Ref. Ares(2025)2069276 - 14/03/2025

DIMAT MODELING AND DESIGN SUITE

14/03/2025

Grant Agreement No.: 101091496
Call: HORIZON-CL4-2022-RESILIENCE-01
Topic: HORIZON-CL4-2022-RESILIENCE-01-25
Type of action: HORIZON Innovation Actions

D5.1 DIMAT MODELLING AND DESIGN SUITE

Grant agreement number	101091496	Acronym	DiMAT
Full title	Digital Modelling and Simulation for Design, Processing and Manufacturing of Advanced Materials		
Start date	01/01/2023	Duration	36 months
Project url	[HTTPS://CORDIS.EUROPA.EU/PROJECT/ID/101091496](https://cordis.europa.eu/project/id/101091496)		
Work package	WP 5 - BUILD: DiMAT Modelling and Design		
Deliverable	D5.1 – DiMAT Modelling and Design Suite		
Task	T5.1 – Multidisciplinary Materials Design Framework T5.2 – Materials Modelling and Behavior Prediction T5.3 – Materials Design, Characterization and Synthesis		
Due date	30/06/2024		
Submission date	14/03/2025		
Nature	OTHER	Dissemination level	Public
Deliverable lead	6-CETMA		
Version	1.1		
Authors	Leonardo Cosma (CETMA), Yoav Nahshon and Lukas Morand		

	(Fraunhofer IWM), Yannis Papadimitriou (CERTH)
Contributions	CETMA, Technical Partners, Toolkit Development Leaders
Reviewers	Leonardo Cosma (Cetma), Yoav Nahshon and Lukas Morand (Fraunhofer IWM), Yannis Papadimitriou (CERTH)
Abstract	This document presents the status of the implementation of the toolkits implemented within the DiMAT Modelling and Design Suite.
Keywords	MODELLING AND DESIGN, KNOWLEDGE SYSTEM, VIRTUAL TESTING

Document Revision History

Version	Date	Description of change	List of contributor(s)
0.1	01-May-2024	ToC	UPV, CERTH
0.2	07-Jun-2024	1st Draft available for internal review	CERTH, CETMA, IWM
0.3	13-Jun-2024	Internal review	DRAXIS, SUPSI, UPV
0.4	21-Jun-2024	2nd Draft addressing the comments from internal reviewers	CERTH, CETMA, IWM
1.0	28-Jun-2024	Final quality check and issue of final document	CERTH
1.1	14-Mar-2025	Update of deliverable according to PO/reviewers' comments and quality check	CERTH, CETMA, IWM

DISCLAIMER

COPYRIGHT NOTICE

EXECUTIVE SUMMARY

This document describes the implementation status of the first release of the toolkits that compose the DiMAT Modelling and Design Suite as of M18 of the DiMAT project. The Suite aims to help its users to design materials in terms of their internal structure and properties, analyzing all factors that influence their behaviour and optimizing their performances. The Suite is composed of three toolkits, the DiMAT Material Design Framework, the DiMAT Materials Modeler and the DiMAT Materials Designer.

The DiMAT Materials Design Framework is an ontology-based open knowledge system aimed at supporting the materials design process. In this release, a first version of its correlation's component is implemented as well as a basic ontology to be used as a material relation manager. Future development planned for the next releases will deal with the improvement of the correlation component and the development of a query application.

The DiMAT Materials Modeler is a computational toolkit that integrates advanced Machine Learning and AI technologies to predict the material behaviour in different conditions. In the first release, different features are implemented, such as missing data manipulation, exploratory data analysis, and an integrated Large Language Model (LLM). Future developments will include expanded capabilities for the LLM.

The DiMAT Materials Designer is a tool that executes virtual testing on complex materials, defined in terms of the base components and their internal microstructure, with the aim to calculate their actual properties. In the first release, the user has the possibility to calculate elastic properties of materials and store the obtained results locally. Plans for future release include the connection to other toolkits for data exchange and storage, as the possibility to calculate other kind of material properties.

The source code of all the three toolkits is available in the respective GitLab repository.

This deliverable is an updated version that addresses all comments provided by the reviewer and the Project Officer. To ensure clarity and transparency, we have created a detailed table summarizing the changes made to this document in response to the reviewer's feedback. The Table 1 can be found in the appendix, providing an overview of the revisions and additions for easy reference.

TABLE OF CONTENTS

Co-funded by the European Union

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

AI	Artificial Intelligence
CPU	Central Processing Unit
EDA	Exploratory Data Analysis
FEM	Finite Element Method
GB	Gigabyte
GPU	Graphics Processing Unit
GUI	Graphical User Interface
KNN	K-Nearest Neighbors
LLM	Large Language Models
MICE	Multivariate Imputation by Chained Equations
ML	Machine Learning
NB	Naïve Bayes Algorithm
RAG	Retrieval-Augmented Generation
RAM	Random Access Memory
SVM	Support Vector Machine

DiMAT Toolkits

CMDB	Cloud Materials Database
KAF	Knowledge Acquisition Framework
MEC-LCA	Materials Environmental and Cost Life Cycle Assessment
MDF	Materials Design Framework
MM	Materials Modeler
MD	Materials Designer

Co-funded by the European Union

MMS	Materials Mechanical Properties Simulator
MPS	Materials Processing Simulator
DTPC	Digital Twin for Process Control

1 DIMAT MATERIALS DESIGN FRAMEWORK – DI^{MDF}

1.1 OVERVIEW

The Di^{MDF} is an ontology-based open knowledge system to support the material design process. It provides three components that together aim to support users in finding information (exploring dataspaces and potentially other DIMAT toolkits) and generating new knowledge based on the available information (e.g., find correlations using machine learning and data-science tools and gain new insights on this basis). Di^{MDF} in its first version is running on DiMAT Di^{CMDB} that provides data and information from specific application cases. In addition, Di^{MDF} provides knowledge from the materials modelling domain relevant to the material design process. In this regard, Di^{MDF} summarizes and provides easy access to materials modelling knowledge, which is constantly growing and, therefore, challenging to overview, especially for users with other backgrounds than materials modelling. The information from the materials modelling domain, however, is particularly valuable for understanding the behaviour of materials and, therefore, for the materials design process.

Specifically, the three components the Di^{MDF} consists of are: The exploration component that supports the user to better find information in Di^{CMDB} and potentially in other toolkits, the materials relation manager that provides engineers with knowledge from the materials modelling tasks, and a correlation component that aims a supporting the engineers in correlating information stored in the Di^{CMDB} toolkit.

1.2 FEATURES

The following features are included:

- Exploration component: Query Di^{MDF} for data/ information contained in Di^{CMDB} and potentially in other DiMAT toolkits. It is designed so that also complex queries are supported.
- Materials relations manager: Application that provides an overview of existing materials relations from the materials modelling domain to support the materials design process by providing knowledge about the behaviour of materials and how they are modelled. The knowledge is stored using a knowledge graph in Di^{MDF} or Di^{CMDB} alternatively.
- Correlations component: Application that identifies and visualizes correlations in data, for example the one stored in the Di^{CMDB}. This component specifically aims at providing

state-of-the-art data science approaches tailored to any type of data available in the Di^{CMDB} toolkit.

1.3 TECHNICAL SPECIFICATIONS

- Provides a friendly graphical user interface (GUI) based on Angular.
- Runs in Docker containers, ensuring compatibility across different computing environments.
- Requires an internet connection for updates and authentication via Keycloak.
- Minimum hardware Requirements: 1-core CPU, 8GB RAM.

1.4 IMPLEMENTATION STATUS

1.4.1 Current Implementation

The Di^{MDF} components are still under development. The current implementation of Di^{MDF} includes the following:

- First version of the correlation's component in the form of a Jupyter Notebook that operates on data stored in the Di^{CMDB}. This notebook will be embedded in a user-friendly interface (Angular based) in a way that enables users without programming skills to interact with it. The Jupiter environment is provided by Di^{CMDB}, allowing python programmers to directly integrate their code and make is usable for other users. The correlation component will be based on state-of-the-art statistics/ machine learning methods, such as for the calculation of correlation matrices or for clustering. Methods are not only aimed to be applied to individual datasets but also to compare different datasets in the Di^{CMDB}.
- Basic ontologies that can be used for the material relation manager. These are the continuum modelling ontology and the simulation ontology developed within the EU project marketplace.

Link to the GitLab repository of the solution:

https://gitlab.cc-asp.fraunhofer.de/dimat/modelling-and-design-suite/materials-design-framework

Link to demo video:

https://www.youtube.com/watch?v=BeIXAzhYkmY&ab_channel=DiMAT

Co-funded by the European Union

1.4.2 Next developments

The future development and implementation of Di^{MDF} will include the following (all of high priority):

- Improved version of correlation component embedded in a Di^{MDF} frontend
- Materials relation manager with first test case
- Exploration component
- General Di^{MDF} frontend, from which all the three components are accessible.

These features are expected to be implemented by the next release of the Suite.

2 DIMAT MATERIALS MODELER – DI^{MM}

2.1 OVERVIEW

DiMAT Materials Modeler (Di^{MM}) is an innovative computational toolkit tailored to enhance the prediction and analysis of material properties in manufacturing. This solution integrates advanced machine learning and AI technologies, offering predictions based on the material behaviour under various conditions. Di^{MM} stands out with its user-friendly interface, simplifying complex data analysis and making high-level data science accessible even to non-experts. The toolkit not only automates the selection of optimal algorithms and hyperparameters but also includes a local open-source large language model for real-time, context-aware insights, enhancing therefore the interpretability of results.

Di^{MM} is defined as a dynamic, data-driven environment that aids in modelling material properties influenced by microstructural and raw materials attributes and processing conditions. By accelerating the discovery and application of new material insights through digital means, Di^{MM} aims to assist material design and production processes. Its primary objective is to democratize access to sophisticated data analysis, enabling professionals at various expertise levels to extract actionable insights from complex data sets efficiently. Ultimately, Di^{MM} strives to enhance industrial innovation, providing the tools necessary to design superior materials more efficiently and cost-effectively, thereby promoting a shift towards more advanced manufacturing paradigms.

2.2 FEATURES

Di^{MM} offers advanced capabilities for manipulating, analyzing, and predicting material properties. Key features include:

- **Data Manipulation:** Users can select methods like KNN or MICE for handling missing data, along with setting a threshold for missing data percentage beyond which rows are discarded.
- **Exploratory Data Analysis (EDA):** Includes correlation calculation and presentation, providing insights into the relationships between variables.
- **Feature Importance Calculation:** Allows analysis at both individual feature levels and grouped features (materials raw parameters and process parameters).
- **Materials Property Prediction Model Development:**
  - Dynamically selects between classification or regression models based on the target parameters.

    - Automatically selects from a range of algorithms (e.g., Random Forest, SVM, ElasticNet) based on performance, including K-fold validation and hyperparameter tuning.
    - Supports both Univariate and Multivariate predictions, providing metrics on model performance.
- **Local Open-Source Large Language Model Integration:** Users can query a local Large Language Model (LLM) to gain insights and answers directly related to their data.

2.3 TECHNICAL SPECIFICATIONS

Hardware Requirements:

- Minimum: 4-core CPU, 16GB RAM (no LLM functionality).
- Recommended: 8-core CPU, 32GB RAM, and an Nvidia GPU with 16GB VRAM.

Software Requirements:

- Fully developed in Python (both front and back-end). Runs in a Docker container, ensuring compatibility across different computing environments.
- Requires an internet connection for updates and authentication via Keycloak.

2.4 IMPLEMENTATION STATUS

2.4.1 Current Implementation

The implementation of Di^{MM} is currently in the testing phase with all primary features implemented and undergoing evaluation to ensure reliability and accuracy before deployment. It includes the following:

- Missing Data Manipulation:
    - Selection of methods, such as KNN or MICE, for handling missing data.
    - Ability to set a threshold for the missing percentage of data, discarding rows that exceed this threshold.
- Exploratory Data Analysis (EDA):
    - Calculation and presentation of correlations between variables to understand their relationships. Using open-source libraries like scikit-learn and matplotlib, the correlations of the materials and process parameters present in the data are calculated and visualized.

- Feature Importance Calculation:
  - Analysis of individual features as well as grouped features (materials parameters and raw parameters).
- Materials Property Prediction Model Development:
  - Dynamic selection between classification or regression models based on target parameters.
  - Automatic selection from a wide range of algorithms including Random Forest, SVM, ElasticNet, KNN, Gaussian NB, Extra Trees, XGBoost, based on their performance.
  - K-fold validation to select the best hyperparameters.
  - Ability to select between Univariate and Multivariate predictions when applicable, with detailed metric presentation to evaluate model performance.
  - Presentation of the final model and its hyperparameters.
- Local Open-Source Large Language Model Integration:
  - Capability to ask questions directly related to the data provided, enhancing user understanding and interaction with the data. Please note that the choice of the open-source LLM is left to the users, who can select and integrate the API of their preference based on their specific needs and use cases. The system is designed to be flexible, allowing compatibility with various LLM providers to accommodate different performance, licensing, and deployment requirements.

Link to the GitLab repository of the solution:

https://gitlab.cc-asp.fraunhofer.de/dimat/modelling-and-design-suite/materials-modeler/mm

Link to demo video:

https://www.youtube.com/watch?v=GeNiFAALJME&t=135s&ab_channel=DiMAT

2.4.2 Next developments

Based on user feedback, further enhancements will be made to the existing features. The future development and implementation of Di^{MM} will include (all of same priority):

- Possible enhancements of the LLM pipeline and expanded capabilities for the LLM, aiming to provide even more sophisticated tools and insights to users.
- Also based on the user feedback, features such as inverse search and solution landscape investigation will be explored, leveraging on the Machine Learning model (ML model) developed in-app to predict ranges of parameters for desired outcomes.

These features are expected to be implemented by the next release of the Suite.

3 DIMAT MATERIALS DESIGNER – DI^{MD}

3.1 OVERVIEW

The DiMAT Materials Designer (Di^{MD}) is a tool that provides the user the capability to define complex materials, like composite materials, in terms of their base component and internal microstructure and obtain the mechanical properties of the new material.

Properties are calculated via a virtual testing approach, which consists of building a Finite Element Method (FEM) model and solving its equations to extract the requested properties from the results.

3.2 FEATURES

The features of the Di^{MD} are the following:

- Material definition: the toolkit provides a web-based interface for defining and modifying composite material, starting from the base component materials to the internal microstructure and layup arrangement of the composite.
- Properties calculation: the toolkit is connected to a calculation server which executes virtual testing of the defined material and calculates its elastic properties.
- Data visualization: the user can select one or more materials and visually compare their properties and download text files with all available information for local use.
- Data storage: all data related to the material definition and their calculated properties are stored on the server running the toolkit, so it can be used stand-alone without relying on other storage solutions.
- Access management: materials definitions and calculations results are shared, by default, with working groups to allow different users to have access to the same data if the toolkit administrator gives them access.

3.3 TECHNICAL SPECIFICATIONS

The Di^{MD} is implemented as several Docker containers each delivering specific functionalities:

- The frontend, based on Nicegui framework.
- The internal database, based on PostgreSQL.

- The Api server and the scheduler/interface with the calculation code, implemented using the Fastapi framework.

Alle the containers are orchestrated using Docker compose.

The calculation code, responsible for executing the actual calculations, is Hexagon Digimat-FE® and is installed and licensed separately from the Di^{MD}.

The Di^{MD} is a web-based tool, so it needs a working internet connection to be reached by users and to communicate with the calculation code. Moreover, authentication is managed via a Keycloak server running externally and maintained by DiMAT partners.

3.4 IMPLEMENTATION STATUS

3.4.1 Current Implementation

The current implementation of Di^{MD} includes the following:

- User interface.

The user interface serves as the primary access point to the toolkit's functionalities. All operations supported by the toolkit, including definitions, modifications, deletions, calculation requests, and result analysis, can be conveniently executed through the interface. As illustrated in Figure 1, the main toolbar allows users to select between working with mechanical or thermal properties (even if thermal properties calculation is not yet supported in this version), as well as choosing the appropriate working group, while main functionalities can be access from the left sidebar.

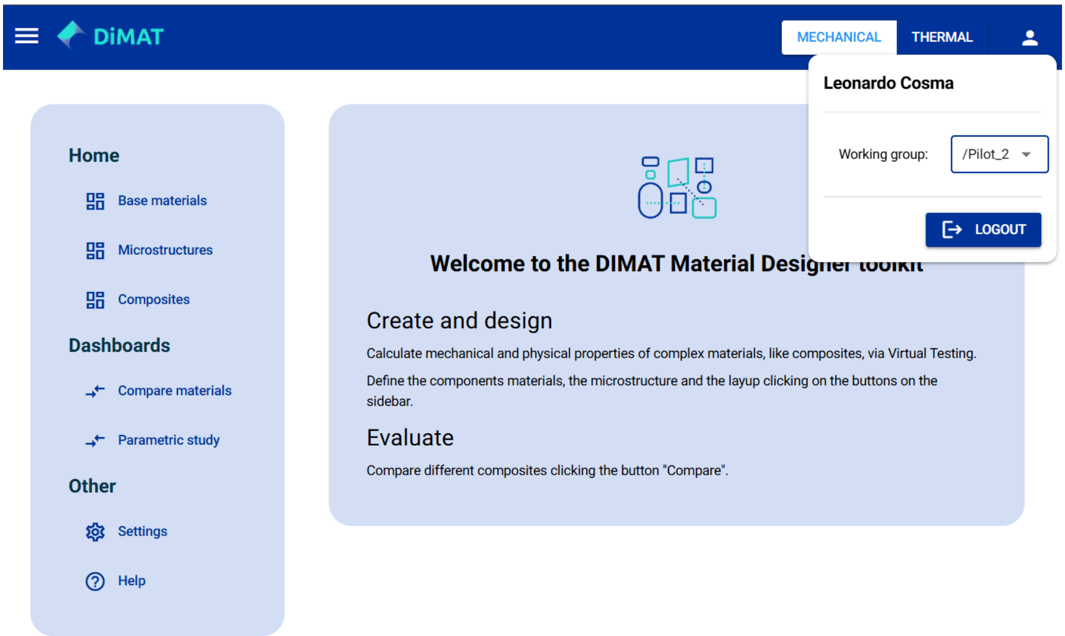

Figure 1: Home page of Di^{MD} toolkit

- Material definition and storage.

Base materials, microstructures and complex materials can be defined in the interface, listed, modified and deleted. As an example, composite page is shown in Figure 2.

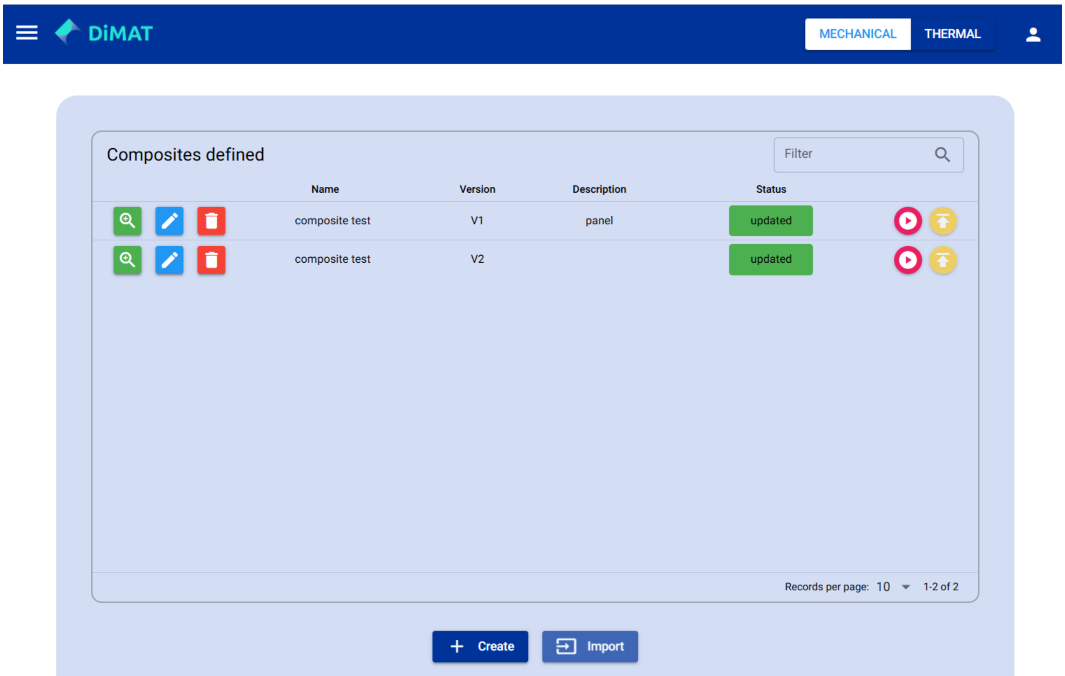

Figure 2: Composite page of Di^{MD} toolkit

- Material elastic properties calculation.

The toolkit can calculate elastic properties for different types of microstructures. Currently only long fibres are supported, but more types will be supported by future releases of the toolkit. Calculated properties are shown in the relative page, as shown in Figure 3.

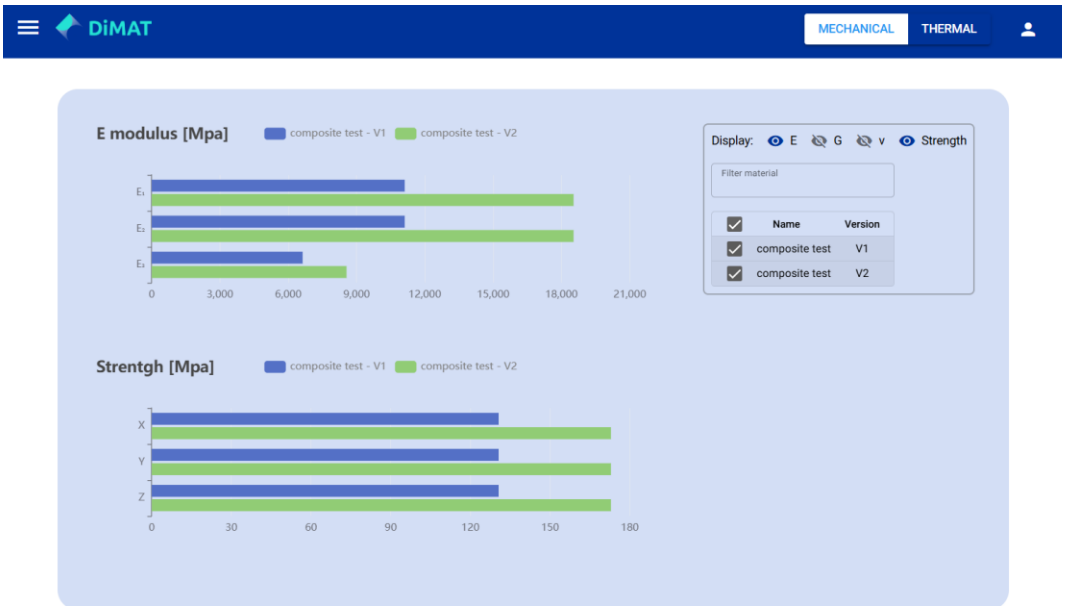

Figure 3: Results page of Di^{MD} toolkit

- Local (server-based) data storage.

All data, including user inputs and results calculated by the toolkit, are stored by a dedicated database service. Communication between the main server and the database is implemented asynchronously to ensure optimal performance of the application backend.

- Containerization and orchestration of the toolkit components.

As specified in "Technical specification" section, the solution is implemented in several Docker containers orchestrated via Docker Compose. The toolkit is deployed and accessible for users at the following link: http://md.dimat-tools.eu.

Link to demo video:

https://www.youtube.com/watch?v=Ir73SxeZFLQ

Link to the GitLab repository of the solution:

https://gitlab.cc-asp.fraunhofer.de/dimat/modelling-and-design-suite/materials-designer/md

3.4.2 Next developments

The future development of the Di^{MD} will be guided by the users' feedback during Pilot evaluation, especially regarding bugfixes or improvements of user interface usability.

Planned future developments and implementations are:

1. Modelling of different types of microstructures based on Pilots requirements.
2. Implementation of a connection with the Di^{CMDB}, to import and export data.
3. Implementation of a parametric study feature into the toolkit user interface.
4. Implementation of a Software Developer Kit that could be used by other toolkits to request data and calculation to the Di^{MD}.

The most important items of this list are items 1 and 2, since they are directly related to the fulfilment of the requirements of the DiMAT Pilots: for this reason, their implementation will have the maximum priority. Items 3 and 4 are certainly important to improve the usability and versatility of the toolkit, but are not essential to fulfilment of Pilot's requirements, so their implementation will follow that of the first two items. However, all features are expected to be fully implemented by the next release of the Suite.

Co-funded by the European Union

APPENDIX A

Nº	Comment	Action	Paragraphs
1	"However, there are few details on the implementation status. As the links in the document are directing for private sites it is hard to have a real understanding of the state of implementation. This should be corrected either by including some screenshots in the document or include links for videos in a public space."	We have added more info about the implementation in the sections "Implementation status". In the same sections we have also added screenshots and links to youtube videos. Access to GitLab repository will be given to the reviewers.	1.4.1 2.4.1 3.4.1
2	"The next developments are presented with very few details. In this regard, the preparation of a timeplan is beneficial as well as the justification for the prioritisation choices being followed in the developments"	We have added more details about future developments, prioritisation and time plan	1.4.2 2.4.2 3.4.2
3	"The mechanisms used to identify and visualise data correlation should be better detailed as well as which open source LLM are being used"	We added details about data correlations and open source LLMs being used	2.4.1

Table 1: Changes in the deliverable based on the reviewers' and Project Officer's comments