



MED1stMR
Mixed Reality Training

TRAIN
[SKILLS.
RESILIENCE.
PERFORMANCE]
SAVE LIVES

D1.3

Data Management Plan

Version
V1.0

Authors

Jaison Puthenkalam (AIT)
Helmut Schrom-Feiertag (AIT)

Project
MED1stMR

Project number
101021775

Type of Action
RIA

Start date of project
01.06.2021

Duration
36 months

Deliverable
D1.3

Deliverable Lead
AIT

Related work package
WP1

Dissemination level
Public

Due submission date
30.11.2021

Actual submission
30.11.2021

Versions

Version	Date	Author(s)	Description
V0.1	01/10/2021	Jaison Puthenkalam (AIT)	Document Structure
V0.2	02/11/2021	Jaison Puthenkalam (AIT)	Draft Version, shared with consortium for additional inputs & feedback
V0.3	24/11/2021	Jaison Puthenkalam (AIT)	Included all feedback from consortium partners and AIT DPO
V1.0	25/11/2021	Helmut Schrom-Feiertag (AIT)	Final revision, content check and quality assurance
V1.0	30/11/2021	Vendula Rajdlova (AIT)	Final formatting and upload

Report Review

Version	Date	Reviewer(s)	Statement
V0.2	19/11/2021	Quynh Nguyen (AIT) Rafael Wespi (UBERN) Yannick Hill (UHEI)	Document review with minor changes, additions, notes and approvals.
V1.0	26/11/2021	Markus Murtinger (AIT)	Final review

List of Acronyms and Abbreviations

Acronym/ Abbreviation	
DPO	Data Protection Officer
EPME	Effective Performance in Medical Emergencies
MFR	Medical First Responder
MR	Mixed Reality
MS	Microsoft
Obj.	Objective
PSS	Protected Sharepoint Server
SOTA	State of the Art
VR	Virtual Reality

Terms and Definitions

Term	
Protected Sharepoint Server (PSS)	A locally hosted instance of Sharepoint within AIT, with heightened security regulations and limited access, for more sensitive data (e.g. raw study data including personal data points)

Relation to Objectives


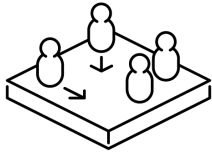
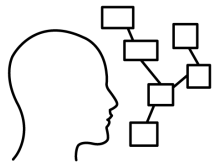
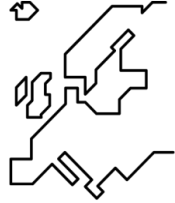
Objective	Description
	<p>Obj. 1: Pioneering MR training approach for enhanced realism</p> <p>Definition of necessary data types and formats to achieve project goals</p>
	<p>Obj. 2: Effective training scenarios and a training curriculum</p> <p>Description of the data management process during the assessment and development of training scenarios</p>
	<p>Obj. 3: Physiological signal and trainee behaviour feedback loop and smart scenario control</p> <p>Process for secure storage for sensitive data such as physiological parameters of trainees or test persons.</p>
	<p>Obj. 4: Position the pioneering MR training approach across Europe</p> <p>Framework for managing and storing contact details and similar personal data for exploitation and dissemination purposes</p>

Table of Contents

Executive Summary	3
1 Data Summary	4
2 Data Management	11
3 Allocation of Ressources & Data Security	13
4 Ethical Aspects	14
5 Conclusions	15

List of Figures

Figure 1: AIT Sharepoint folder structure based on work packages for MED1stMR	12
---	----

List of Tables

Table 1: The work and the document build on results from the following deliverables.	3
Table 2: The results of this work will be incorporated into following work and developments.....	3
Table 3: Data summary	5

Executive Summary

This Data Management Plan (DMP) describes the handling and lifecycle of data created and processed during the Med1stMR project. The structure of the deliverable is based on the DMP template provided by the European Commission and includes information on the types and purposes of the data collected, the measures taken in the project to allow for re-use of the project data (where applicable), and information about the protection of personal data acquired and managed during the project. It also describes ethical guidelines for the handling of more sensitive data points (e.g. personal data), and the security measures taken to protect them, such as saving them in a locally hosted Sharepoint Server with strictly limited access.

MED1stMR has opted into the Open Research Data Pilot for providing open access to research data, where possible. The open data repository “Zenodo” has been chosen to distribute the data sets that can be shared openly.

As this is a ‘living’ document, it will be adapted and enhanced as the project develops, to ensure that it remains up to date with any new developments.

Relation to other deliverables and tasks in MED1stMR

Table 1: The work and the document build on results from the following deliverables.

No.	Title	Information on which to build
D9.3	POPD – Requirement No. 3	General processes for protection of personal data

Table 2: The results of this work will be incorporated into following work and developments

No.	Title	Basis for
D2.2	End users Point of View: Requirements Report, Stakeholder Map and Expectation Summary for Smart Wearables, MR Training Framework and Curriculum	Data lifecycle for study data collected in WP2
D2.4	High-Level System Architecture	Data collected and processed in the project also include data generated through the use of the training system and its processing and storage. This must be taken into account in the system architecture.
D3.1	Guidelines and Inputs for the future Training Scenarios	Enhancement of SOTA findings with data collected from observations & interviews with trainers & trainees

-	Most of the other future deliverables dealing with developments, studies, data acquisitions and data analysis	As an overarching document, this deliverable defines the data management process for MED1stMR. Wherever data is collected, processed and stored. This includes all the data from studies, the data used for the developments and from the field studies.
---	---	--

1 Data Summary

Med1stMR aims to gather new knowledge about training processes for medical first responders, and to improve on existing training frameworks using a combination of VR and technologically advanced manikins. For this purpose, a variety of data needs to be collected and evaluated during the project. A summary of the data being generated, as well as their origin and purposes, can be found in Table 3 on the following pages.

Table 3: Data summary

Data types / category	Format	Origin / Source	Purpose & Relation to Obj.	Utility beyond project
Personal data, such as: <ul style="list-style-type: none"> • Contact information • Age • Sex • Role / work experience • Other relevant data points (e.g. education level) • Experience with VR/MR 	Word/Excel files, databases	Questionnaires, existing contacts and participant databases in the consortium	Coordination & selection of participants for studies, workshops and focus groups; contact information for longitudinal studies (all objectives, esp. obj. 2)	Data that can directly identify participants will not be shared outside the consortium and will be securely stored in the Protected Sharepoint Server (PSS). Only aggregated or pseudonymized data will be included in publications and public deliverables.
Results of SOTA analysis regarding <ul style="list-style-type: none"> • Disasters • Existing products and projects • Medical training practices and VR training systems • AR/VR/MR as a method for training of MFRs 	Word, PDF, databases	Online research, book reviews	Creation of a common understanding about the status quo; quality assurance and decision-making support for next project steps; preparation of the end-user workshops (obj. 2)	The results of the SOTA analysis are meant to be published as papers and will provide valuable aggregated information about the current status in the mentioned fields.

Data types / category	Format	Origin / Source	Purpose & Relation to Obj.	Utility beyond project
<p>Knowledge and information from MFRs on disasters, needs, training curricula, processes, etc.</p>	<p>CSV/XLS Exports, R-files, Word documents, photos / images of workshop results</p>	<p>Contextual interviews, focus groups, questionnaires, co-creation workshops</p>	<p>Investigation of the needs of MFRs & their involvement in the ideation process; creation of first-hand knowledge on current training processes for the EPME model (obj. 2)</p>	<p>To ensure that MFRs provide authentic feedback without holding back, individual data points will only be shared in the consortium. However, aggregated data will be used in publications so as to provide and share insights into the needs of MFRs for MR training. In addition, as these data points directly influence the next project steps, they will be shared by inclusion in the following work packages.</p>

Data types / category	Format	Origin / Source	Purpose & Relation to Obj.	Utility beyond project
<p>Training assessments</p> <ul style="list-style-type: none"> • Performance ratings • Activity logs • Mannikin "vitals" (over time) • Training observations and observation protocols • Trainee movement data • Attention • Cognitive workload • Distraction • Drowsiness 	<p>Word, Excel, CSV, images, videos</p>	<p>Ratings by trainers during training sessions for MFRs; sensor recordings on trainees, data provided by mannikins, eye-tracking system, EEG, etc.</p>	<p>Assessment of training processes (obj. 2); dashboard for scenario control and automated scenario control (obj. 3); establishing relevant behavioural indicators for the EPME model and smart scenario control (obj. 3)</p>	<p>While data from individual training sessions cannot be shared for privacy reasons, aggregated data can provide guidelines for MR training.</p>
<p>Stress measurements, e.g.</p> <ul style="list-style-type: none"> • Self-assessment • Heart rate variability • Skin conductivity • Cortisol & alpha-amylase 	<p>Time series, sensor data, biological samples</p>	<p>Questionnaires, wearable sensors, ECGs, saliva samples collected during / after training sessions</p>	<p>Assessment of existing training processes; evaluation and comparison with VR & mannikin based (mixed reality) training (obj. 2); physiological measurements are shown on a trainer dashboard and used as feedback loop for smart scenario control (obj. 3).</p>	<p>The results of the assessment can provide important insights into the value of mixed reality training for first responders and provide guidelines for their effective design in the future.</p>

Data types / category	Format	Origin / Source	Purpose & Relation to Obj.	Utility beyond project
Training history / trainee profiles	Databases	Previous training assessments & personal data points	Provide an overview of the development / “growth” of trainees over time to better assess training effectivity (obj. 2 and 3)	As the trainee profiles are likely to include personal data (related to the trainee’s age, health, etc.), this data will only be used in the project software tools and will not be shared outside the project.
Experience measurements: <ul style="list-style-type: none"> • Quality of Experience (QoE) • Quality of Learning Experience (QoLe) • Presence 	Word/Excel files	Contextual interviews, focus groups, questionnaires	Assessment of factors affecting user experience and presence in VR (obj. 1 and 2)	The results of the assessment can provide important insights into the value of mixed reality training for first responders and provide guidelines for their effective design in the future.

Data types / category	Format	Origin / Source	Purpose & Relation to Obj.	Utility beyond project
<p>Psychological Characteristics</p> <ul style="list-style-type: none"> • Perceived resilience • Mental toughness • Mindfulness • Personality • Coping styles • Rate of perceived exertion • Team performance • Sense of unity • Evaluation of the training • Evaluation of the MR experience 	<p>Word/Excel, Database, CSV</p>	<p>Questionnaires, interviews (open questions)</p>	<p>Establishing psychological correlates with performance and psychophysiological stress to optimize performance (obj. 3).</p>	<p>Psychological training in addition to practical training can enhance EPME and recovery from stressful experiences.</p>

Data types / category	Format	Origin / Source	Purpose & Relation to Obj.	Utility beyond project
<p>Evaluation (through subjective experience and ratings) of the user-centred research and development process</p>	<p>CSV/XLS Exports, R-files, Word documents, photos / images of workshop results</p>	<p>Contextual interviews, focus groups, questionnaires, co-creation workshops</p>	<p>Related to obj. 2 as it provides evaluation for how the agile end user centred research (AEUCR) methodology is working and how to improve it for the remainder of the project</p>	<p>Based on the aggregation of this data type, improvements will be made to the AEUCR methodology. The findings from the evaluation of the experience of participants of the methodology will also be shared as methodological findings in publications. It helps to determine the necessary developments for a final product after the project.</p>
<p>Simulation Data</p> <ul style="list-style-type: none"> • Positions of trainees • Headings of trainees • Eye tracking • Time needed for tasks / sub-tasks • Movement patterns • Voice recordings 	<p>Time series, sensor data, bio signal samples (CSV / XLS)</p>	<p>VR-application, sensor data, wearables</p>	<p>Related to obj. 3 for team performance evaluation purpose.</p>	<p>Based on this data it's planned to develop a team performance tool for VR applications, which could also be of interest for future VR applications.</p>

2 Data Management

MED1stMR opted into the Open Research Data Pilot, and aims to follow the FAIR (Findable, Accessible, Interoperable & Reusable) principles wherever possible. While certain data (such as personal and health data of the participants) cannot be shared, actions are already planned from the start to make other project results more easily available.

To ensure discoverability of the data both within and outside the project, a consistent naming scheme needs to be followed. Whenever multiple versions of a document/data set are published, version numbers or timestamps (for data sets) shall be included in the title, in order to facilitate easier acquisition of the newest data as well as comparisons between different versions. For software versions, a version naming scheme similar to the semantic versioning schema (<https://semver.org/>) will be followed where possible, except for example when existing software already utilises a different naming scheme, where changes may lead to confusion.

- **Deliverables:**
Med1stMR_DX.X_Title_VersionNumber
- **Documents/data:**
Med1stMR_WPX_Title_OptionalTimestampOrVersionNumber

Within the project, Microsoft Teams and an underlying Sharepoint server are used to manage communication and data sharing within the consortium. The MS Teams is structured according to Work Packages, allowing for easier communication within work packages and higher discoverability of associated files, as they can be structured according to the relevant work packages. Beside the work package folders/channels, additional ones are provided for certain groups and purposes as shown in Figure 1 (e.g. executive board, PhD Roundtable, etc.). Access to Microsoft Teams and the Sharepoint server is managed by AIT, with access being limited to consortium members. It is also possible to create sub-folders in each of the Work package channels to allow for better structuring and finding of data.

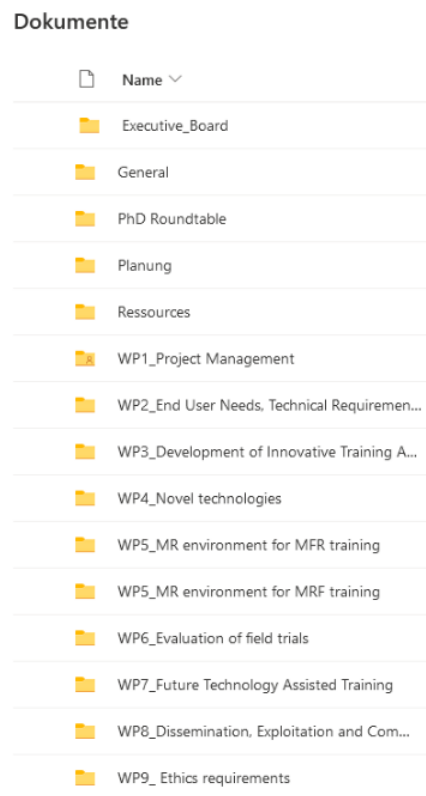


Figure 1: AIT Sharepoint folder structure based on work packages for MED1stMR

More sensitive data, like raw study results, video/audio recordings etc. will be separately hosted on a Protected Sharepoint Server (PSS) with additional security precautions and more limited access. More details about these measures and the server will be provided in the data security segment.

For collecting data of all state-of-the-art analyses being conducted specifically, the service Mendeley is being used. The data is structured in folders depending on the topic, and keywords are provided to facilitate easier recoverability of information.

For providing open access to all parts of data that are not sensitive and can be openly shared, we reviewed multiple options using services such as the re3data.org, openaire.eu and the suggestions in the [Open Data Guidelines](#) provided by the European Commission.

The consortium finally decided on utilising [Zenodo](https://zenodo.org). As the data collected and created in MED1stMR spans a variety of domains and consists of vastly different data types, the use of a purely domain-specific repository was discounted in favour of a more multidisciplinary provider. Zenodo supports a variety of different data types and offers many important services, such as providing a persistent identifier (PID) for easier retrieval. It is hosted by CERN, a renowned science institution located in continental Europe, which also guarantees [long-term availability](#) of all data being hosted. Beside Zenodo,

relevant data sets (or their parts) may additionally be published in databases that are specific to certain domains, to make them more easily findable and impactful in the respective community.

To maximize the utility and discoverability of uploaded data, it will be enhanced with descriptive metadata and keywords. For databases and tables that are being published, column names will either have descriptive names that make their meaning understandable, or they will be appended with metadata that explains their meaning, to increase their use outside the consortium. The description of data sets may also include additional information where reasonable, e.g. about how the data was collected, data standards, items that were excluded due to privacy reasons etc.

Certain data will not be made openly available, as described in the data summary chapter. This is especially the case for data sets containing data about directly identifiable participants or bio signal data about specific participants (e.g. acquired via ECG). To protect the privacy of our participants, items relating to personal data will only be published in an aggregated form or under pseudonyms (where applicable) and will never be attributable to a specific person.

As there is no common open standard for mixed reality training with manikins yet, we will use our findings to define a suitable communication protocol between the various modules and systems. Proprietary data formats will be forgone in favour of open file formats where possible, to increase possible interoperability and the opportunities for expansion in the future.

In order to adhere to the rules of good scientific conduct, raw study data that has been used as a basis for publications in MED1stMR will be stored for 10 years after the end of the project. Access to these files will be restricted to potential investigations regarding the scientific conduct. Any such requests need to be coordinated with the Data Protection Officer of AIT and need to provide a clear reference to the scientific publication that used the data and the specific complaint.

3 Allocation of Resources & Data Security

For personal data that is collected during the project, such as audio and video recordings of studies, AIT is hosting a protected sharepoint server (PSS) on its IT premises, located in Giefinggasse 4, A-1210 Vienna. The data is only saved on the local servers and not transferred to the Microsoft cloud and is regularly backed up to ensure that data is not lost. The sharepoint server is managed by AIT's IT department and is part of its security concept. The hosting costs for the service are covered by the overhead costs of AIT.

The data is located in a separate extranet section of the AIT IT infrastructure, can only be accessed via an encrypted connection ("https") and is secured by multiple firewalls and other security measures. Access to this server is only provided to the research partners in the project (namely AIT, UBERN, UHEI, UMU and IDENER), and require authorisation.

A registration process has been enabled to control the users that receive credentials to log into the system. Physical access to the server is strictly secured as well and limited to IT staff at AIT. In case that study data is stored locally for analysis reasons, the respective partner employees will handle the data with necessary care and corresponding security measures. With regard to materials that needs to be deleted within a certain timeframe (e.g. video/audio recordings where the informed consent has defined their deletion after 90 days), the partner employees are also obligated to delete locally saved copies of these materials.

In the unexpected case that access to a specific file from the PSS is critical for a non-research consortium partner:

- A research partner may present the relevant data to them (e.g. during a conference call) without sending out the file itself, or
- A request for access to a specific file (with an explanation on why it is necessary) needs to be provided to and approved by the Executive board. In case of approval, the requested files must be provided to the participant without sharing any more files or data points than necessary for the determined purpose.

4 Ethical Aspects

As detailed in the previous chapter of this deliverable, we place great emphasis on protecting personal data. When personal data is referenced in results that are published (e.g. in public deliverables, scientific publications, open repositories etc.), participant names will be pseudonymized and/or data will be aggregated to protect the privacy of our participants. Participation in all studies and similar activities in the project will be voluntary, and participants are informed beforehand about the purpose of each study, the data being collected, the storage duration of the data and their rights to withdraw at any time or request the deletion of the data. Contact data of the participants will only be shared strictly on a need-to-know basis, where it is vital for the recruitment, coordination or conduction of studies or similar processes.

To assess the psychophysiological stress responses, saliva samples are meant to be collected for cortisol and alpha-amylase level analysis. While these samples would theoretically provide the potential for being used to analyse the DNA of participants, their use during this project is strictly limited to identifying the aforementioned stress hormones. The samples will be destroyed after the stress hormone levels have been detected.

Additional details about the mechanisms each research partner is taking to ensure the protection of personal data can be found in Deliverable D9.3 (POPD).

5 Conclusions

Table 3 of this document defines the various types of data and data sets that are generated/collected in MED1stMR, how the data is saved and handled, as well as how it can be used both within and outside of the project. The open data repository “Zenodo” has been chosen to distribute the data sets that can be shared openly, while also defining processes to safely secure more private data that needs additional protection (such as raw study data and recordings of studies/workshops). For the latter, a separate sharepoint that is locally hosted by AIT (the Protected Sharepoint Server or PSS) has been set up, with clear data security and ethical guidelines on how the data is secured and handled.