European Commission, Horizon 2020

D6.4 Human Factor Impact Analysis and Machine Learning Models

SHOTPROS

Deliverable	D6.4
Deliverable Lead	AIT
Related work package	WP6
Author(s)	Till Bieg Sebastian Egger-Lampl Helmut Schrom-Feiertag Daniele Pretolesi Bruno Gardlo
Dissemination level	PUBLIC
Due submission date	31.01.2022
Actual submission	31.01.2022
Project number	833672
Instrument	RIA
Start date of project	01.05.2019
Duration	42 months
Version log	V1.0

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 833672. The content reflects only the SHOTPROS consortium's view. Research Executive Agency and European Commission is not liable for any use that may be made of the information contained herein.

Versions

Vers.	Date	Author	Description
0.1	13/08/21	Till Bieg (AIT)	First Draft
0.2	27/09/21	Sebastian Egger-Lampl (AIT)	Updated document structure, aims and goals
0.3	28/09/21	Sebastian Egger-Lampl (AIT)	Feedback from Markus Murtinger included
0.4	20/12/21	Daniele Pretolesi (AIT) Till Bieg (AIT)	Rewriting of the document and Finalisation.
0.5	10/01/22	Daniele Pretolesi (AIT)	Feedback from Markus Murtinger included
0.6	24/01/22	Daniele Pretolesi (AIT)	Incorporating comments and feedback from partners
0.7	26/01/22	Helmut Schrom-Feiertag (AIT), Markus Murtinger (USE)	Quality assurance and final content check
1.0	31/01/22	Valerie Schlagenhaufen (USE) Gerhard Helletzgruber (USE)	Final check, formatting and submission

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 833672. The content reflects only the SHOTPROS consortium's view. Research Executive Agency and European Commission is not liable for any use that may be made of the information contained herein.

List of Acronyms and Abbreviations

Acronym / Abbreviation	
AI	Artificial Intelligence
ELT	Extract Transform Load
ML	Machine Learning
NPC	Non-Player Character
RAT	Risk Assessment Tool
VAS	Visual Analog Scale
VR	Virtual Reality

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 833672. The content reflects only the SHOTPROS consortium's view. Research Executive Agency and European Commission is not liable for any use that may be made of the information contained herein.

Table of Contents

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 833672. The content reflects only the SHOTPROS consortium's view. Research Executive Agency and European Commission is not liable for any use that may be made of the information contained herein.

Table of Figures

Tables

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 833672. The content reflects only the SHOTPROS consortium's view. Research Executive Agency and European Commission is not liable for any use that may be made of the information contained herein.

Executive Summary

The aim of this deliverable is to describe a machine learning based concept for stress prediction that can be further extended for scenario control. To achieve this, the proposed concept builds upon outcomes from the Real-Time Training Progress Assessment Tool (D4.5) and the Risk Assessment Tool (RAT – see D4.7) and uses the data from the Human Factors Studies (D6.1) conducted as part of the SHOTPROS project which have been collected and processed in D6.3 Human Factors Study Transcripts.

Previous research on the current role of Artificial Intelligence (AI) in VR applications has shown that the inclusion of AI techniques in a VR experience is limited to two main applications: as a tool to analyse the data produced by the users and as a behavioral tree to control Non-Playing Characters (NPCs) in a simulation. The SHOTPROS project offers an ideal starting point to consolidate the role of AI for VR applications. In particular, machine learning could be employed to manage participants stress and provide autonomous changes to the simulation when necessary. In order to achieve this result, a modelling pipeline for stress prediction is described by implementing the stress prediction model of D4.5 and the RAT discussed in D4.7. Since a ML approach requires a considerable amount of data for model training, a standardised data structure is proposed which allows data from previous SHOTPROS studies to be used for model training. Compared to the RAT, the modelling concept described in this deliverable allows for more accurate stress predictions and opens up further integrations of autonomous systems in the project. In particular, advantages for both trainees and trainers are identified. For example, live prediction of stress during the training exercise could be implemented and a dashboard, to support trainers in understanding the prediction of the model, designed to encourage a trustful relationship between the user and an autonomous system.

This project has received funding from the European Union’s Horizon 2020 Research and Innovation Programme under grant agreement No 833672. The content reflects only the SHOTPROS consortium's view. Research Executive Agency and European Commission is not liable for any use that may be made of the information contained herein.

1 Added Value

1.1 Relation to SHOTPROS Work Packages (WPs)

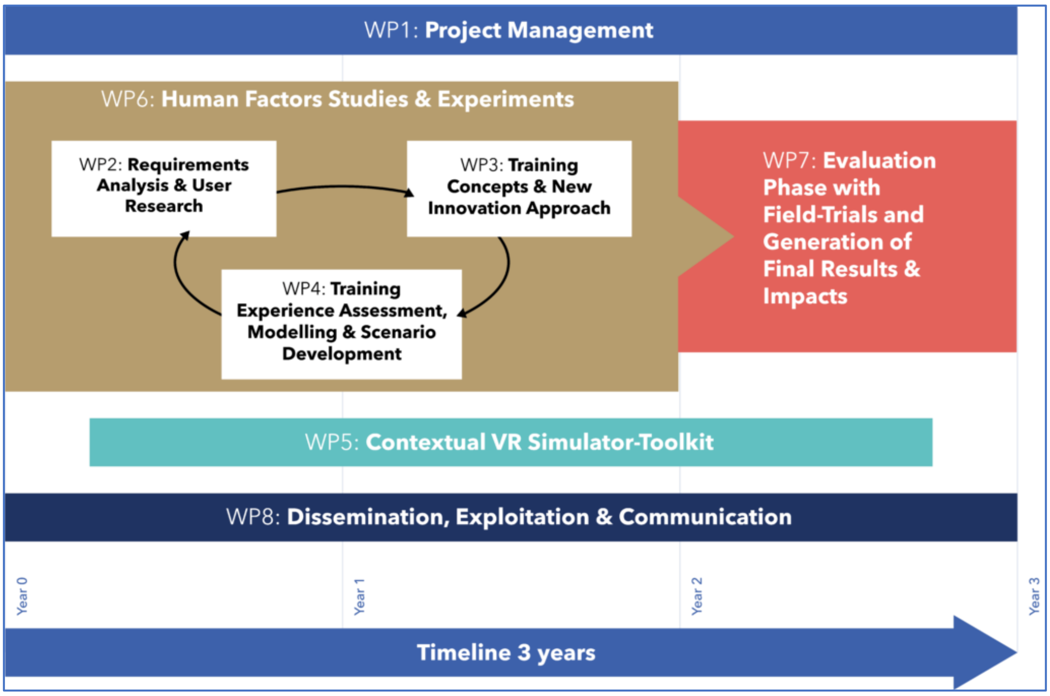

Figure 1: SHOTPROS work packages.

WP6, the container of this document, combines the Human Factor (HF) studies and Experiments executed within SHOTPROS to set up requirements, a training concept and additional modelling and development for a VR police training solution. Data gained throughout the different HF studies (listed in D6.1) and requirements and needs from LEAs (as described in D4.6) were used to implement the deliverable at hand. Machine learning techniques were used to derive respective prediction models for DMA in addition to the model defined in D3.2. It herby delivers insights to beyond state-of-the-art approaches to scenario-based VR learning.

1.2 D6.4 is informed by the following deliverables

Deliverable	How did theses deliverables influence D6.4
D3.2	The foundation of the DMA training is laid by the scientific model and therefore influences the analysing methods used in D6.4.
D4.1	The Cue Repository for Personalization and Customization of VR Training Scenarios delivered the list of stressors that informed the machine learning

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 833672. The content reflects only the SHOTPROS consortium's view. Research Executive Agency and European Commission is not liable for any use that may be made of the information contained herein.

	approach to use stressors for the steering of stress persception of police trainees.
D4.3	Concept for Physiological Measurement Suite for Stress Assessment delivered requirements to set up the physiological stress measurement tools and calculation within the solution to enable execution the HF studies.
D4.5	The Real Time training Assessment Tool delivered information on how to set-up the machine learning approaches within the SHOTPROS VR solution and within the planned Dashboard view.
D4.6	Requirements towards simpler scenario set up and simplified scenario steering influenced the machine learning approach of D6.4
D4.7	The Risk Assessment Tool delivered criteria on the selection of different stress levels (risks) in scenarios and are used to predict and steer the scenario on machine learning based data.
D6.1	This deliverable list all HF studies and therefore builds an overview on the objectives and methods of all HF studies and informed D6.4 to choose the relevant data from certain studies.
D6.3	The data set from all HF studies were screened for appropriate use in the context of machine learning.

Table 1. Influence of SHOTPROS deliverables on D4.6.

1.3 D6.4 consequently feeds into the following deliverables

Deliverable	How does D6.4 influence other Deliverables within SHOTPROS
D4.6	Based on the suggestions raised in D6.4 regarding further use cases for VR training, additional requirements towards a future VR solution have been added to the product backlog.
D5.1	As 6.4 provides options on how to integrate machine learning as a feature to VR training, it has to be considered in D5.1 how such a feature could be used and integrated.
D7.5	As D6.4 delivers opportunities to automise and support the training process, this should be considered during the set up of an ideal training framework.

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 833672. The content reflects only the SHOTPROS consortium's view. Research Executive Agency and European Commission is not liable for any use that may be made of the information contained herein.

D7.6	D4.6 influences D7.6 the final guidelines for VR training with options of supporting the trainer in setting up ideal scenarios before and adapt them during the training based on machine learning data – for final VR guidelines this is an important input.
D8.6	For a Exploitation Plan and Business outlook of SHOTPROS it is important to consider beyond-state-of-the-art approaches like AI based scenario steering suggested in D6.4

Table 2. Influence of D4.6 on other SHOTPROS deliverables.

1.4 SHOTPROS Objectives Relation

When considering the SHOTPROS objectives, the D6.4 on a first attempt is mainly important for a future VR training environment. But as the support of AI with machine learning attempts also to support the trainer, this deliverable is also aiming for the objectives of the training framework (objective 3) and potential VR guidelines (objective 4). Within the European VR network (objective 5) this topic could raise additional knowledge exchange and further research projects with other LEA partners.

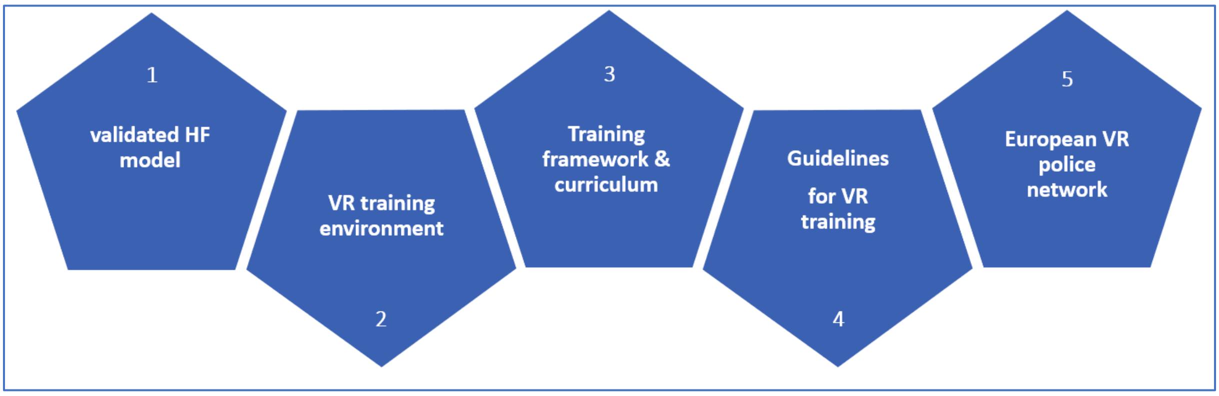

Figure 2. SHOTPROS objectives – overview.

2 Introduction

During the observations and training with the SHOTPROS system, the need for further automation of aspects of the training process became clear (see requirements in D6.4), in particular how the system can support trainers and trainees to achieve and maintain an

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 833672. The content reflects only the SHOTPROS consortium's view. Research Executive Agency and European Commission is not liable for any use that may be made of the information contained herein.

optimal stress level during the exercise. With this in perspective, several workshops and discussions among the technical partners of the project were carried out in the summer of 2021. The result converged towards the use of AI, in particular of machine learning, to improve the quality of the RAT tool and allow for more accurate stress prediction and visualisation as well as scenario steering in the trainer Dashboard (D4.5). In addition, the prospect of an AI implementation in SHOTPROS uncovered original applications for further automation of the training practice.

As basis for the machine learning (ML) framework described in this deliverable the RAT (D4.7) serves as a starting point. The RAT can be used by trainers to systematically pre-determine the amount and complexity of stressors needed to create a scenario of optimal complexity and stress (see D3.3, chapter 3.3.4) that challenges the trainees but does not overcharge them with respect to stress levels. More specifically, the trainer specifies characteristics of the training scenario (e.g., location, weather conditions, presence of other people) and the trainee (e.g., experience level), and based on this input the complexity of a scenario is estimated by the tool (the complexity in turn serves as an indicator for induced stress). Thus, the tool makes it possible for trainers to systematically pre-determine the type and number of stressors to achieve an appropriate complexity level in a scenario that fits pre-defined training objectives. The RAT is based on an evaluation system derived from an empirical study on the perception of stress among police officers that determined static weights of individual stressors for estimating the complexity of a scenario. To overcome this limitation of static weights, ML-based methods are a feasible approach. The aim of the proposed framework is therefore to derive weights of the different features (e.g., characteristics of a scenario) iteratively by an algorithm based on existing data (e.g., data from past trainings representing features of a scenario, features of the trainee as well as associated stress levels). To achieve this objective, the document is structured as follows: Chapter 3 gives an overview of existing work and approaches for the utilisation of ML and artificial intelligence for the estimation and the management of stress and experience in training and technology experience related contexts. Based on the related work, chapter 4 proposes a framework for a learning model that can be utilised for assisting a human instructor in scenario control or automated control of scenarios based on stress level optimisation and management. Chapter 5 and 6 describe possible future outcomes for further implementation of an autonomous system in the project and proposes novel solutions for both trainers and trainees. Finally, chapter 7 draws a conclusion on the proposed model and outlines the main advantages for its implementation.

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 833672. The content reflects only the SHOTPROS consortium's view. Research Executive Agency and European Commission is not liable for any use that may be made of the information contained herein.

3 AI and VR overview with the final aim for stress and experience management in training

3.1 Current state of AI in VR applications

The increased popularity of VR applications for training purposes has led researchers to explore how these systems could benefit from Artificial Intelligence (AI) to support and create personalized and more effective training scenarios. The synergy between AI and VR has been explored and adopted especially in medical applications (Winkler-Schwartz et al., 2019; Harmon et al., 2020). As shown in Bissonnette et al., (2019) a ML algorithm (Support Vector Machine) to predict training levels was successfully implemented during a VR surgical skill training. Moreover, the algorithm was able to identify 12 original metrics which highlighted key training elements to help distinguish between senior and junior surgeons. While VR is primarily used to elicit a higher sense of presence and immersion in users, AI algorithms have been applied to immersive scenarios with different goals. For example, in the work of Sadeghi and colleagues (2021), AI was used to segment patients’ pulmonary scans which could then be visualized in a VR environment allowing surgeons an easier and insightful look into patients' anatomy which may prove critical when planning surgical procedures. Differently, in the work of Riva et al., (2019) ML models are proposed to optimize and personalize patients' psychological treatment by feeding the algorithm behaviour related data collected during VR therapy sessions such as movements, eye movements and physiological data (i.e., electroencephalograms, heart rate, etc.). In the works presented so far, AI models are mostly used as an additional tool together with VR to collect user's performance data and return meaningful insights, but they have no direct effect on the virtual environment presented to the user.

Outside of the medical domain, AI and VR share a closer relation in which AI algorithms can affect and intervene in the simulation based on users' behaviour. An example of this relation can be found in Li et al., (2020). Li and colleagues developed a VR safety training system in coal mines that features two AI-controlled NPCs. The NPCs have been designed to simulate the behaviour of real miners and guide the trainees during the training exercise. Both NPCs are equipped with automatic navigation which allows them to adjust their movements based on the changes in the scenario furthermore, the NPCs behaviour tree allows them to execute multiple tasks in parallel and adapt their actions based on the user's decisions during the training. Similarly, Gluck and colleagues (2020) integrated AI methods in immersive VR training for soldiers. In this scenario, a team of soldiers (Blue Team) works in collaboration with an AI assistant drone to reach undetected the finish point avoiding the enemy team (Red Team)

This project has received funding from the European Union’s Horizon 2020 Research and Innovation Programme under grant agreement No 833672. The content reflects only the SHOTPROS consortium's view. Research Executive Agency and European Commission is not liable for any use that may be made of the information contained herein.

which consists of AI-controlled NPCs. The AI drone provides the Blue Team with information on where the enemies are located and which spots on the map are safe for hiding. On the other hand, the Red Team is composed of two types of enemy AIs, a patrolling enemy, which follows a static path until an opponent is detected, and a lackey enemy which randomly moves around the map. When an NPC spots the player, it will diverge from its path and move towards the location where the player was last seen at the same time, it will communicate to the other NPCs to converge on the same position.

As seen so far, the role of AI in VR applications has been explored by researchers to gain insights into users' behavior in VR or to enhance the simulation by introducing autonomous elements. However, the overview provided in this chapter shows that the role of AI, when included in a VR experience, is somehow limited to two aspects: as a tool to analyze the data produced by the participants and as a behavioral tree to control NPCs in a simulation.

3.2 Possibilities for AI and VR in SHOTPROS

To further expand and consolidate the role of AI systems in VR scenarios, the SHOTPROS project offers an ideal starting point to look into the opportunities and limitations of this integration. Thanks to the considerable amount of data collected during SHOTPROS HF studies (see D6.1 and D6.3), a ML approach could be designed to tackle aspects of the training that may benefit from the support of an AI algorithm. The next paragraphs will discuss some of the possible approaches and advantages a ML algorithm could bring to the training exercise.

3.2.1 Stress management

Firstly, a ML approach could be used to estimate, manage and maintain the optimal stress requirement during each training session. Interviews conducted with trainers and experience from other tactical training over the course of the project point out that police training, in particular VR training, should not overstress an officer. Poor training performance due to over-stress during a scenario may compromise the self-esteem of the trainee, which could then result in potentially life-threating mistakes while on duty. To achieve appropriate stress requirements, a ML algorithm could be applied to the RAT to calculate meaningful stress scores for the stressors. Since the scores currently used in the RAT were assigned based on an empirical study of police officers' perceptions of stress, the algorithm could be used to derive the appropriate weights for each variable based on previous trainings, thus improving the

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 833672. The content reflects only the SHOTPROS consortium's view. Research Executive Agency and European Commission is not liable for any use that may be made of the information contained herein.

validity of the RAT and the quality of trainings. A theoretical overview of how this result could be achieved is presented in the following chapters.

3.2.2 Changes to the virtual environment

The implementation of an AI algorithm to improve the RAT could be coupled with the possibility to apply immediate changes to the virtual environment to increase or reduce the stressors used in the scenario based on the live data collected during the exercise. For example, the AI system could decide to change the weapon assigned to an NPC to reduce the stress or add an additional offender if the participants are not stressed enough. Other changes could relate to the setting of the scene (e.g., number of civilians involved) and its environments, such as weather and visibility conditions or time of day. In doing so, the system can make sure that each participant achieves the best results from the training, providing higher levels of engagement and immersion.

The table below gives an overview of the stressors as listed in the RAT, as well as possible variations that the AI could apply if needed

Stressor	Description
NPC Civilian	Calm, Scared, Injured, Sick, Filming, Crazy and unresponsive, Fatigued, Non-Cooperative, Under strain, Asking questions, Speaks foreign language, Dead
NPC Offender	Unarmed, With weapon, Aggressive, Non-Cooperative, Speaks foreign language, Drawing weapon, Shooting, Holding hostage
NPC Animal	Calm, Aggressive, Agitated, Barking (Dog), Large size, Small size, Injured, Dead
NPC Crowd	Calm, Aggressive, Scared, Filming, Asking questions, Throwing objects, Non-Cooperative
Weather	Sunny, Rainy, Foggy, Windy, Sunrise, Daylight, Nightfall, Night
Sound	Scream, Cry, Door banged shut, Firearm, Loud music, Complete silence, Explosion, Animal noise
Visuals	Blood, Weapons, Bullets, Drugs, Injures, Smoke, Fire, Collapsing building, Car traffic

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 833672. The content reflects only the SHOTPROS consortium's view. Research Executive Agency and European Commission is not liable for any use that may be made of the information contained herein.

Smell	Strong or foul body odor, Chemical odor, Gas, Gunpowder, Drinking alcohol, Dead or decomposing body, Smoke, Fire, Urine, Garbage, Rotten food, Cannabis
Weapon	Firearm, Knife, Baseball bat, Drug needle, Bottle, Beer glass, Furniture, Crossbow, Chainsaw, Metal bar, Screwdriver, Vehicle

Table 3: Overview of the stressors found in the RAT and the possible variations that the systems could apply autonomously.

3.2.3 Trainer support

The benefits of including an AI system in SHOTPROS are not only limited to the trainees but could also be extended to the trainers. Trainers are highly experienced individuals but may lack the knowledge of how AIs work as a result, they could refuse the suggestions provided by an automated system because they might not trust it enough. To avoid this, it is important to establish a collaborative relationship between the AI and the trainer end user. In particular, during the debriefing moment, the AI could provide the trainer with a summary of meaningful correlations between variables captured during training (e.g. stress level, heart rate, attention), presented through insightful visualisations. In doing so, the knowledge and experience of the trainer can be enriched with the precious insights that only an AI is capable of highlighting in the data. As a result, a positive and more trustworthy collaboration can be established between the trainer and the AI system.

3.2.4 Automated creation of the training scenario

Lastly, AI could be used to create new and personalised training scenarios based on the stress requirements selected by the trainers. In the future, an AI algorithm could be directly integrated into the game engine and be responsible for the creation of a scenario. While the current training makes use of pre-defined scenarios, which could reduce participants engagement and immersion over time, AI-generated environments could provide trainees with original and challenging simulations. By implementing a similar approach to the one explained in 3.2.1 and 3.2.2, the AI could generate the most suitable environment based on trainees' previous sessions and automatically provide expected stress levels, per participant, using the RAT variables introduced in the scenario.

4 Stress management model description

The aim of this section is to describe how a ML model could utilise prior data and outcomes from previous activities in the SHOTPROS project. In particular, the goal of the model is to

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 833672. The content reflects only the SHOTPROS consortium's view. Research Executive Agency and European Commission is not liable for any use that may be made of the information contained herein.

predict subjective stress scores building upon the RAT for scenario control (D4.7) and findings regarding stress predictions described in D4.5. In this way, the approach proposed in this section extends outcomes aimed at a more accurate prediction of stress in training scenarios.

4.1 Data structure for a ML approach

An important prerequisite for building ML models is data that is used to train models. In the context of D6.4, this entails data from training sessions or similar encoding information on variables for prediction ("predictors" - e.g., specific features of the scenario) and predicted variables ("outcomes" - e.g., stress). The presence of the outcome variables in the data is mandatory at least for the case that a supervised learning approach is chosen. For the purpose of the model described in this chapter, the RAT and the data from previous SHOTPROS observations represent the first building blocks of the ML solution.

The following sections will define a high-level approach to implement a ML model for stress prediction. Chapter 4.1.1 provides a general overview of the aim of the model and which data it will use to achieve it. 4.1.2 focuses on the preprocessing pipeline, what considerations need to be taken and what shape the data should have before training the model. In chapter 4.1.3 the modelling workflow is presented. Lastly, in 4.1.4 the output produced by the model is discussed and feature importance implementations are analysed.

4.1.1 General outline of the model aims and which data it should use

In general, the various features (scenarios, trainees, etc.) used in the RAT (D4.7) are planned to be included in the model training when possible. However, additional features that are not represented in the RAT might be needed to more accurately predict stress as an outcome. For example, it is planned to use aggregates of physiological parameters such as heart rate variability in a model based on the results of D4.5. A complete overview of predictors that are planned to be included in the model (in addition to variables used in the RAT) can be found in Table 5 in the appendix. The outcome to be predicted is operationalized as a (subjective) stress score based on a Visual Analogue Scale (VAS) on a scale from 0 to 100. For further details regarding stress scores based on the VAS refer to D4.5 (see D4.5 chapter 7.1.2 and chapter 11.1.3).

4.1.2 Preprocessing of the data

Before feeding data from previous training into the model some adjustments to their structure needs to be performed. In particular, we suggest using a standardised flat structure in which the previously collected observations should be stored. The proposed format includes characteristics of the trainee, characteristics of the training session and scenario and the

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 833672. The content reflects only the SHOTPROS consortium's view. Research Executive Agency and European Commission is not liable for any use that may be made of the information contained herein.

variables to predict. Future activities concerning the collection of data that is used as input for model training should explicitly consider this structure and choose variables to collect accordingly. This novel data structure should be saved in a standardised format such as comma-separated values (.csv) to ensure portability and usage across different modelling approaches. Alternatively, data could be stored in a (relational) database and retrieved via a standardised ETL process for modelling.

participant_id	age	years_experience	...	training_system	location	weather	objective_known	...	mean_stress
1	34	12	...	System_A	park	sunny	TRUE	...	7.5
2	58	29	...	System_B	bank	rainy	FALSE	...	4.1
...	...	...	...	...	...	...	...	...	...

Charateristics of the trainee (predictors)
Charateristics of the training session / scenario (predictors)
Variables to be predicted (outcomes)

Table 4: Example of the standardised structure in which data collected in different studies is organised for modelling

It should be taken into account that, depending on the chosen ML framework (e.g., scikit-learn, tidymodels), the proposed data structure might need to be converted into a receivable format or structure once imported in the computational environment. For the purpose of this deliverable, we will not describe approaches based on specific ML frameworks, but present a general description that is not dependent on the framework. Thus, further preprocessing of the data proposed in this section might be necessary (depending on the framework used).

Once the data have been into transformed the appropriate shape, it should be considered how many missing values are present in the dataset. Missing values represent a threat to many ML models as they can alter the predictions, lead to misinterpretations or simply will not be accepted by the model. For the modelling concept discussed here this is of particular importance as data from many different sources is combined, which is likely to introduce substantial amounts of missing data. In order to take care of this issue, several methods for handling missing values could be employed. Missing values could be completely removed from the dataset. However, this could drastically reduce the amount of usable data. Alternatively, missing values could be replaced using higher imputation methods such as random forest algorithms (Tang & Ishwaran, 2017).

4.2 Modelling workflow

After the preprocessing stage is concluded the data can be prepared for fitting to the model. The first step is to divide the dataset into a training and a testing set so that our model can be trained to generalise better and perform with new data. Two options to perform this division

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 833672. The content reflects only the SHOTPROS consortium's view. Research Executive Agency and European Commission is not liable for any use that may be made of the information contained herein.

can be equally applied to the dataset: training, test and evaluation split or training and test split with application of cross-validation to the test split.

With the training test and evaluation method data are divided across three subsets, the training split is used to train the model, the evaluation split refines the model and the test split checks how well the model performs on never seen data. On the other hand, in the training and test split with cross validation method, the training set is divided into several folds that are used alternately for training and testing during model refinement. This allows more precise conclusions regarding the generalisability of the model. In the final step, the model refined using cross-validation is applied to the test set to check its performance on new data.

Since defining a ML model is an iterative process and it is not possible to clearly define a model in advance without training, we suggest here a standard ML workflow to identify a model for stress prediction on the given data. To do so, in the next section we describe some of the possible approaches to tackle this issue.

Before testing candidate models, it is best practice to explore the data further to identify additional processing steps that might lead to a better outcome of the model. Some of these techniques include:

- **Dimensionality reduction** for reducing the number of variables and handle collinearity.
- **Feature engineering** to extract new features that are more informative for the model.
- **Feature selection** to limit the predictors to the most informative ones.

After applying these approaches to the data, we need to feed our subsets to the models we have decided to test. As a starting point we provide some models that we consider promising for approaching this task. Promising models for the task at hand include the following:

- Random forest (Breiman, 2001)
- Gradient boosting (Friedman, 2002)
- Elastic-net regression (Zou & Hastie, 2005)

For each of these models, performance is assessed using the evaluation subset or the cross-validation method described above. By comparing the models against each other we can identify their performances based on several metrics, for example root mean-squared error and mean-absolute error. For the models with the best performance, hyperparameters will be tuned (e.g., based on grid search) to further improve their performance and determine a final model. Once the best performing model has been identified it can finally be evaluated against the test subset to get a conclusion about the model's performance on data it has never

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 833672. The content reflects only the SHOTPROS consortium's view. Research Executive Agency and European Commission is not liable for any use that may be made of the information contained herein.

seen before. If the results on the test set are consistent with what has been observed above, the model can be used to predict the stress score based on characteristics of the training scenario and the trainee and is ready to for deployment in the SHOTPROS framework.

4.2.1 Outputs and further analysis

Having successfully trained and tested our model we shall now discuss the outputs obtained and the steps that should be taken to capitalise on them. The first important result is the predicted stress scores based on the characteristics of the training scenarios and trainee. Furthermore, the model can provide additional insights regarding the importance of specific features. A flexible approach to determine the importance of certain features is the permutation importance method (Pedregosa et al., 2011). Given a two-dimensional data matrix, permutation importance is obtained by randomly shuffling the rows of one feature column, then using the trained model to make predictions on this "corrupted" dataset. The obtained performance metric (e.g., R^2) is than compared against the model performance on the original dataset. This process is then repeated for each column in the dataset. For each feature, the difference between the performance on the original model and the respective permutated dataset indicates the importance of the respective feature. This allows for a broader interpretation of the results and provides insights into which features are relevant to manage stress during a training scenario.

4.2.2 Advantages and disadvantages

The implementation of a ML model is associated with several advantages. Firstly, ML has become a powerful method in recent years which has the potential to be used on large datasets, like the one described above, to obtain more accurate stress scores compared to the RAT. This methodology benefits greatly from large amounts of data and is likely to improve its predictions if more data are collected. Secondly, this approach provides a unique insight on which features are pivotal for stress management during training exercises and can highlight features previously underestimated for stress prediction. Lastly, the approach proposed in this deliverable allows to further expand the model to provide the capabilities of AI to other aspects of the project. More details on future implementations will be described in the next chapter.

Although a ML model has significant advantages, drawbacks are also present and must be considered if this approach should be implemented in the future. As discussed in the previous chapters, data collection represents the corner stone of a ML pipeline, hence, a significant effort needs to be taken to ensure a high quality in the data. At the moment, the lack of a unified framework for the data collection process requires a considerable time investment to

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 833672. The content reflects only the SHOTPROS consortium's view. Research Executive Agency and European Commission is not liable for any use that may be made of the information contained herein.

convert the existing data to the quality necessary for AI operations. Further, collecting new data that fit into this new structure represent also an added effort that should be considered. Finally we discuss a possible approach to the modeling workflow and propose some models that could lead to significant results - however, there is no guarantee that a model can be found that provides results that are accurate enough to be viable for practice. This risk arises because the underlying data may not contain sufficient information to accurately predict stress. Thus, the approach proposed in this work may not be viable. Hence, other approaches may be considered to reach a satisfactory result (e.g., refinement of hand-crafted decision rules to predict stress similar to the RAT).

5 Future thinking

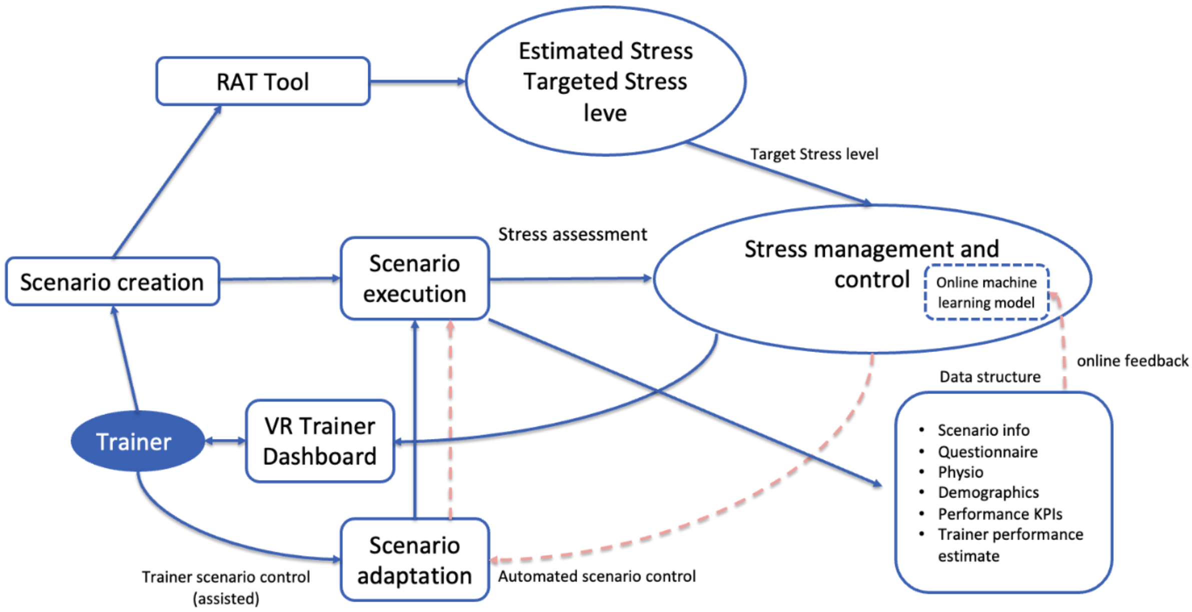

Figure 3: Structure of ML based scenario control loop (incl. data storage).

In this section, we discuss ideas to further expand the initial model described earlier and integrate it into existing training frameworks for decision-making and acting. Having discussed the possible implementation of a ML model for stress management, the SHOTPROS framework could be further developed to support other automated features. As already mentioned in the previous chapters, a ML model could be used to predict stress levels of trainees during a scenario. In particular, a model could be designed (by further expanding the modelling described in chapter 4) to support trainers with the topic of stress management and support to maintain a constant level of stress with the trainees (Figure 3). This output could be achieved by developing an online model which could keep track of a trainee's stress during

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 833672. The content reflects only the SHOTPROS consortium's view. Research Executive Agency and European Commission is not liable for any use that may be made of the information contained herein.

the scenario and provide suggestions for changes to the trainer or automatically adjust stressors to maintain a desired stress level.

In order to make the decisions or suggestions of the model transparent for the trainers, a dashboard is proposed (Figure 4 und 5), which gives the trainers an additional insight into the work of the AI. As the model will take actions autonomously or at least provide suggestions, the trainer must be informed about changes and needs to know what the reasons behind them are. The dashboard provided to the trainers will allow them to understand how the model is acting and what variables are being analysed during the training, this trainer-model interface will be constructed using data visualisation techniques to allow a faster and more comprehensible overview of the data. Moreover, the trainer's dashboard could be an important tool during the debriefing moment, providing the trainer with an at-a-glance knowledge of the trainees' physiological response to the training which was not available so far. To further improve the model, trainer decisions in reaction to past model suggestions could also be used. For example, when the trainer accepts or rejects a suggestion, this decision is recorded by the system and can be used later for model training. Once enough data has been collected on the trainer's responses to the suggestions, the model could fine tune its recommendations and automatically apply some of them while the attention of the trainer is focused on the exercise and less time can be spent controlling the dashboard. Another advantage of this approach would be that less experienced trainers are supported in accessing existing knowledge based on past trainings.

Finally, the interface for scenario configuration from the RAT, currently available as a stress calculator and model, could also be embedded into the trainers' dashboard providing the trainers with a complete tool that could be used in all phases of the training process, from designing the scenario to introducing changes during the training and evaluate the trainees afterwards. Figures 4 and 5 show a prototype of the interface and some of the core options that the trainer could access from it.

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 833672. The content reflects only the SHOTPROS consortium's view. Research Executive Agency and European Commission is not liable for any use that may be made of the information contained herein.

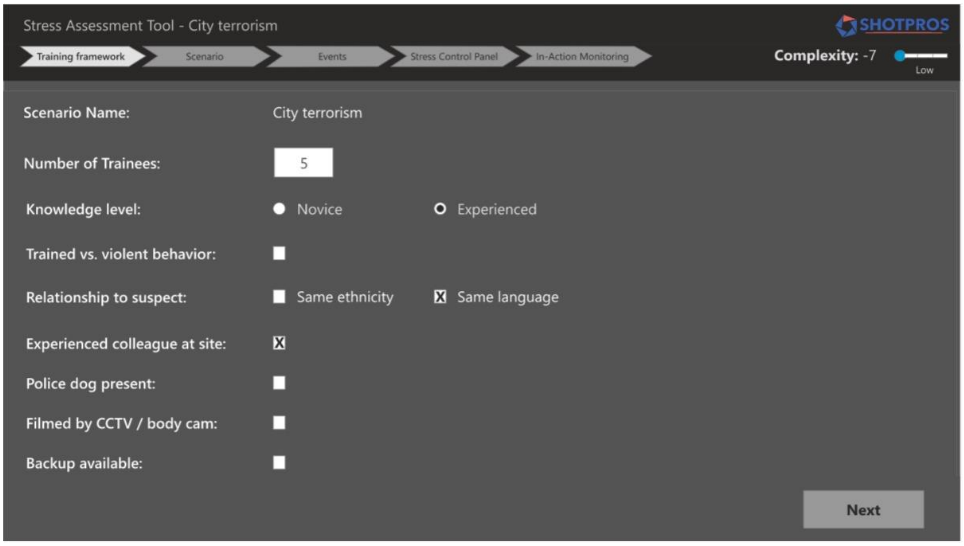

Figure 4: Starting screen of a wizard process to setup a training from the trainer perspective. The VR trainer has the possibility to setup the main information for the upcoming scenario and see the calculated complexity score of the scenario.

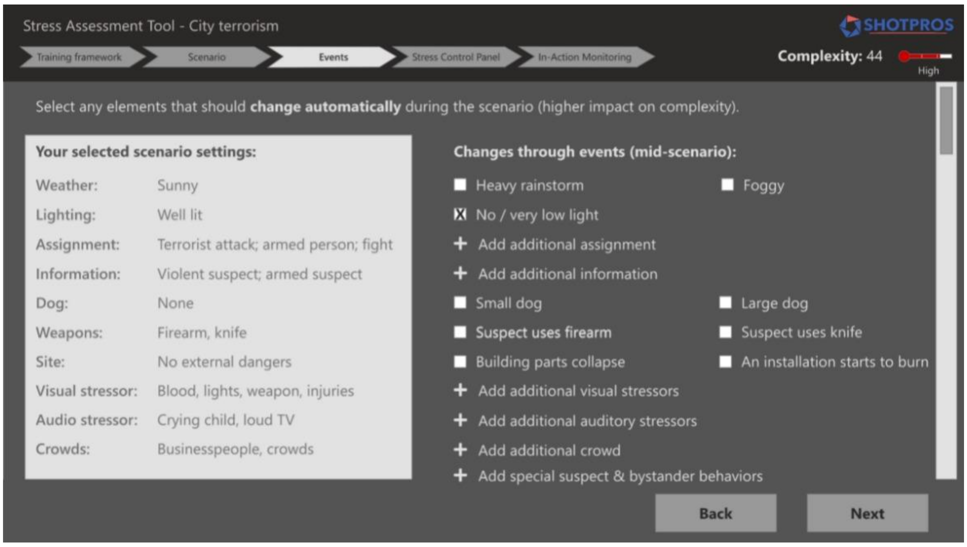

Figure 5: Event overview screen: Within the wizard process the trainer see the suggested events and situations for the scenario (e.g. sunny, used weapon, etc.) and can easily change the settings and add new stressors.

As seen so far, ML relies on a great amount of data to make accurate predictions. Thus, to ensure better and more reliable results, new data should be continuously fed to the model. This data can be collected during future training sessions or studies. To minimise the effort in data collection, it is preferable that for each training sessions, configurations made by the trainer (i.e., concerning features of the scenario) are automatically logged and transferred to a database in a standardised format. Moreover, self-reports from trainees should be gathered

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 833672. The content reflects only the SHOTPROS consortium's view. Research Executive Agency and European Commission is not liable for any use that may be made of the information contained herein.

using digital survey tools also storing collected data in a standardised format. The less time-consuming the data recording is for trainers and other parties involved, the more regularly and thoroughly data can be collected that can later be used for model training. Moreover, automatically storing all data in a standardised format minimises the need of further data pre-processing and consolidation.

6 Future use case

To better present the thoughts, ideas and approaches for a future AI-based VR training, together with the LEAs and technical partners, we have generated the following future user story:

Use Case: For today's training, the plan is to schedule a team of 4 experienced police officers with a focus on DMA training in complex situations under stress and high risk.

6.1 Setup of an AI-based VR-Training

The VR trainer collects the names and badge numbers from the participants and starts the VR training setup. Due to the innovative automation, it is sufficient to start this 10 minutes before the actual training. The trainer enters the 4 names and service numbers into the setup program and is immediately notified that the participants are experienced officers with more than 15 years of service each. Based on the data stored from their last training, the system recommends training in the medium to the high-stress range.

Using a wizard process, as depicted in Figure 4 and 5, the trainer is now guided through the individual steps with a concrete system recommendation (based on participant characteristics such as age, experience, degree of specialisation, etc.) as well as the existing training data (such as bio-signals, behaviour during training, etc.). These suggestions are generated from the existing data using the ML approach illustrated above, resulting in the following scenario suggestion: training in the medium to high-stress range in an urban area with a focus on a rampage in a school with an unknown number of suspects. Today, however, the trainer changes the scenario from a school to a busy shopping street, due to an actual incident that occurred a few days ago.

In the next step, based on an expected stress score for the participants, the AI system now suggests a selection of stressors be used as a starting point. The proposed stressors have been selected by the system from the RAT, following the pipeline described in chapter 4, accounting for previous performances and the current input from the trainer. Additionally, for didactic reasons, the system recommends a 2-step training sequence with run 1, a medium stress

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 833672. The content reflects only the SHOTPROS consortium's view. Research Executive Agency and European Commission is not liable for any use that may be made of the information contained herein.

scenario and following the After Action Review a 2nd run with a higher expected stress score (due to the experience of the team, a run with a low-stress score is waived). The trainer automatically receives the scenario description of both runs, the explanations of the starting points (including the radio message) and the list of proposed stressors. The system suggests the following stressors for scenario 1:

- Scenario during the day, at midday
- Normal weather situation (sunny) in summer
- Maximum of 2 suspects in the scenario
- Manageable space with few winding corridors
- Little danger of escape
- Due to the time of day, the square is less busy than usual (no big crowd)
- Little traffic on the streets

After confirming these stressors, the trainer starts the AI scenario editor and scenario 1 is automatically created and a suggestion for scenario 2 is made by the system:

- Early evening scenario
- Bad weather situation in winter (foggy, very dark)
- Minimum 2 suspects with live adjustment up to 4 suspects (determined by team performance)
- Many cursing possibilities in surrounding small alleys and pubs/bars
- Due to the time of year (just before Christmas the square is busy and a small Christmas market is set up)
- A lot of traffic on the surrounding streets due to rush hour

The trainer reviews the suggestion given by the AI. In this case, he decides to make some adjustments (e.g. start with 3 suspects, several injured already when the police team arrives and adds wind to the weather). The system records that the trainer has applied a different value than the suggested one and will use this information to update future training design.

Now the trainer is satisfied with the two scenarios and after less than 10 minutes in total, 2 complex tasks have been generated by the system and the training can start.

6.2 During the VR-Training

In Scenario 1, for example, individual stressors are adapted based on a good performance and KPIs (such as fast trading, paying attention to safety distances, tactical movement, etc.) within the Trainer Dashboard. The bio-signals collected during the exercise inform the model that the participants are below the required stress level thus, an adjustment to the scenario needs

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 833672. The content reflects only the SHOTPROS consortium's view. Research Executive Agency and European Commission is not liable for any use that may be made of the information contained herein.

to be made. The model communicates, via the dashboard, to the trainer that the participants are not stressed enough. The model suggests increasing the number of suspects and switching the weather from sunny to rainy, to achieve the required stress level. Once the changes are confirmed and applied, the AI will inform the trainer when the stress target is achieved.

In Scenario 2, a too high-stress score is reached early on, e.g. good radio contact is not possible due to the wind and foggy weather make training extremely difficult due to the lack of communication and visibility. Similarly, to the previous scenario, the system reacts automatically, in this scenario triggered by a rise in errors and elevated heart rate. This time, the AI suggests to the trainer to remove individual stressors and to attenuate the behaviour of NPCs. For example, instead of searching for a hiding place when they are spotted by an officer, offenders' NPCs stay in place and their reaction time is increased. These changes are also applied to the behaviour of the surrounding crowd which now is calmer and will comply with the trainees' instructions. Once the changes are approved, the system will apply them and inform the trainer when stress is returned to optimal levels. By taking actions to reduce the stress during the exercise, the performance and thus the learning success of the team increases again.

6.3 After the VR-Training

Once the training exercise is concluded, it is time for the trainer to evaluate the performance of the trainees, provide feedback and suggestions to avoid mistakes in real-life situations. Since police training only happens a few times a year, this moment is crucial to correct and improve officers' skills.

The body of experience collected by police trainers is unique and hardly this knowledge can be transferred to a machine as it consists of qualitative observations gathered over many years of police service. However, the scope of the AI system proposed in this work is not to replace the trainer but rather to enhance its knowledge by showing it the hidden correlations among data collected during the simulation. To better understand how this interaction happens let us imagine the debriefing of Scenario 2.

During the After-Action Review, the AI supports the trainer by providing an overview of the collected bio-signals through meaningful and interpretable visualisations. The system highlights moments where too high-stress levels were recorded and pairs each moment with a relevant video clip of what happened in the simulation. The AI shows when changes to simulation were applied and what impact they had on the trainees' performance. Through the dashboard, the trainer can easily review group and individual performances. Thanks to this novel overview, the trainer can better contextualise the actions of the individual within the

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 833672. The content reflects only the SHOTPROS consortium's view. Research Executive Agency and European Commission is not liable for any use that may be made of the information contained herein.

group performance and notice how the behaviour of a too stressed officer negatively affected their closer teammate.

7 Conclusion

In this deliverable, ideas for ML based scenario control based on outcomes of the SHOTPROS project were introduced. In particular, a modelling concept for stress prediction was outlined which extends the RAT (D4.7), the stress prediction model described in D4.5 and the collected data from several user studies (D4.6).

The modelling approach proposed in the present deliverable would use a standardised data structure that allows leveraging data from previous human factor studies in the SHOTPROS project for model training. This approach is promising for more accurately predicting stress in comparison to the RAT. Finally, suggestions for further expanding the initial model and integrating it into existing training frameworks were discussed. This includes ideas for future data collection and the integration of a model into training systems building upon the proposed modelling approach. In the future, this could allow to predict stress more accurately during live training and on this basis provide meaningful suggestions for trainers to adjust the training scenario.

An AI or ML development itself is not part of the SHOTPROS project, but this deliverable should serve as a starting point for future developments in the field of AI/ML for VR training topics. The document provides concrete approaches from the project that were developed together with end users and technical partners and thus a foundation for future considerations. The technical partners from the project will also take up these approaches for the final go-to-market strategy or incorporate them into the final product strategy. In any case, based on the approaches explained, we see a great potential for the environment of AI and ML topics in innovative VR training simulations.

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 833672. The content reflects only the SHOTPROS consortium's view. Research Executive Agency and European Commission is not liable for any use that may be made of the information contained herein.

8 Appendix

Table 5 describes predictor variables as part of the standardised data structure for modelling in addition to variables of the RAT. For variables already contained in the RAT refer see D4.7 (Risk Assessment Tool Manual).

Variable name	Description
Age	The age of the participant
Gender	The gender of the participant
Years of professional experience as a police officer	How many years the participant has been a police officer
Affinity for technology interaction	Tendency to engage in interactions with technology based on the Affinity for Technology Interaction scale (Franke et al., 2019)
Experience with VR	Past experience with VR technology expressed on 5-point Likert scale
Experience with the training system	Previous experience with the VR training system expressed on a 5-point Likert scale
Number of human-controlled (virtual) characters	How many NPCs are present in the simulation
Mean heart rate variability	Mean score of the variation in time between heartbeats
Mean heart rate	Mean score of the speed of the heartbeat
Mean respiratory rate	Mean score of breaths per minute

Table 5. Predictor variables to be included in the standardized data structure for modelling (in addtion to variables in the RAT).

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 833672. The content reflects only the SHOTPROS consortium's view. Research Executive Agency and European Commission is not liable for any use that may be made of the information contained herein.

9 References

Bissonnette, V., Mirchi, N., Ledwos, N., Alsidieri, G., Winkler-Schwartz, A., & Del Maestro, R. F. (2019). Artificial Intelligence Distinguishes Surgical Training Levels in a Virtual Reality Spinal Task. *The Journal of Bone and Joint Surgery. American Volume*, *101*(23), e127. https://doi.org/10.2106/JBJS.18.01197

Breiman, L. (2001). Random forests. *ML*, *45*(1), 5-32.

Franke, T., Attig, C., & Wessel, D. (2019). A personal resource for technology interaction: development and validation of the affinity for technology interaction (ATI) scale. *International Journal of Human–Computer Interaction*, *35*(6), 456-467.

Friedman, J. H. (2002). Stochastic gradient boosting. *Computational statistics & data analysis*, *38*(4), 367-378.

Gluck, A., Chen, J., & Paul, R. (2020). Artificial Intelligence Assisted Virtual Reality Warfighter Training System. *2020 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*, 386–389. https://doi.org/10.1109/AIVR50618.2020.00080

Harmon, J., Pitt, V., Summons, P., & Inder, K. J. (2021). Use of artificial intelligence and virtual reality within clinical simulation for nursing pain education: a scoping review. *Nurse Education Today*, *97*, 104700.

Li, M., Sun, Z., Jiang, Z., Tan, Z., & Chen, J. (2020). A Virtual Reality Platform for Safety Training in Coal Mines with AI and Cloud Computing. *Discrete Dynamics in Nature and Society*, *2020*, 1–7. https://doi.org/10.1155/2020/6243085

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, E. (2011). Scikit-learn: *ML in Python. Journal of ML research*, 12, 2825-2830.

Riva, G., Wiederhold, B. K., Lernia, D. D., Chirico, A., Riva, E. F. M., Mantovani, F., Cipresso, P., & Gaggioli, A. (n.d.). *Virtual Reality meets Artificial Intelligence: The Emergence of Advanced Digital Therapeutics and Digital Biomarkers*. 8.

Sadeghi, A. H., Maat, A. P. W. M., Taverne, Y. J. H. J., Cornelissen, R., Dingemans, A.-M. C., Bogers, A. J. J. C., & Mahtab, E. A. F. (2021). Virtual reality and artificial intelligence for 3-dimensional planning of lung segmentectomies. *JTCVS Techniques*, *7*, 309–321. https://doi.org/10.1016/j.xjtc.2021.03.016

Tang, F., & Ishwaran, H. (2017). Random forest missing data algorithms. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, *10*(6), 363-377.

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 833672. The content reflects only the SHOTPROS consortium's view. Research Executive Agency and European Commission is not liable for any use that may be made of the information contained herein.

Winkler-Schwartz, A., Bissonnette, V., Mirchi, N., Ponnudurai, N., Yilmaz, R., Ledwos, N., ... & Del Maestro, R. F. (2019). Artificial intelligence in medical education: best practices using ML to assess surgical expertise in virtual reality simulation. *Journal of surgical education*, *76*(6), 1681-1690.

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the royal statistical society: series B (statistical methodology)*, *67*(2), 301-320.

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 833672. The content reflects only the SHOTPROS consortium's view. Research Executive Agency and European Commission is not liable for any use that may be made of the information contained herein.