DOCUMENT

Executive Summary Report

MoBaTe – Enabling Model-Based Testing and Automated Test Case Generation for Ground Segment Data Systems

Prepared by	**MoBaTe team**
Reference	**77-058_MoBaTe_FR_FinalReport**
Issue	**1**
Revision	**0**
Date of Issue	**2022-12-14**
Status	**Final**
Document Type	**TN**

European Space Agency
Agence spatiale européenne

Distribution

APPROVAL

Title Final Report	
Issue 1	**Revision 0**
Author MoBaTe team	**Date 2022-12-14**
Approved by	**Date**

CHANGE LOG

Reason for change	Issue	Revision	Date
Initial draft	0	1	2022-12-01
Version for review	0	2	2022-12-06
Version for AR	1	0	2022-12-14

CHANGE RECORD

Issue	Revision		
Reason for change	**Date**	**Pages**	**Paragraph(s)**

Table of contents:

1 INTRODUCTION 4
1.1 Purpose 4
1.2 Scope 4
1.3 Document Overview 4
2 APPLICABLE AND REFERENCE DOCUMENTS 5
2.1 Applicable Documents 5
2.2 Reference Documents 5
3 TERMS, DEFINITIONS AND ABBREVIATED TERMS 6
3.1 Acronyms 6
3.2 Definition of Terms 6
4 EXECUTIVE SUMMARY 7
4.1 Overview 7
4.2 Background 7
4.3 Objectives 7
4.4 Analysis and Strategy Definition 8
4.5 Prototype Development 10
4.6 Validation and Training 11
4.7 Results and way forward 12

1 INTRODUCTION

1.1 Purpose

This document presents the Executive Summary Report of the "MoBaTe – Enabling Model-Based Testing and Automated Test Case Generation for Ground Segment Data Systems". It presents a concise version of the activity performed, results obtained and recommendations and ideas for future development.

1.2 Scope

This document constitutes deliverable ESR, linked to the outcome of "WP 400– Management". It covers the entire scope of the activity, providing a concise description of all work performed, giving an overview of the context of the activity and the achieved results.

1.3 Document Overview

Section 1 - Introduction (this section) provides the purpose, scope, and this document's overview.
Section 2 - Applicable and Reference Documents provides the list of reference documents.
Section 3 - Terms, Definitions and Abbreviated Terms provides a list of acronyms and terms used throughout this document.
Section 4 **- Executive Summary** – presents the summary and the results of the activity.

2 APPLICABLE AND REFERENCE DOCUMENTS

2.1 Applicable Documents

Ref.	Document Title	Issue and Revision, Date
[AD-1]	Technical Proposal - Enabling model-based testing and automated test case generation for ground segment data systems (MoBaTe) EXPRO+	1.0, 21.07.2021
[AD-2]	Statement of Work - Enabling model-based testing and automated test case generation for ground segment data systems, EGOS-GEN-MBTA-SOW-1001	1.3, 29.04.2021

2.2 Reference Documents

Ref.	Document Title	Issue and Revision, Date
[RD-1]	Formalized Definition of Application Contracts to Support Test Driven Methodologies (FORDAC) – Final Report	V1.0, 08.11.2016
[RD-2]	ADVANCED DIGITAL GROUND SEGMENT ENGINEERING - TECHNICAL NOTE ON GS ENGINEERING PROCESS ANALYSIS	V1.2, 28.05.2021
[RD-3]	ADVANCED DIGITAL GROUND SEGMENT ENGINEERING - TECHNICAL NOTE ON DATA MODEL	V1.3, 09.07.2021
[RD-4]	ALBERT: A Lite BERT for self-supervised learning of language representations, Zhenzhong Lan et al.	ICLR conference, 2020
[RD-5]	Drain: An Online Log Parsing Approach with Fixed Depth Tree, Pinjia He et al.	IEEE 24th International Conference on Web Services, 2017
[RD-6]	TAMDEM – Software Design Document, Berend Semke	V2.1, 2017.06.16
[RD-7]	AtoS TEMMPO Designer (IDATG) Version 16.7 – Handbook Documentation oft he IDATG XML Format	January 2018
[RD-8]	MoBaTe - TN1 – Summary of different approaches	1.0, December 2022
[RD-9]	MoBaTe – TN2 – Strategy and architecture	1.0, December 2022
[RD-10]	MoBaTe – TN3 – Deployment and config	1.0, December 2022
[RD-11]	MoBaTe – TN4 – Software user manual	1.0, December 2022
[RD-12]	MoBaTe – TN5 – Detailed design	1.0, December 2022
[RD-13]	MoBaTe – TN6 – Prototype validation	1.0, December 2022
[RD-14]	OPEN Developer Guide (space-codev.org) https://open.space-codev.org/open-community/git/open/doc/src/open-sum/	

3 TERMS, DEFINITIONS AND ABBREVIATED TERMS

3.1 Acronyms

Acronyms	Description
ADGE	Advanced Digital Ground Segment Engineering
AI	Artificial Intelligence
ART	Automated Regression Testing
E2E	End-to-End
GSRF	Ground Segment Reference Facility
BERT	Bidirectional Encoder Representations from Transformers
EUD	Egos User Desktop
FORDAC	Formalised Definition of Application Contracts to support Test Driven Methodologies
GSEF	Ground Segment Engineering Framework
LDM	Logical Data Model
MBSE	Model-Based Systems Engineering
ML	Machine Learning
MMIT	Man-Machine Interface Testing
NLP	Natural Language Processing
PLGSE	Paperless Ground Segment Engineering
RCP	Rich Client Platform
RS	Recommender System
SoW	Statement of Work
SUT	System Under Test
SWT	Software Widget Toolkit
TTI	Test Tool Interface
UML	Unified Modelling Language
W2V	Word to Vector

3.2 Definition of Terms

Terms	Description
EUDART	ESOC software based on EUD that controls the execution of GSSW functional tests
OPEN-M	Software application based on the OPEN preparation framework for ESOC's Flight Control Teams.
MaLTa	Machine Learning for Test Automation (study)
TAMDEM	Demonstrator for Test Automation using Graphical User Interfaces (study)
laNGoSTA	New Ground Segment Test Automation (study)

4 EXECUTIVE SUMMARY

4.1 Overview

The "MoBaTe – Enabling Model-Based Testing and Automated Test Case Generation for Ground Segment Data Systems" activity started in December 2021 and was concluded, after 12 months, in December 2022.

The overall work logic of the activity was broken down into the following phases:

1. WP 100 - Analysis and Strategy Definition
2. WP 200 - Prototype Development
3. WP 300 - Validation and Training

The produced deliverables were the following:

- TN1 – Summary of Different Approaches
- TN2 – Strategy and Architecture
- TN3 – Deployment and Configuration
- TN4 – Software User Manual
- TN5 – Design of specific elements
- TN6 – Prototype Validation
- SW1 – OPEN-M based test modelling environment
- SW2 – Recommender system

4.2 Background

The driver behind the MoBaTe activity was to propose ways to improve the current software development and validation lifecycle using MBSE techniques to shift some of the system validation activities to the early stages of development. The roadmap to achieve this included identifying strategies, required model properties, modelling tools and languages needed to move from a requirements-based approach to a model-based approach. All this considering existing and established workflows and tools currently used by the Agency.

4.3 Objectives

The main objective of this activity was to identify possible ways to extend existing test automation capabilities with AI supported recommendation features and an integrated test modelling environment based on MBSE without disrupting the existing workflow.

The focus was placed on:

- Model-based testing
- (semi) automated test case generation

4.4 Analysis and Strategy Definition

The first phase of the activity focused on analysing the landscape of previous and ongoing ESA studies to identify the reusability potential, and the strengths and shortcomings that could be relevant to or addressed by the MoBaTe prototype if in-line with the study objectives. The following studies addressing MBSE frameworks, Machine Learning and Natural Language Processing applied to test case generation were evaluated

- FORDAC
- GSEF
- ADGE
- OPEN
- Capella
- ART-MMIT
- MaLTa
- NLP (state of the art)

From this analysis OPEN was selected as the candidate to form the baseline of the test modelling environment prototype. A strategy was devised to develop the prototype based on what OPEN already provides and what the user could benefit from in this new test modelling workflow (Se Figure 1). Additionally, a state-of-the-art analysis of Natural Language Processing techniques was performed to identify the best ways to develop a recommendation system to intelligently support the user in the modelling of test cases.

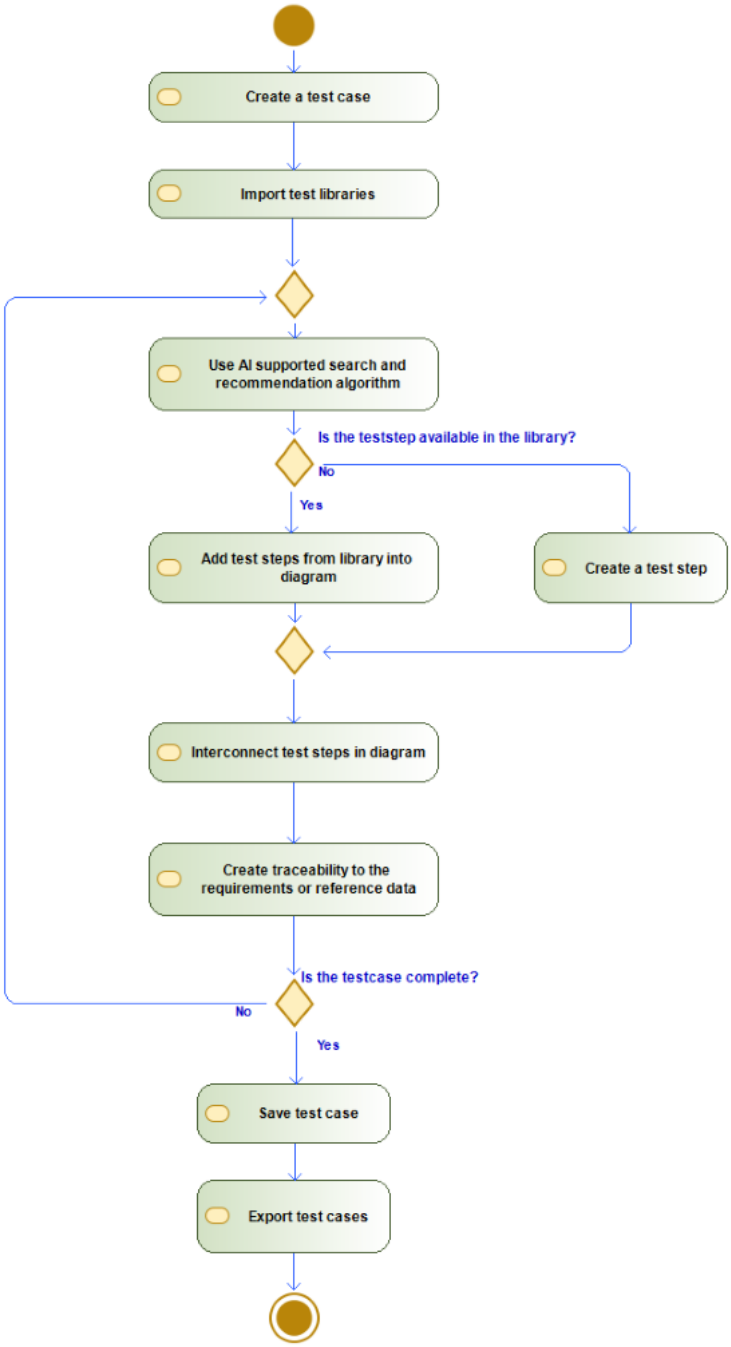

Figure 1: New test modelling workflow

4.5 Prototype Development

The development of the prototype followed an Agile approach to deliver the most functionality possible in the short project duration. The prototype was designed with four main components

- A test modelling environment (TME)
- An AI recommendation system (RS)
- Import feature
- Export feature

The test modelling environment was developed on top of the OPEN preparation framework product OPEN-M. This choice allows for an interface to the EGS-CC tailoring data model and for some out-of-the-box functionalities that can be extended for the purpose of this activity, like the import of files (needed to import the test libraries) and the export of files (needed to export the produced test schedules).
To provide a test modelling environment, a new view/perspective was implemented based on Eclipse Sirius.

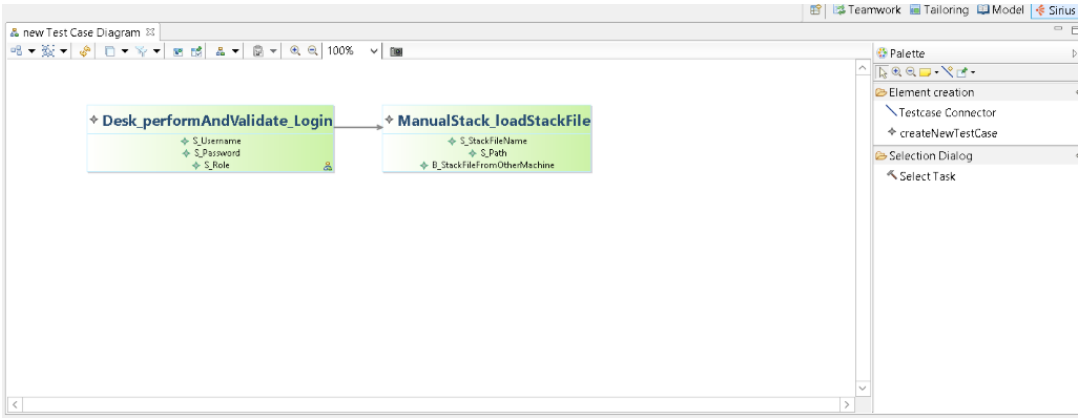

Figure 2: Sirius-based test modelling environment

The test modelling environment relies on a new meta model that defines the data types and the relationships of the entities. This metamodel was designed based on the TEMPPO XML format and structure because the data that is imported into the system is in this format. (Test libraries are TEMPPO XML files).

The TME implements three types of diagrams to help the user navigate the test case creation:

- Test Case Diagram
- Test Step Diagram
- Traceability Diagram

The Recommendation System component was designed and developed as a standalone application that is loosely coupled from the TME. It provides a GraphQL API that can be queried by the TME for test step recommendations. The RS component is composed of a search engine and a data processing unit and was delivered with two ML models – a contextuary model based on a text2vector approach and a pre-trained BERT model. Both models can be used and were used by the TME for evaluation. An arbitrary number of models can be added to the prototype.
The RS was integrated into the TME UI through a search window that can use a prefix "AI:" to query the RS.

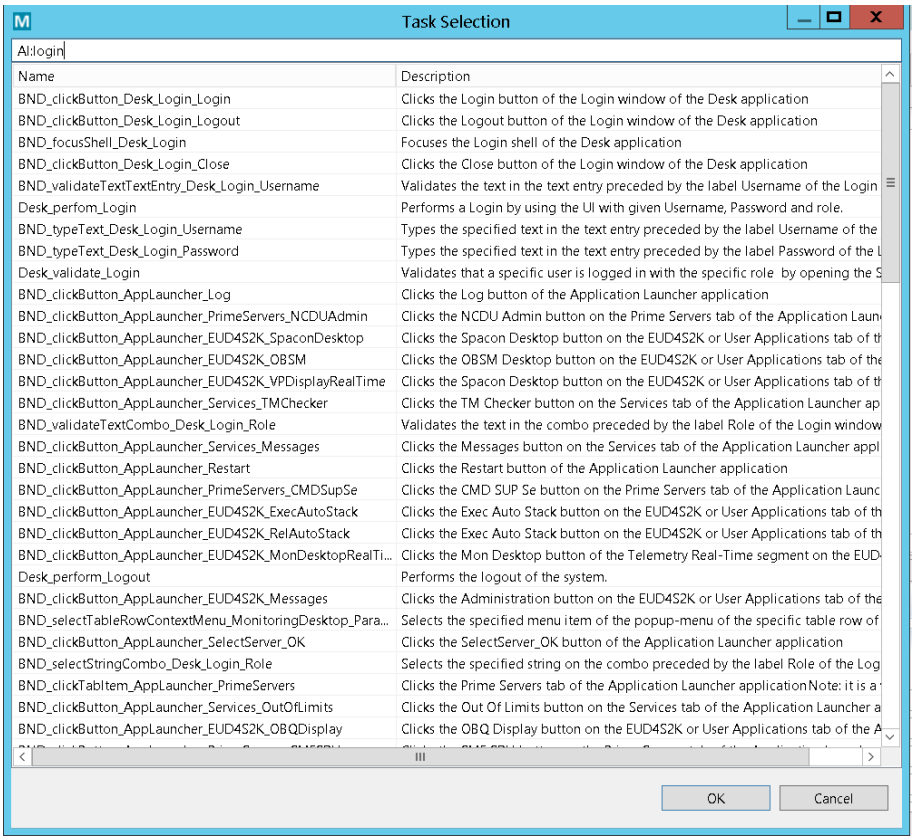

Figure 3: AI-based test step search

4.6 Validation and Training

The validation of the prototype was performed in collaboration with users experienced in the creation of test schedules in the EDLab environment. As a validation scenario, a tentative of recreating the TC BASIC Sentinels5P test schedule was performed. For this a set of search queries (in natural language) was defined by the users to search for each test step of the test case. An example of such a query is "add command with parameter" to find the test step "EditReleasing_addCommandWithParameterDetail". The focus of the validation was to evaluate which kind of test steps would be recommended by the RS. This

validation scenario was run using both ML models and then the results were compared to each other.

The outcome of the validation showed that the BERT model performed significantly better than the contextual text2vector model in this specific scenario. In more than 90% of the test steps used in the validation, the correct test step was proposed in the first or a higher position in the list by the BERT model compared to the contextual model. In several cases, the contextual model failed to return the desired test step as part of the suggested test steps. Both models have the potential to be fine-tuned to better fit the problem at hand.

4.7 Results and way forward

The MoBaTe activity aimed at developing a prototype system that has a higher level of (artificial) intelligence compared to the current systems and workflow. The prototype demonstrates the potential of adding MBSE and AI techniques to the current validation workflow at ESOC. A lot can be gained as well from linking the validation preparation environment to the systems engineering environment. (Linking to spacecraft data, perform traceability from the validation to the requirements, …)

The implemented AI, in the form of a recommender system, shows great potential to support a user with the creation of tests. The possibility to search for test steps using natural language and receive a pre-selection (recommendation) of potentially relevant test steps could simplify the validation workflow and reduce efforts greatly. To make the recommendation more useful, much more data must be collected, user interactions and decisions could be taken into account and the models can be finetuned to fit the ESA corpus and objectives more closely.

Some of the initially planned features were not implemented due to several scope changes and time constraints but are still believed to be important. The following is a list of such features

- Accessing spacecraft tailoring data
- Requirements traceability diagram
- Import feature
- Export feature
- Configuration management
- Finetuning of ML models