

Benefits provided by IMAP

Using the MOVE-II project as a reference, an in-depth analysis of the procedures, deficiencies and lessons-learned of the traditional development approach has been performed. In addition, a potential future project (MOVE-III), using an application of the IMAP platform, has been compared against. As a result, the following key improvements have been identified and confirmed by the development team members:

- significant reduction of time for searching/gathering data
- improved communication across teams, based on visualizations and models
- automated tests on available data, specified by experts, proceeded by test engineers
- end to end data browsing, independent from entry point
- faster onboarding of people due to easy-to-use frontends
- completely new analysis capabilities based on MachineLearning

Due to the fact that no formal criteria for efficiency measurement and performance gain have been established yet, only estimates, based on the above challenges and before and after analysis have been derived. The overall efficiency gain (time decrease, cost decrease, quality increase) is estimated to be in the range of 30%-40%. This may be degraded slightly, depending on the effort to be spent on initial MachineLearning training (supervised vs. unsupervised) in relation to the confidence level of the resulting prediction models.

[7] https://github.com/big-data-europe