Ref. Ares(2022)369618 - 18/01/2022

D8.5 Lessons Learnt from Pilots

Document Date: 30/11/2021

Work Package 8: Pilots

Document Dissemination Level:
Public (PU)

TRESSPASS
robus**T** **R**isk bas**E**d **S**creening and alert **S**ystem for **PASS**engers and luggage
is funded by the Horizon 2020 Framework Programme of the EU for Research and Innovation.
Grant Agreement number: 787120

Abstract

The TRESSPASS project introduces a new Risk Based Border Management (RBBM) concept which has been developed to perform pre-travel screenings and border scans and checks at border crossing points at Europe's external borders.

For the implementation of this concept a series of various technological components and sensors have been developed by the technical partners in the project.

For the testing and validation of this concept and tools, three pilots were planned to take place in different EU pilot BCP sites, each of them representing different use cases and modalities, namely the Dutch Pilot (air case), Polish Pilot (land case) and the Greek Pilot (sea-cruise case).
Based on the RBBM concept, different detailed use case threat scenarios were developed for each pilot to test and validate the TRESSPASS concepts and technologies in different trial configurations, with the aim to provide validated configurations for the pilots' execution.

In light of the above, T8.5- Lessons Learnt gathered end-users' feedback with the overall various experiences gained and lessons learnt of all stakeholders involved in all three pilot cases and evaluate the TRESSPASS technology comparing the pilot results with the initial end-user requirements, needs and expectations and with the actions and objectives set in the project and in the original plan in T8.1 (D8.1).

A list of all the conclusions for each separate pilot will be included in the deliverable, along with the overall successes, issues, and obstacles that users came across with the use of TRESSPASS technology in the three border crossing pilot cases.

The current report (T.8.5) presents the results of the evaluation of each Pilot, carried out by the respective responsible partners, including lessons learned for future improvements. The evaluation carried out based on the process described in the Evaluation phase (3rd phase) of the Pilot Guidance Methodology (PGM), following the Planning and Execution phases of the pilots. Detailed description of PGM is provided in D8.1- Pilots Planning and Training (P.18).

Project Information

Project Name	robusT Risk basEd Screening and alert System for PASSengers and luggage
Project Acronym	TRESSPASS
Project Coordinator	National Center for Scientific Research "Demokritos", EL
Project Funded by	European Commission
Under the Programme	Horizon 2020 Secure Societies
Call	H2020-SEC-2016-2017 (SECURITY)
Topic	SEC-15-BES-2017 "Risk-based screening at border crossing"
Funding Instrument	Innovation Action
Grant Agreement No.	787120

Document Information

Document reference	
Document Title	**Lessons Learnt from Pilots**
Work Package reference	WP8
Delivery due date	30/11/2021
Dissemination Level	Public
Author(s)	**KEMEA**
Contributor(s)	**HP, PPA, WAT, PBG, RNM, TNO, FHG, NCSRD, CASRA, NUIM, Z&P, ED, VV, ALU-FR,**
Document Review Status	✓ Consortium ✓ WP leader ✓ Technical Manager ✓ Quality and Risk Manager ✓ Project Coordinator

List of Acronyms and Abbreviations

ACRONYM	EXPLANATION
API	Advanced Passenger Information
BCP	Border Crossing Point
CONOPS	Concept of Operations
DP	Data Protection
e-Gates	Electronic (Passport) Gates
EES	Entry Exit System
ETIAS	European Travel Information and Authorization System
GUI	Graphical User Interface
ICT	Information Communication Technology
ISS	Interview Support System
KPIs	Key Performance Indicators
PNR	Passengers Name Records
RBBM	Risk Based Border Management
RNM	Royal Netherlands Marechaussee
TCN	Third County National
TGM	Trials Guidance Methodology
TRAM	TRESSPASS Risk Assessment Method
PGM	Pilots Guidance Methodology
SBC	Schengen Border Code
SIS	Schengen Information System
SOP	Standard Operating Procedures
VIS	Visa Information System

Table of Contents

LIST OF FIGURES

1 INTRODUCTION

This report consolidates the results of the TRESSPASS Pilots which took place at the Dutch, Polish and Greek pilot BCPs respectively as a part of WP8 (T.8.2-4). The report also constitutes the final and closing step of the Evaluation phase of the Pilots Guidance Methodology (PGM), which follows the Planning and Execution phases of the Pilots. PGM derived from a customized version of Trials Guidance Methodology introduced by the Driver+ Crisis Management Project[1], tailored for the needs of TRESSPASS.

The Evaluation phase of PGM and the overall assessment of all the pilots presented here, is a result of the combination of external and internal evaluation[2] consisting of a) the feedback collected from the participants who were trained and tested the system during all pilots, and b) internal self-assessment-based feedback from the consortium members. Considering the above, the present report consists of a consolidated evaluation deriving from all individual pilots' evaluations which took place at the execution phase for each pilot run.

As per the evaluation methods used, several quantitative and qualitative questionnaires were utilized for gathering pilot participants feedback regarding components, training methods and tools, and executions methods which were implemented/ applied for the execution of the TRESSPASS pilots.

Specifics about the characteristics which were considered for the evaluation of the TRESSPASS components, and methods used to evaluate the overall performance and acceptance of each pilot, and the training have been described in Chapter 14 of TRESSPASS D8.1 Planning deliverable.

1.1 Aim of Task T8.5 and D8.5

Task T8.5, as described in GA, has started running after the end of the first pilot planned and ended after the end of all pilot cases. It gathers end-users' feedback with the overall various experiences gained and lessons learnt of all stakeholders involved in all three pilot case scenarios used in the 3 pilot sites and evaluate the TRESSPASS technology, concepts and aspects. It lists the conclusions for each separate pilot and overall, about the success story, issues, and the obstacles that users came across with the use of TRESSPASS technology in border crossing pilot cases.

In respect to the above, D8.5's aim is to present the pilots' results and compare them with the initial end-user requirements, needs and expectations, as well as actions and objectives set in the original plan in T8.1. Towards that end and according to PGM (T8.1), the following chapters present seven (7) major assessment aspects which were examined for each pilot by its respective stakeholders towards the overall evaluation of the Pilots (see Pilot sections below per pilot). By exploiting the assessments, this deliverable will then examine common aspects, successful elements and limitations between the three Pilots and draw conclusions for dissemination purposes and future implementation.

[1] https://www.driver-project.eu/trial-guidance-methodology/

[2] TRESSPASS D8.1 p.235

1.2 Contributions - Inputs/Outputs of the Task

As mentioned above, D8.5 is part of the Pilots Guidance Methodology's steps described in D8.1. As depicted with green color in the PGM Wheel figure below, D8.5 is the report produced in the third Evaluation phase and consists of the users' evaluation and the pilots assessment derived from the previous Pilots Execution phase of the 3 pilots.

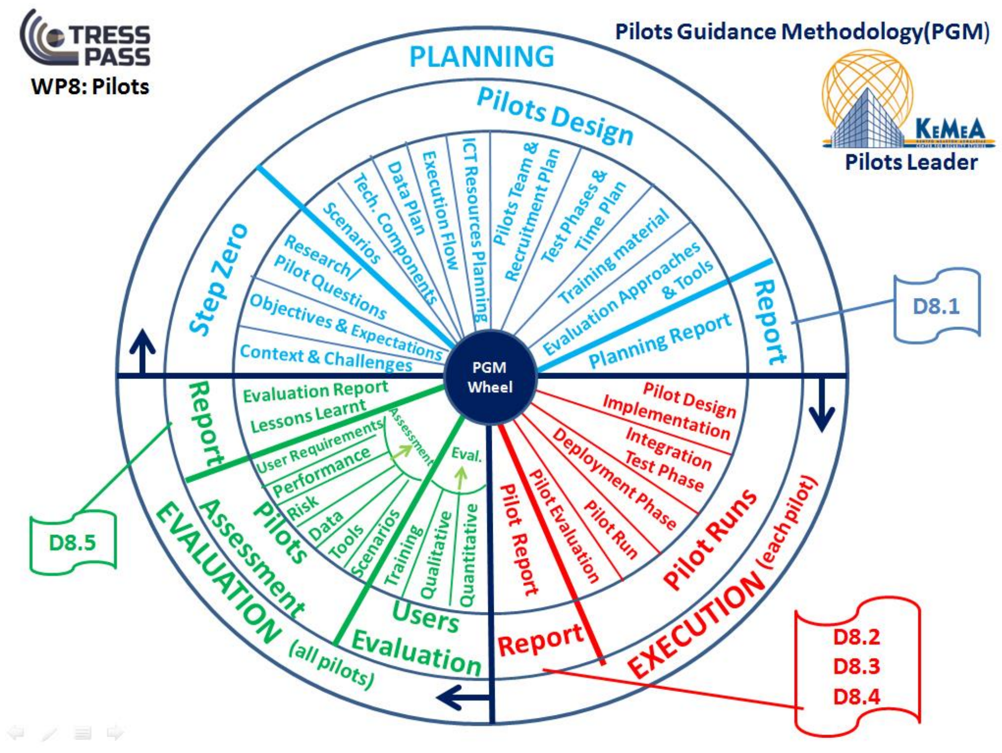


FIGURE 1 OVERVIEW OF TRESSPASS PILOTS GUIDANCE METHODOLOGY (PGM)

More specifically according to PGM, the evaluation phase is based on the following concept described in D8.1 Planning and overall includes the following parts.

- **Users Evaluation**

This part contains the evaluation based on the feedback which was collected from the participants who were trained and tested the system during all pilots. This evaluation is a merged summary evaluation from all individual pilots' evaluations taken place at the execution phase for each pilot run. Users' evaluation includes the following 3 parts described in the next sub-sections.

- ***Quantitative Evaluation***

This type of evaluation is an overall users' evaluation for all pilots based on the scores extracted from the individual surveys conducted through score-based questionnaires at each pilot evaluation in the execution phase of each pilot run.

- ***Qualitative Evaluation***

This type of evaluation is an overall users' evaluation for all pilots based on the feedback collected from the individual contacts through interviews, text-based questionnaires and comments provided at each pilot evaluation in the execution phase of each pilot run.

- ***Training Evaluation***

This type of evaluation is an overall users' evaluation for all pilots based on the simulation and training tools and training material used by the pilots' participants prior and during the pilots' execution phases.

- **Pilots Assessment**

Based on the users' evaluation in the previous steps and on internal self-assessment-based feedback from the consortium members, an overall assessment for all pilots is prepared to evaluate the following pilot elements against those targeted and planned in the preparation step zero and design phases. The assessment of all pilots includes 6 elements.

- ***Scenarios Assessment***

Assessment of all pilot scenarios and therefore the corresponding represented threats, how these have been handled and covered by the TRESSPASS system.

- ***Technology Assessment (Tools)***

Assessment of the TRESSPASS technology for each component and their contribution to all pilots and the results produced as well as the ICT infrastructure and resources deployed.

- ***Data Assessment***

Assessment of sources and data (real and mockup) including the Risk and Bona fide Indicators used, accessed, collected, processed and stored and the positive or negative impacts, if any, for the Data Protection and legal, ethical and societal issues.

- ***Risk Assessment***

Assessment of the risk-based concept, the indicators and the risk assessment methods used in pilots compared to those targeted and planned.

- ***Performance Assessment***

Assessment of Key Performance Indicators (KPIs) against those targeted to achieve.

- ***End Users' Requirements & Resources Assessment***

Assessment of user requirements against those targeted to achieve in D1.1 and of the Human Resources used in pilots in the role of Border/Security actors.

The diagram below illustrates the position and the inputs and outputs of task T8.5 in Work Package 8 (Pilots) taking inputs from the three pilot tasks (T8.2, T8.3 and T8.4) via the quantitative and qualitative user and internal evaluations and the corresponding deliverable reports (D8.2, D8.3, D8.4), producing as output the final Lessons Learnt D8.5 report comparing them at the same time with the initial planning (D8.1).

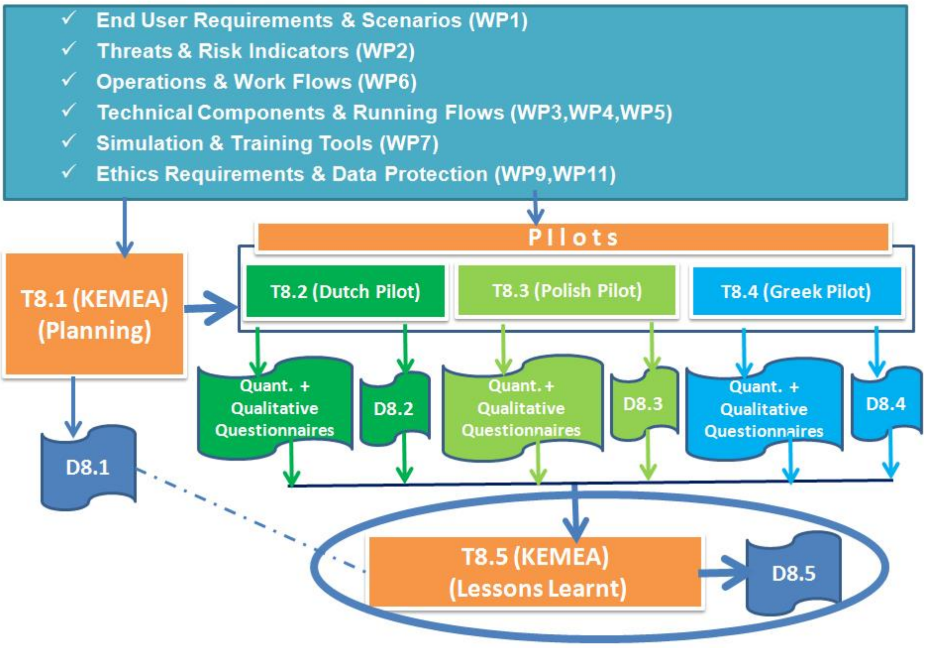


FIGURE 2 CONTRIBUTIONS & INPUTS/OUTPUTS OF TASK T8.5

1.3 Structure of the document

The rest of the document starts with 3 chapters, one per pilot:

Chapter 2: Dutch Pilot: Lessons Learned

Chapter 3: Polish Pilot: Lessons Learned

Chapter 4: Greek: Pilot: Lessons Learned

Each pilot chapter above consists of the following nine (9) sections which correspond to the 9 assessment aspects mentioned in the previous section plus an extra section for the overall lessons, limitations and obstacles observed and extracted from each pilot.

1) The operations involved in the pilot comparing them with the current and future ones
2) The Threat Scenarios applied and tested in each pilot
3) The TRESSPASS Technology (components, sensors and ICT infrastructure) used and tested
4) The mockup and real Data, including the Risk and Bona fide indicators used and tested
5) The Risk based concept and the Risk assessment methodologies applied
6) The Key Performance Indicators (KPIs): Effectiveness, Flow Rate, Efficiency
7) The end users Training
8) The Simulations
9) Overall Lessons, Limitations and Obstacles observed or faced

The document closes with the chapter 5 summarizing the overall lessons learnt and the conclusions from all pilots and the project.

2 DUTCH PILOT: LESSONS LEARNT

This chapter breaks down in sections each aspect of the Dutch Pilot assessment, the lessons learned, the experience gained, and the obstacles observed by the pilot's stakeholders upon the completion of the pilot's' execution, the collection and processing of the external and internal feedback via the evaluation tools (questionnaires) used and the relevant pilot deliverable report (D8.2).

The main research objective of the Dutch Air Pilot was formulated as:

To test the performance and potential of the TRESSPASS risk-based border management (RBBM) concept and system in terms of (the balance between) the following objective elements: effectiveness, efficiency, acceptance, flowrate and feasibility.

In other words, with the pilot the RNM wants to discover if RBBM is indeed a valid concept for the future when the flow of passengers will increase, and available personnel will be scarce. Border control is executed for the protection against threats. One of the threats that is capacity consuming but not only, is irregular migration. With RBBM the effectiveness of detecting all relevant threats should be equal or better than the current situation and/or adjustable based on varying the risk acceptance upon actual urgency and circumstances.

The Dutch Air Pilot comprised two main test activities:

1) The Air Pilot Recording Day at Lelystad Airport in November 2020.
2) The Experiment with border guards in April 2021.

Main purpose of the Air Pilot Recording Day was to generate representative airport passenger security camera footage in order to test the VTC system and the RTBA system with real input data.

The purpose of the Experiment with border guards was to **1) assess the decision making of border guards (on travellers on the threat irregular migration) in different conditions** and **2) evaluating the feasibility of the TRESSPASS system through the C2 and SPA interfaces**. Through the experiment we compared three conditions:

- The current system with the BCS main interface and data.
- The TRESSPASS system and data with the C2 interface.
- The TRESSPASS system and data with the SPA interface.

TABLE 2 EXPERIMENTAL CONDITIONS

Condition	Information overview		
	Traveller story (duration, purpose, means)	Information set based on current system	Risk indicators & risk scores
Current system (BCS)	X	X	
RBBM system (C2)	X	X	X
RBBM system (SPA)	X	X	X

2.1 Operations

As stated in D6.1 Observational studies methodology and research framework, a CONOPS "*is not a detailed user-requirements document but functions as a high-level description of the proposed end-state which helps to guide the technology development and implementation process*"[3]. The CONOPS represents a dynamic concept: indeed, it does not provide a static representation of the final state of a system but indicates the direction of the technological development.

The work carried out within T6.2 has led to the development of the CONOPS for the Dutch border. The CONOPS has been adopted as a starting point to understand where the TRESSPASS innovation would have been inserted and how this change would have affected the operational environment.

Besides the technological aspects, a crucial aspect is related to the human factor: the human component, indeed, is still of utmost importance in the TRESSPASS system; therefore, it is mandatory to fully comprehend the relation human-technology and the impact in terms of workload, procedures, and tasks. As the theoretical underpinning of the TRESSPASS CONOPS, activity theory views each actor in the system as interdependent and inseparable from each other and therefore each activity has a consequential impact on the other system dependencies.

[3] TRESSPASS, *D6.1 Observational studies methodology and research framework.*

This was a fundamental consideration during the development of the TRESSPASS CONOPS and provided a framework to discuss the operational changes such as those described for the TRESSPASS air BCP pilot.

The Dutch pilot was planned to take place at the Schiphol Airport in Amsterdam. This airport is one of the biggest European airports for passengers and cargo traffic and it is registering an increase in traveler numbers, take-offs and landings. According to the Royal Schiphol Group's annual report, 499,444 take-offs and landings occurred in 2018, compared with 450,697 in 2015. 71.1 million passengers transited the airport in 2018, an increase of 12.9 million compared with 2015. Currently, approximately 200,000 passengers per day travel through the airport.

The airport consists of several actors: private and public partners, defense/RNM, governmental and administrative bodies, the general public and the media. This complexity poses challenges in terms of coordination and tracking of ongoing issues.[4] Additionally, although the pilot tests were eventually conducted at Leylstad Airport and RNM Barracks, the core challenges, concerns, and research questions structuring the pilot tests were developed from the context of the Schiphol airport BCP.

The diagram below is taken from D6.2 and shows the CONOPS As-Is of Schiphol Airport, for the AIR BCP pilot.

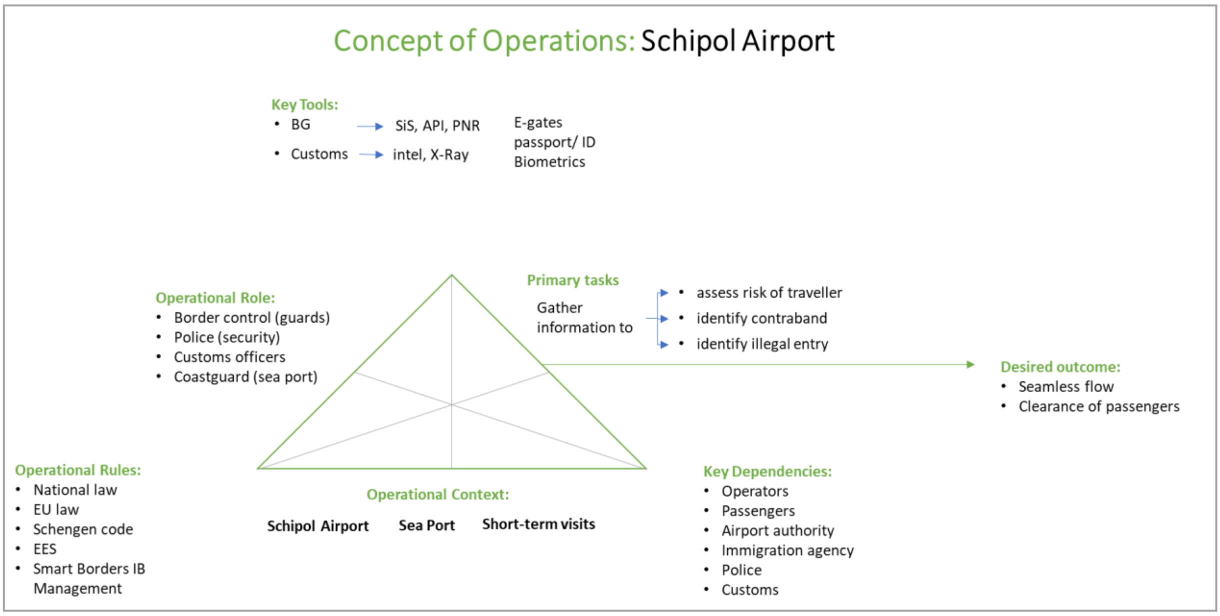


FIGURE 3: CONOPS SCHIPHOL AIRPORT

The analysis of the situation at the Dutch air BCP and the interaction with the end-users has pointed out some core challenges:

- Find a balance between mobility and security, considering that the flow of passengers will increase while the available personnel will be scarce:
- speed up the control process without compromising the check and jeopardizing the security of operators and passengers.

[4] TRESSPASS. *D6.5 Evolving CONOPS framework (2nd version).*

- enforce border crossing points, as terroristic attacks keep threatening the national security of EU MSs.
- push towards the digitalization of processes at the airport, including the border checks procedure.

The main research objective of the air pilot was to evaluate the performance and potential of the TRESSPASS RBBM concept and system in terms of the overall project KPIs, (and how to strike a balance between them) of effectiveness, efficiency, flowrate, and feasibility.

The greatest focus was given to the human role (i.e., the border guards decision making process), thus exploring how the introduction of the TRESSPASS solution would affect and change their operational environment and its effectiveness in supporting the decision-making process.

Based on the abovementioned challenges some expectations emerged amongst the end-users:

- Maintain at least the same or increased security level by adding on site behavior data, analysis methods, data from external systems like VIS/SIS/PNR etc.
- Enhance border guards' performance by leading them focus on the risk and not on the typical process.
- Reduce the false positive labels as the system matures and the real threats are more accurately identified.
- Improve the traveler's experience by assigning them to control checks accordingly to their specific risk level, and thus facilitating the process for bona fide travelers.[5]

2.1.1 CONOPS consideration post- pilot

Taking the above air BCP challenges, Dutch end-user expectations and project KPIs into consideration and using a human centered activity theory perspective, the TRESSPASS CONOPs described the proposed changes that would occur to the operational environment from three perspectives: technological, procedural, and organizational, with the introduction of the TRESSPASS system. These changes were updated post-piloting phase and the main overall conclusions of these changes for the Dutch air pilot are described below.

I. In terms of technologies, the use of self-service registration gates is expected to increase, thus automating border checks for all travelers (not just EU citizens as today). Consequently, staff personnel performing manual checks at this stage will likely decrease.

II. From a procedural perspective, border guards will reduce direct contact with travelers – due to the introduction of the registration kiosk. DPM interviews will be conducted only for those passengers assessed as a potential risk. Due to data retrieved from multiple sources, the border guards conducting this interview will be supported by additional information.

III. These changes will inevitably lead to organizational adjustments in terms of BG workload, the way of working and will require ad *hoc* training for the border security operators.[6]

---

[5] TRESSPASS website. *The Dutch use case: Airport border crossing point*. https://tresspass.eu/Pilot-1

[6] TRESSPASS. *D6.6 Evolving CONOPS framework (Final).*

The end-users strongly support the integration of TRESSPASS solution with the existing system. However, they also recognized the significant operational changes that the introduction of the TRESSPASS solution will bring. Therefore, it is recommended that such transition will be supported by adequate training and that a balance between the technologies and the human operator will be maintained, as the border guards want to retain control of final decision making during the border management security procedures.[7]

2.1.2 CONOPS Conclusions

The conducted TRESSPASS air pilot aimed to address core challenges of the BCP, namely finding a balance between mobility and security, increasing the flow rate across the BCP without compromising security and whilst maintaining the integrity of the BCP and attempting to improve the digitalization of border processes at the airport.

Certain expectations emerged from the pilot design and air team, mainly surrounding the desire for improved performance regarding both BGs decision making and the overall BCP security performance whilst enhancing the traveler experience.

From an operational perspective, overall findings demonstrate an increased automation of the air BCP, with all passengers, including extra-EU to now use the self-service gates and for BGs to have less direct interaction with the bona fide traveler.

These system changes will have varying impacts upon human factors, such as the aforementioned workload and the nature and way of working for the BG. Additionally, if implemented, the TRESSPASS system changes will have an impact upon the traveler and their experience of crossing the BCP, with the nature of their interaction, becoming more machine and less human based. It remains to be determined if this change will improve the overall passenger experience.

These changes will have a knock-on effect broader impact to the wider system dependencies, such as, private, and public air BCP partners, governmental and administrative bodies including human resources and even public opinion and the media perception. It is evident that these system changes will require adequate and specialized training for relevant personnel and that the border security operator’s decision is retained as the foremost and final decision in the border checking process.

---

[7] *Ibidem*.

2.1.3 End Users' Views

The RBBM evaluation through survey research (comparing the systems per question) shows that the border guards appreciate the usability of the RBBM way of working into a large extent. Additionally, they find the new systems less stressful to work with. RBBM could replace the current way of working for the better.

From border guard point of view the RBBM concept is more convenient for non-malicious and equal for malicious travellers. They find RBBM from a privacy point of view accountable and proportional and appropriate from an ethical point of view, including the use of individual social media information.

By evaluating RBBM with the focus group approach not completely new but additional and supporting observations are found. The most distinctive are mentioned below.

- When sorting out possible mala fide suspects do not create a clearly visible row for them, this could lead to a "walk of shame" or the impression of ethnic profiling. Also do not forget there will be non-malicious samples in there too.
- However, the border guards said they mainly based their decisions on the DPM interview they found the extra information with plusses and minuses useful. They were not overwhelmed by the amount of information, but they would like to have the more important risk indicators highlighted. They think it is better to have the information during the DPM interview than before, to prevent biases.
- The border guards trust the RBBM system in sending mala fide assessed travellers to the manual desk.
- The border guards trust to RBBM system in sending bona fide travellers to the ABC-gate and a sample portion (10-30%) to the manual desk.
- The border guards have strong hesitation to pass the majority of neutral assessed travellers (not tested but this would be the largest category by far). There is not yet enough trust in the RBBM system; they miss the non-verbal, visual and DPM narrative of the traveller where they are now used to rely on. Beside this they don't trust yet the authenticity checking of the traveller and his document at kiosks and ABC-gates.
- One expects RBBM will significantly increase the chance of being caught when combined with the current way of working (decision support at manual desk for DPM with all travellers). They also expect the processing to get increased with the risk that some border guards might become sloppy.

From the border guards' point of view the RBBM concept is more convenient for non-malicious and equal for malicious travellers. In case of creating a separate row for mala fide assessed travellers and samples a "walk of shame" or the impression of ethnic profiling need to be prevented.

The border guards expect a higher chance of being caught but have a strong hesitation in automatically passing neutral assessed travellers to the ABC-gates without seeing them in person. They miss the non-verbal, visual and DPM narrative of the traveller where they are now used to rely on. There is already enough trust in the system for mala fide and bona fide assessed travellers. A sudden switch to an RBBM situation where the majority of people will not see the border guard anymore, cannot be expected. This will need a change of mindset and a gradual transition where first more experience and trust can be obtained.

2.2 Threat Scenarios

The threat use case for testing RBBM in the air pilot was the irregular migration.
Irregular migrants are non-EU nationals who don't fulfil, or no longer fulfil the conditions of entering to the territory of a Schengen State as set out in the Schengen Borders Code (Regulation 562/2006), or other conditions for entry, stay or residence in an EU State.

The use case for the pilot was narrowed down to detect travellers with the possible intention to irregular migration but in possession of valid travel documents and visa (if needed). With experienced border guards, intelligence officers and subject matter experts 34 mala fide and 11 bona fide indicators were distracted from practice and implemented in the TRESSPASS system (Data Fusion Analytics) for the risk calculation (Dynamic Risk Assessment System), using the TRESSPASS Risk Assessment Methodology (TRAM). The indicators are confidential.

Irregular migration was seen as a use case for the pilot but not as the only threat TRESSPASS should be suitable to work with for the RNM.

2.3 TRESSPASS Technology (Components/Sensors, Infrastructure)

The list of the components applied in the air pilot included: TRA, VTC/RTBA, WI, C2, SPA, LSI, DRAS, DFA, DMS.

2.3.1 TRA

Particularly for the first air pilot, the TRA application was used by 21 participants:

- No technical issues observed using TRA in the air pilot, while the application was tested with three different types of devices (laptop, tablet, desktop).
- The time to complete the TRA account creation phase and the trip pre-registration phase were short, which is important for a user application not tiring the user, while targeting regular use of the application. In detail, most participants (66.67%) were able to complete the TRA account creation phase in less than 10minutes (note that this phase is executed only once per traveller). Moreover, 76.20% of the participants needed less than 10 minutes to create a new trip in the TRA application (half of them needed less than 5 minutes), while 95.25% needed less than 15 minutes.
- In general, the users could easily insert information in the application. 85.71% of them found easy to select destination countries, while all participants had no issues while providing information related to transit countries or used vehicles. Finally, the vast majority of participants (95.23%) was able to provide travel document information without difficulty.
- Most of the participants (61.91%) could easily navigate using the user interface of TRA.
- In general, according to the collected feedback, TRA seems to be a promising application for use in actual border crossing conditions since most participants (76.19%) stated that they would use TRA application to accelerate the border crossing, 52.38% stated that they believe that TRA application would improve security, while 66.67% stated that they would use TRA regularly.

2.3.2 VTC/RTBA

Although VTC does provide promising results the application of VTC in a risk-based framework in the future has to address the following issues

- **VTC is not a ready-to-use component:** Besides a considerable calibration effort, also a lot of effort is needed to fine tune the system.
- **Detection of behavior needs proper camera placement:** For instance, row switching was impossible to recognize by VTC/RTBA, because the camera observes the row from the front.
- **Lighting conditions have to be considered:** Lack of light as well as backlight hinder re-identification. Even a human in the loop is not able to recognize if only a silhouette is visible.
- **On-line VTC application is hindered by security and privacy concerns**: These concerns did not allow a live connection with the existing CCTV Infrastructure in the pilot.
- **Human in the loop support would be very valuable**: Border guards would like to interact with an on-line VTC application.
- **A tool supporting the automatic determination of technical and operational factors for the application of VTC would be very valuable as a work around to avoid (too much) manual effort in the future**: Such a tool would support the automatic external calibration during installation and after maintenance, such as the automatic and continuous monitoring of lighting conditions in relation to operational margins, and the automatic acceptance testing of camera placement. We refer the reader to den Hollander et al.[8] for an overview of state of-the-art technologies supporting the feasibility of the implementation of such a tool.

2.3.3 WI

Web Intelligence was deployed to facilitate the information gathering of the involved passengers. For the needs of the scenario execution, the pilot owner facilitated the creation of synthetic social media (Facebook) profiles, which were populated according to the profile of the participating mala fide passenger personas. As soon as the travel details and social media information of each of those passengers was entered in TRA, WI was triggered to assess their digital footprint by calculating the risk indicators raised for their profiles. The risk indicators calculated by WI in this scenario were AIR RI-007, AIR RI-008, AIR RI-012 and RI-014.

[8] den Hollander, R. J., Bouma, H., van Rest, J. H., ten Hove, J. M., ter Haar, F. B., & Burghouts, G. J. (2017, October). Automatically assessing properties of dynamic cameras for camera selection and rapid deployment of video content analysis tasks in large-scale ad-hoc networks. In Counterterrorism, Crime Fighting, Forensics, and Surveillance Technologies (Vol. 10441, p. 1044108). International Society for Optics and Photonics.

2.3.4 C2, SPA and Back End Components (LSI, DFA, DMS, DRAS)

During the Dutch Pilot experiment with the border guards, the current Netherlands border guard system was compared with the TRESSPASS C2 and SPA systems. The systems were evaluated after the border guards assessed and decided on 10 travelers with each system. The underlying systems (LSI, DFA, DMS, DRAS) were not directly visible for the end-users and therefore also not evaluated. The evaluation of the C2 and SPA compared to the current system in use was done by answering one or more questions about 7 constructs (dimensions) of the usability of the system. This was done with statements where they could react with the 5-point Likert scale: completely disagree – completely agree.

The table below presents the outcomes of the surveys per system.

TABLE 2 SYSTEM EVALUATION RESULTS

Construct (Dimension)	System	N	Average	Std. Deviation
Perceived usability of system	Current	30	3,18	0,83
	C2	30	3,88	0,69
	SPA	30	4,07	0,65
Perceived usability of the information in the system	Current	30	3,71	0,51
	C2	30	3,77	0,49
	SPA	30	4,02	0,39
System understanding	Current	30	4,12	0,67
	C2	30	3,45	0,75
	SPA	30	4,22	0,36
System performance	Current	30	3,27	0,86
	C2	30	3,45	0,75
	SPA	30	3,80	0,47
Trust in automation	Current	30	3,40	1,04
	C2	30	3,33	0,80
	SPA	30	3,73	0,64
Perceived ethics	Current	30	3,67	0,88
	C2	30	4,00	0,87
	SPA	30	4,13	0,57
Intention to use the system	Current	30	2,83	1,05
	C2	30	3,70	1,02
	SPA	30	4,20	0,55

The border guards appreciated the usability of both the C2 and SPA systems as high. Regarding acceptance, the border guards would accept the RBBM systems as the standard way of working in the future, and the SPA is most appreciated. The border guards trust the C2 system almost as equal than the current system (medium/high) and the SPA system a little more (high).

2.4 Data and Risk/Bona Fide Indicators

In total 45 risk indicators in travel intelligence and border operations, were derived from several workshops by RNM and TNO experts. The indicator set was very well recognised by the tactical and operational border guards, and they represented their implicit way of current working at the border.

In the preparation phase TNO and RNM created a data set for 300 mock-up travellers. This mock-up data was used by ICTS to generate fake passports for the travellers. The passenger data was stored in the TRESSPASS system components. (ICTS, C2, DFA, VTC, RTBA).

For the purpose of the pilot three sets of 10 travellers (total of 30 travellers) were created. For each traveller a profile and a story with bona fide and mala fide risk indicators and scoring was created and filled in the TRAM implementation tool (the Excel tool that represented the DRAS implementation for the Dutch pilot).
For each traveller complete data (flights, names, id, dates, visa, way of paying, etc, etc) was created in the LSI databases (VIS, ETIAS/TRA, EES) and the RTBA component (results of the risk indicators based on their behaviour on the airport). In this way it was possible to operate the LSI, DFA and DRAS components so the right risk assessment and information could be displayed in the C2 and SPA applications by the border guards during the pilot. They consulted the system, searching by name or ID number the person in the application and getting the results.

In order to be able to test the decision-making of the border guards in the current situation, with the C2 and with the SPA, false negatives, false positives, true negatives and true positives were created.

The travelers with the scoring indicators and the story (text) were judged by specialists (as well in international border control operations and intelligence) for creating a “ground truth” based upon their knowledge and experience.

2.5 Risk Assessment (Concept & Methods)

During the pilot experimentation border guards have observed the moment that a traveller is standing at the manual desk and the border guard had to collect information and decide upon. It was seen a very high detection probability for malicious travellers, same as in the current systems. There is a medium possibility of false alarms for non-malicious travellers that are stopped by the border guards, which could lead to inconvenience at non-malicious travellers. There are almost no false alarms caused by malicious travellers not being stopped.

The RBBM systems, same as the current systems, are supporting the reasoning for decision making in a sufficient way, where the SPA could lead to more suspicion. The border guards had confidence in their decision making and did not find it difficult regardless the system.

The appreciated usability (overall score of the constructs) of the RBBM systems by the border guards was a little higher than the current system, including the perceived ethics. They would have the intention to use such systems in the future.

2.6 Research Questions and Key Performance Indicators (KPIs)

	#	Research questions
Effectiveness	1	*What is the effectiveness of the TRESSPASS RBBM concept (and corresponding available information) in comparison to the current rule-based approach (and corresponding available information)?* **Appreciation: Very high for the decision making of the border guards. Equal to the current system.**
	2	*What is the probability of a false positive (i.e. people that are bona fide are send to the manual desk) and a false negative (i.e. people that are mala fide are not send to the manual desk)* **Appreciation: There is a medium probability that border guards stop non-malicious travellers and a very low probability that they will not stop the malicious travellers.** **This is good in fighting the threat but inconvenient for some non-malicious travellers.**
	3	*To what extent can the information of the TRESSPASS system potentially help to better execute the process of mitigating the threats handled by your pilot (irregular migration, smuggling etc)?* **Appreciation: The current and the RBBM systems are supporting the reasoning in an equal sufficient way. The border guards have sufficient confidence in taking their decisions regardless the system. The border guards do not find it difficult to take their decisions regardless the system.**
Flow rate	4	What is the effect of the TRESSPASS RBBM concept on the flowrate of different categories of travellers in comparison to the current rule-based approach? **Appreciation: The ultimate added value of implementing RBBM is speeding up the traveler flowrate in border crossing with a higher change of getting caught and with less border personnel. It is not the intention to add an extra system and make it more complex and expensive. Therefore, the TRESSPASS simulations had to be carried out and analyzed carefully to get more insights in what the impact of RBBM is in expected efficiency, personnel and resources. The results of the simulations are in the deliverable 8.5 and show clear signs of a positive business case, which may vary depending on the implemented process (where a higher or lower percentage are allowed for self-service border crossing).**
Efficiency	5	What are the resources needed to perform the RBBM-concept to achieve a desired level of flow rate and/or effectiveness? **Appreciation: The conclusion of the Dutch pilot was that the RBBM equally or better supports the border guard in his work in various domains. The results of the simulation showed possible very large efficiency gains. For that the border crossing need to be automated in such a way that a majority of the travelers will not see the border guard anymore. For that the trust in the RBBM must grow further and call for a change of mind and a gradual transition where more experience can be obtained.**

Acceptance	6	*To what extent perceive the participating border guards the TRESSPASS RBBM concept and system as usable?* **Appreciation: From the point of view of privacy and proportionality the border guards appreciate the RBBM concept far more than the current way of working. The usability of the C2 and SPA systems is high. More than the current system. From a viewpoint of acceptance, the appreciation was very high. The majority would replace the current way of working for it.**
	7	*To what extent perceive the participating border guards the TRESSPASS RBBM concept and system as trustworthy?* **Appreciation: The border guards trust the C2 system almost as equal than the current system (medium/high) and the SPA system a little more (high).**
	8	To what extent perceive the participating RNM-border guards the TRESSPASS RBBM concept and system as efficient? **Appreciation: High. More than the current system.**
	9	What are the implications of a RBBM concept for the personnel at strategic, tactical and operational level? **Appreciation: It is not realistic to suddenly replace the current systems and processes at once, it is too risky, technical unfeasible and could probably count on resistance of the current border guards and society. So, a gradual introduction is advised where increasing RBBM parts will be added as decision support in the current way of working. Starting with single extra indicators in the current primary border guard system and expanding this to complete profiles, experience and trust can be built and lead the way to seamless flows where bona fide assessed travelers can pass the border in self-service concepts without seeing any border guards. To support this implementation process it is advised to have a living lab with a simulated BCP where these gradual steps can be experienced by the border personnel at strategic, tactical and operational level, experimental user interfaces can be tested and impacts on the efficiency and effectiveness can be calculated.**
Balancing	10	What is the trade-off between the effectiveness, flowrate and efficiency of the TRESSPASS RBBM concept? **Appreciation: Balancing different KPI’s**

2.7 Training

Instructions for the pilot's participants to understand the TRESSPASS (RBBM) concept as needed for the experiment were given through presentations and discussions. This was an overall instruction on the level of secondary vocational education and up and till the level of using the risk indicators.

The instructions for the usage of the C2 and SPA were practical based on a short presentation and some hands-on testing experience with travellers. The use of the applications seemed to be easy with little instructions.

The insights gained about the e-learning modules for the C2, and SPA module are as follows:

- Both e-learning modules (SPA, C2) were okay on the functionalities which were addressed
- Border guards experienced the questions at the end of both e-learning modules as difficult
- When trainees pushed the 'enter' button to go to the next slide, the 'click' sound was found distracting and it is recommended to leave that out in all e-learning modules. (Considering 10 people in a room all using the e-learning module at the same time, lots of 'clicks' are heard all the time)
- Additional instructions / exercises were orally given for the C2 and SPA modules (as it was also recommended).

It was found out that the additional instructions/exercises in the 'classroom' were much more efficient and effective than the e-learning module itself, the same feedback came also from the border guards. It was thought that this is also due to the fact that we only left the relevant functionalities for the Dutch pilot in the e-learning module, so it was only a couple of instructions they got via the e-learning module.

It has been concluded that for the next pilot days we only wanted to give oral instructions / exercises for the C2 and SPA module, this also gave some extra time within the daily program as we found out that our agenda was quite ambitious for 1 day.

2.8 Simulations

Simulations of an airport risk-based BCP have been performed with both the iCrowd simulator and the FHG-simulator. In both simulators the situation of the arriving travellers from Schengen and non-Schengen countries at Schiphol Airport was taken as the base. Travellers on these flights have nationalities from non-Schengen as well as from Schengen countries.
Schiphol Airport is a representative situation for international flights in relation to the combination of available space, maximum number of available manual desks for the border guards and flow rates. The input parameters as used within the simulations are based on educated guesses and do not represent the real actual values of the current Schiphol Airport situation.

2.8.1 Simulation Research Questions

The main research question for the air pilot simulations was:

What is the effect of the TRESSPASS RBBM concept on the flowrate of different categories of travelers in the context of irregular migration, in comparison to the current rule-based approach?

Expected effects were assessed in terms of:

a. Queue lengths and queuing times for the travellers
b. Resources needed to handle different numbers of incoming passengers. (Resources can be number of border guards, enrolment kiosks, ABC gates)

The covered TRESSPASS Key Performance Indicators are:

- Checking time per traveller (KPI-FLW-01);
- Average cumulative checking time per traveller per traveller category (PI-FLW-01-01 & 3).
- Minimum cumulative checking time per traveller per traveller category (PI-FLW-01-02 & 4).
- Number of travellers per unit of time (KPI-FLW-02-01 & 02).
- Waiting time per traveller (PI-FLW-03-01 & 02).

2.8.2 Simulation Research Approach

The simulation approach consisted of running simulations under two main conditions.

1) Maximum flow rates are determined for the rule-based situation and for the risk-based situation, both for a fixed number of five border guards.
2) The minimum number of border guards are determined for the rule-based situation and for the risk-based situation, both for three typical different flow rates of a typical (pre-covid) situation.

All risk-based simulations have been run under three sub conditions:

1) Poor performance of the risk-based assessment system.
2) Medium performance of the risk-based assessment system.
3) High performance of the risk-based assessment system.

We expect a high-performance system to have the combination of very good indicators, profiles and technology. A high performance then means that the risk-based assessment categorises most travellers as either bona fide or mala fide and that the number of travellers in the categories ‘neutral’ and ‘unknown’ are very small (See Table 3). These performance numbers are based on educated guesses as we do not have available detailed performance results. We expect the reliability of risk-based border control to grow in the future.

TABLE 1 PARAMETERS FOR THREE PERFORMANCE LEVELS OF THE RISK BASED ASSESSMENT SYSTEMS (FRACTIONS OF TOTAL NUMBER OF TRAVELLERS)

	Low performance	Medium performance	High performance
Bonafide	15%	40%	80%
Neutral	55%	45%	13%
Malafide	5%	5%	2%
Unknown	25%	10%	5%

2.8.3 iCrowd Simulator Set-Up

Within the iCrowd simulator an arrival hall has been created, based on an already existing model of a Greek airport, for a risk-based border management process. The arriving passengers enter the hall either from the left side or the right side from the different concourses and gates (Concourse and gates have not been created in the simulation model). They enrol themselves at the enrolment kiosks where each passenger receives a directive to go either directly to the ABC-gates (for most of the bona fide and neutral assessed passengers) or to go first to the manual desks (for the passengers that are assessed unknown or mala fide) to have a duration, purpose and means interview (like in the current rule-based situation).

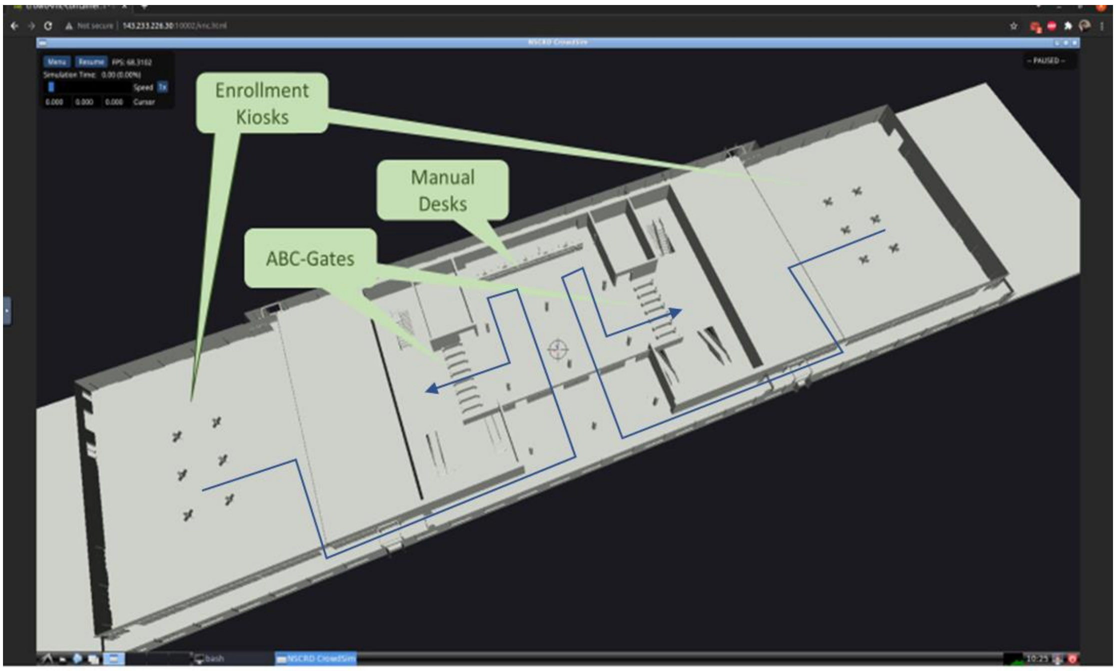


FIGURE 4 AIR PILOT BCP LAYOUT IN THE ICROWD SIMULATOR

A small part (samples) of the bona fide assessed passengers is still sent to the manual desk. This part is called the random selection. Passengers may be mistaken and go directly to the ABC-gates while they were ordered to go first to the manual desks. In this case the ABC-gates will stay closed, and a walking border guard equipped with an SPA (Secure Passenger Application) will interrogate the passengers after which they may enter through the ABC-gates or are sent to the second line.

2.8.4 FHG Simulator Set-Up

Within the Fraunhofer (FHG) simulator also both the rule-based border control processes as well as the risk-based border control processes are modelled. The rule-based set-up consists of four items (See Figure 5):

- Border Guard at the Manual Desk (bg)
- Random Selection (rs)
- ABC-gate (pass)
- Second line check

As input for the simulation, two types of travellers were used (non- malicious and malicious). The border guard (bg) process determines for each traveller the risk assessment (RA) (either trusted=green, or suspect=red). Trusted travellers then go directly to the ABC-gate (pass). The random selection has not been used in the rule-based simulation. Suspect travellers are directed to the second line check where the simulation stops. The second line check was not performed.

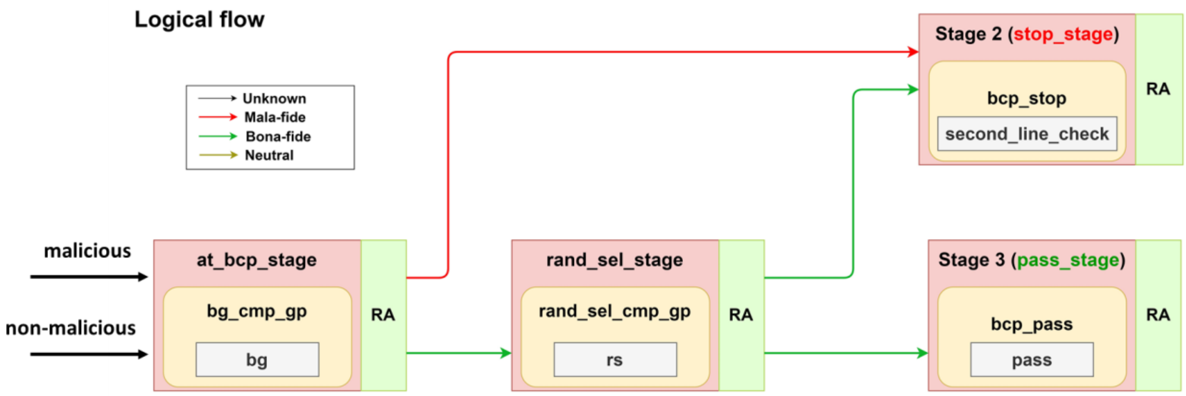


FIGURE 5 FHG-SIMULATOR SET-UP FOR THE RULE BASED BORDER CONTROL

In the risk-based simulation the TRESPASS components have been modelled (See Figure 6). (EK=enrolment kiosk, BDK=border direction kiosk, BG=border guard, SPA=security personnel application, …).

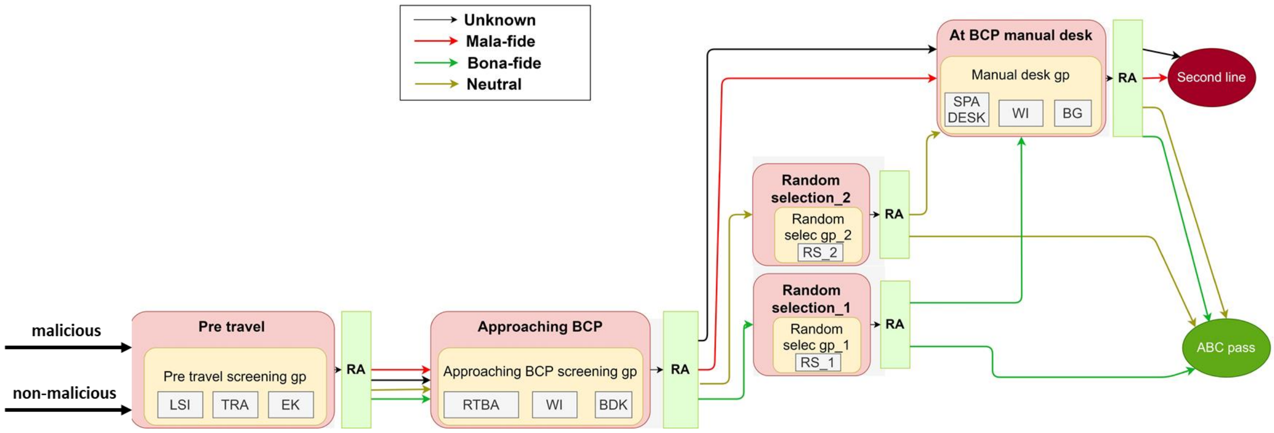


FIGURE 6 FHG-SIMULATOR SET-UP FOR RISK BASED BORDER CONTROL

As input for the simulation runs, we wanted a predefined subdivision into bonafide, malafide, neutral, unknown categories instead of a division between malicious and non-malicious. In order to reach this, we took as malicious the group of malafide plus the unknown travellers. The neutral travellers were simulated by enlarging the proportion of random selection non-malicious travellers that are sent to the manual desk. (RS-1).

2.8.5 Simulation Conclusions

1. The outcomes of the simulations clearly indicated that the risk-based border management concept in an airfield setting can have a positive effect on both the queue lengths and queuing times for the travellers while the average time that a border guard spends at malafide and unknown assessed travellers can increase.
2. The average times that travellers will spend at the BCP decreases significantly in a risk-based setting.
3. Bonafide assessed travellers need 2-3 minutes on average at the BCP, independent of the flow rate, while malafide and unknown assessed travellers need considerably more time. In the order of 8-15 minutes on average, dependant on the number of border guards at the manual desk.
4. The number of needed border guards at the manual desk in a risk-based setting remains equal for low and medium flow rates but much less additional border guards are needed to handle (very) high flow rate conditions.
5. As a rough estimate we state that the risk-based system should perform above a certain threshold. Low performance systems did not show improvements in flow rates and needed resources. As a minimum threshold we can say that the RBBM system should send less than 30% of the travellers to the manual desk for an extra-long interview to outperform the rule-based system. To determine this threshold more precisely, it is recommended to start a pilot at an airport with both the rule based and RBBM and measure both systems' performances in a real-life environment.

2.8.6 Simulations Lessons Learned

For the air-pilot simulations both the iCrowd and the FHG simulator have been used. These simulators differ considerably in their user interface and technical possibilities. The iCrowd simulator has a graphical user interface that shows the airport, the travellers on the airport and that offers via the interface all necessary options to run simulations without much dedicated education and training. After two 2–3-hour training sessions it was good possible to operate the iCrowd system.

The FHG simulator lacked such a graphical user interface which makes it more difficult to operate. For the air-pilot simulations, Fraunhofer's personnel ran the simulations which resulted in a proper way to obtain the simulation results.

For both simulators various improvements can be suggested to make the products more mature. But in general, both products showed sufficient maturity for the air-pilot simulations.

The air-pilot simulations focussed on flow rates and needed resources (number of needed border guards). This left one capability of the FHG simulator unused, namely the determination of effectiveness of the TRESSPASS assessment decisions. This was not taken as part of the air-pilot simulation because in order to obtain valuable results, a sufficient large set of high-quality realistic traveller profiles were needed for that purpose, which were unfortunately not available and would need too much effort to create within the context of the time and budget schedule of the air-pilot.

More in general, for all simulation outcomes, the added value of the outcomes depends on the quality of the input parameters. For the air pilot we were able to provide a set of input parameters that were a mix of statistical data of Schiphol airport for the rule-based situation and a set of educated guesses parameters for the risk-based situation. The reliability of these latter set leaves room for improvement, e.g., by data collection of risk-based border control experiments in possible follow-up projects.

2.9 Overall Pilot Lessons, Limitations and Obstacles

2.9.1 Recording Day

2.9.1.1 Overall Lessons/Conclusions

- Despite the COVID restrictions, the Recording Day was successful and met the primary goal of collecting realistic traveller video recordings that were input to the VTC and RTBA components.
- It is advised to consider follow-up VTC/RTBA tests in a more operational setting such as using actors that simulate abnormal behaviour at an operational airport with regular passengers.

2.9.1.2 Limitations

- The original plan was to have 300 volunteers but due to COVID restrictions only 30 people were allowed in the airport building (volunteers + TRESSPASS people + personnel). Therefore, the number of volunteers was restricted to 24 volunteers and all volunteers were asked to walk 12 times over the airport.
- Due to the COVID 1,5 m social distance rule, the travellers showed less natural behaviour during their walks over the airport. Furthermore, only 30 persons were allowed at the airport due to COVID regulations which is too few for a realistic traveller occupation. COVID restrictions prevented also physical preparation meeting at the airport for all involved partners which hindered the good understanding of the test details.
- No check was performed as to how precisely the volunteers followed their instructions.

2.9.2 Experiment with border guards

2.9.2.1 Overall Lessons/Conclusions

We can conclude that the RBBM systems (C2 and SPA) are perceived just as useful as the current system (BCS) for the border guards in terms of high detection probability, low to medium false alarm rate, good decision support and confidence. This is for the situation when RBBM systems are replacing the current system but within the same process.

The appreciated usability of the RBBM systems by the border guards is higher than the current system, including the perceived accountability, proportionality and ethics (including the use of traveller social media information). Additionally, they find the new systems less stressful to work with. They would have the intention to use such systems in the future.

The border guards continued to hold strongly on the Duration, Purpose, means conversation but find the extra information with plusses and minuses useful. They are not overwhelmed by the information.

2.9.2.2 Limitations

There are limitations to the research approach and the manner that the experiment is conducted (which are partly the effect of the general COVID restrictions to the pilot). These have implications for the validity and reliability of the results of the experiment, and therefore need to be considered when deriving conclusions on the results of the experiment.

Internal validity

The internal validity of an experiment is concerned with the extent that hypotheses can be tested adequately (Greenwald, 1976). Although the pilot is mainly exploratory research to test the potential of RBBM, we are still interested in whether there are (significant) differences in the performance of the border guards under different conditions. Therefore, it is important to reflect on the internal validity of the experimental set-up and to take factors into consideration that could impact the results of the experiment when comparing the 3 different conditions.

The sequence of the experimental conditions during the experiment days were fixed: first the current system was used, second the C2-system, and third the SPA-system. We opted for this sequence as this deemed necessary for the participants to understand the experimental procedure by first conducting the experiment task with their current system, and then to progress to the C2 and SPA conditions. Although this seemed to be most fitting for the learning curve of the border guards during the experiment, this could through learning effects (i.e., the border guards improve their performance during the day as they understand the procedure better through more repetitions). Greenwald (1976) describe this as 'practice' effects (pp. 316 – 317), or 'carry-over' effects (p. 318). Performance could also be influenced through fatiguing effects (i.e., the border guard's performance decrease during the day as a result of repetitive tasks in an environment they are not accustomed with) (Cherry, 2020[9]).

Another dimension of an experiment's internal validity is its construct validity: "[w]whether the inferences from the data are valid for the theory (or constructs) the researcher is evaluating in a theory testing experiment" (Morton & Williams, 2010, p. 256). For this experiment, the items for most of the tested constructs are based on existing scales to enhance the construct validity. However, some items and/or constructs are developed for the purpose of this air pilot, for which not theoretical validation preceded its use. Also, some constructs were measured through a single item. This might reduce the construct of the constructs. However, given the exploratory character of the experiment this is not problematic (as no theories are tested) but nevertheless should be taken into consideration when interpreting the results.

Ecological validity

Ecological validity can be defined as "whether the methods, materials, and settings of the research are similar to a given target environment" (Morton & Williams, 2010, p. 265). The experiment took place in an environment that is not representative for the average operational setting of the border guards at Schiphol Airport. Therefore, the results of this experiment should be interpreted with care.

Another aspect which influences the ecological validity of the experiment, is the fact that the border guards needed to work with fictive travellers (with corresponding traveller stories) that are developed by the pilot team by consulting intelligence specialists of the RNM and tested during the dry run. Although the traveller profiles are deemed as realistic, this only emulates the real information that the border guard receives during their border management tasks (e.g., the posture of the traveller, way of speaking, etc.). Hence, the border guards work with less information during the experiment than in their regular working environment.

[9] Within-Subject Design Experiments (verywellmind.com)

Manipulation check

An important mechanism for an experiment to make statements about causality is to conduct a manipulation check, i.e., “a survey or other method used to check whether the manipulation conducted in an experiment is perceived by the subjects as the experimenter wishes it to be perceived” (Williams & Morton, 2010, p. 108). The purpose of the air experiment is to observe if the other interfaces with additional information (i.e., risk indicators and risk scores) helped the border guards to make better decisions about travellers. However, due to the novelty of the C2- and SPA-interfaces, de pilot team could not fully control for the manner how the risk indicators and -score are interpreted and used in their decision making. Due to the complexity of decision making and the exploratory objective of the pilot, this element (and corresponding manipulation check) is not incorporated in the research approach. This aspect should be addressed by future research after the border guards are more accustomed working with the new systems (e.g., after more extensive training). During the focus groups, this decision-making process was discussed and potentially provides starting points for future studies.

Regarding the survey research the following limitations need to be considered:

Internal validity

The constructs and corresponding items that are used in this research approach are based on the interests of the pilot team, and the research questions and KPIs that are of interest. Therefore, the questionnaire was mainly exploratory and to provide qualitative insights in how the border guards perceived the RBBM concept. Therefore, the constructs and items are not validated by previous research and therefore should be mainly used as measures that helps the pilot team to evaluate the risk-based border management concept.

External validity

The participating border guards can be regarded as a *convenience sample*[10]. Therefore, the results and perceptions are not necessarily representative for the whole population (i.e., all RNM border guards). This should be taken into consideration when interpreting the results and deriving the conclusions based on the data.

[10] https://en.wikipedia.org/wiki/Convenience_sampling

3 POLISH PILOT: LESSONS LEARNT

This chapter breaks down in sections each aspect of the Polish Pilot assessment, the lessons learned, the experience gained, and the obstacles observed by the pilot's stakeholders upon the completion of the pilot's' execution, the collection and processing of the external and internal feedback via the evaluation tools (questionnaires etc.) used and the relevant pilot deliverable report (D8.3).

3.1 Operations

3.1.1 Recap of CONOPS As-Is

As mentioned in section 2.1, the TRESSPASS CONOPS was and is concerned with detailing the operational picture of the TRESSPASS system for each BCP modality in terms of the human factors perspective. Developed from site visits and end user workshops, the third version of the CONOPS guide (D6.6) will provide the final operational description for TRESSPASS system implementation, including the findings from the Polish team land pilot and will present these findings in comparison to the idealized picture.

The importance of the human factor in terms of operations at the land border can be understood from the Polish land pilot findings, which has offered the possibility to test TRESSPASS solution against a land border crossing point.

The land BCP at Terespol, a town located 190 km east of Warsaw, was set as location for the pilot: the BCP includes cars and buses, while the truck terminal is located in Kukuryki (12 km north of Terespol). However, to facilitate the pilot implementation, a terminal for both buses and trucks was temporally arranged in Terespol. The Polish Border Guard (PBG) are the main operators, in charge of both passenger and cargo checks.[11]

The below swim lane diagram captures the 'as-is' or baseline current system process flow (moving across from pre-travel, approaching to at-BCP phases) for lorries arriving to the land BCP at TERESPOL, as shown in the initial version of the evolving CONOPS guide, the proto-CONOPs D6.2[12].

[11] TRESSPASS. D8.3 Land Pilot

[12] TRESSPASS. *D6.2 Evolving CONOPS framework (1st Version).*

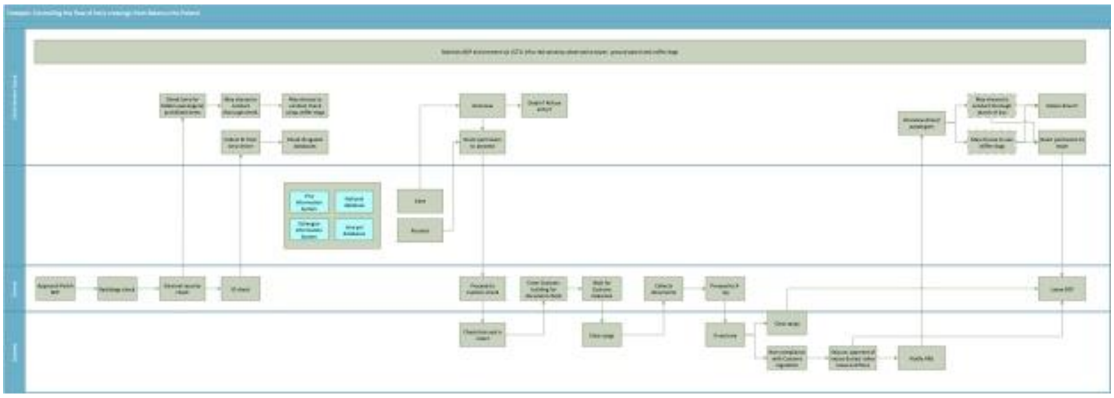


FIGURE 7: LORRY CROSSINGS AT TERESPOL BCP

The core challenges highlighted by land BCP end-users at their BCP can be described as the following:

- Heavy workload and bottlenecks. The Terespol BCP is exposed to high volumes of vehicles and people transiting each day. In addition, as customs control is conducted separately after passport control, a 'bottleneck' is created, slowing down the procedure and the overall flow across the BCP.
- Cross-border smuggling. The *modus operandi* of passengers attempting to smuggle cars and goods (illicit goods and illegal substances) are quite variegated and constantly changing, thus posing a serious challenge for the identification of such criminal activities through standard detection systems and procedures.
- Document forgeries, including i) certificates, which authorize drivers to transport dangerous materials; ii) passport stamps, which help border officers to establish how long a given person stayed in the territory of the European Union/Schengen zone.
- Inspections and Detections. Inspection of people and vehicles is carried out manually, thus being both time-consuming and expensive.[13]

Emerging from the core challenges for the land BCP, and from an operational perspective, it was possible to define three major activities performed by the land BCP operators that would potentially be affected by the introduction of the TRESSPASS system:

1. detection of falsified documents (e.g., passports, residence permits, visas, and other travel documents).
2. detection of people attempting to cross the border hidden in lorries (clandestine entries).
3. detection of third-country nationals trying to cross the green border.

The main research objective of the land pilot was to test and prove how the joint use of multiple technologies and the fusion of information coming from different data sources (i.e., web intelligence, behavioral analysis, Interpol and LEAS dataset) might ensure the operator additional valuable information to perform the risk-based screening.

The Polish pilot focused on the application of the TRESSPASS system towards the specific threat scenarios of inbound illegal entries and cross-border crimes, such as smuggling illicit goods to EU carried and hidden in vehicles.

---

[13] TRESSPASS. *D8.3 Pilot 2: Land Case – Poland.*

Based on this premise, some core outcomes for the land pilot were expected:

- better use of existing infrastructure and facilities by border/custom authorities.
- increase personnel capacities and throughput by eliminating the dependencies based on the types of travelers, their origins, entries or exit types.
- improve existing border and customs control processes with limited resources needed, increasing the throughput, and reducing queues for all the types of travelers.
- Enhance capacity to detect smuggling attempts and document forgeries.
- Increase the throughput of BCPs, especially for the custom checks of cars and trucks.[14]

3.1.2 CONOPS consideration post- pilot

The main operational findings that emerged post-pilot testing at the land BCP were divided into the technological, procedural, and organizational changes that were foreseen to occur at the land BCP with the introduction of the TRESSPASS system. As the CONOPS guide is concerned with human factors, these operational findings are described from the system end-users' perspective, including the BCP operators, traveler, and wider system dependencies.

I. Technological changes: In the current BCP system, border guards have quite a complex system in use during first line checks, for example, databases are in use which access historic data about the passenger and therefore, there was a reluctance to operationalize any additional technological components unless they were easy to incorporate and streamline with the current technology in operation at the BCP.

II. Procedural: From a procedural perspective, the Terespol end-users stressed the importance of security overflow through the BCP or passenger queue time as a main objective. This meant that procedurally, BCP operators considered the TRESSPASS processes during second line checks as of added benefit to their security procedures. For example, the TRESSPASS architecture that assists with micro-behavior risk information during second line interviews could facilitate their work procedures by supporting the decision-making process.

III. Organizational: From an organizational perspective, end-users support the additional security benefits TRESSPASS system components may afford (additional risk information) but it is desired for the system to not add any additional activities or an increase in workload to the BCP operators during the first line checks, where one border guard is managing the security checks per passenger.

From an operational perspective, land BCP end-users considered the most valuable technologies to be those that could provide them with additional information that they currently cannot possess, and therefore, could be crucial in improving the risk screening by improving the breadth of knowledge of the passenger risk score. As current system activities that the BCP operators carry out are complex and time consuming, the desire was for TRESSPASS technologies to provide this additional risk information but at a minimum increase in workload or work activities for the border guard.

---

[14] TRESSPASS website. *The Polish use case: Land border crossing point* https://tresspass.eu/Pilot-2

3.1.3 CONOPS Conclusions

The land BCP at Terespol is unique in that in receives passengers via different forms of transport, bus, train, truck, and car all travelling across the border by land. This BCP, therefore, receives a lot of passengers and often experiences bottlenecks. However, as underlined by end-users prior and during the piloting phase of the project, the effectiveness of the security screening remains a priority for the Polish Border Guard end users at Terespol, rather than the improvement of the flow across the BCP.

3.1.4 End Users' Views

The performance of the tests is assessed positively. Initially, the participating officers from PBG were adequately trained on both the TRESSPASS and the RBBM concept. Then they were practically familiarized with the account creation and travel registration in TRA. People playing travelers and PBG officers played their roles very well, with great authenticity.

It was noticed that PBG officers found it difficult to check people because they had four poorly integrated systems at their disposal (passport reader, face spoofing detector, C2, and SPA) operating on separate devices. It required a lot of mental commitment from them and distracted their attention.

The stable Internet connection is pivotal for proper operation of the TRESSPASS system.

3.2 Threat Scenarios

The pilot scenarios were considered as relevant and important by the participating PBG officers. Detection of the considered threats, i.e., illegal entry and smuggling, is crucial for normal operation of the BCPs in Poland.

Short description of original threat scenarios

The three land pilot threat scenarios focused on:

- inbound irregular migration - illegal entry of a TCN having Belarusian passport,
- inbound irregular migration - illegal entry of a TCN having Polish passport,
- cross border crime (smuggling),

with variations in different modus operandi using different transportation means (bus and trucks) and including a mix of both mala fide and bona fide traveler roles, as drivers and travelers, in each scenario. Due to organizational issues, we merged executions of scenario IA and IB into one bus arrival scenario with passengers having both Belarusian and Polish passports.

<u>*Scenario IA and IB - Inbound illegal entry of TCN having Belarusian and Polish passports*</u>

The scenario focused on detection of illegal entry of two dangerous criminals, wanted by the Belarusian police, travelling by bus from Belarus to Poland via the BCP in Terespol with a driver and five other passengers:

The driver, named Aliaksandr, was a frequent bona fide traveler.

Two people were mala fide travelers in the roles of criminals wanted by the Belarusian police, who try to escape to EU passing the borders from Belarus to Poland via the BCP in Terespol:

The first criminal, genuine name Maxim Matwiejew, has bought from a corrupted official a genuine Belarusian passport with his photo but with false personal details of Ivan Waronka (so called “**multi person**” - intellectual forgery) and thanks to this obtained a genuine visa D (Scenario IA).

The second dangerous criminal, named Mikhail, tried to escape to the EU passing the borders from Belarus to Poland using a **facial mask** based on similarity to the face photos from another person with the Polish passport (Scenario IB).

Five other travelers:

- One of them, named Iryna, had the bona fide status,
- Four of them, named Marcin, Jazep, Cichan, and Maryjana, had the neutral status.

Truck Arrival Scenario 2 - Cross Border Crime

This scenario focused on cross border crime threat category with a mala fide traveler (truck driver), part of a larger organized criminal gang with wider drug smuggling activities. The scenario was referring to an inbound travelling across the relevant land border to enter Poland by truck where the actor tried to carry a cargo of cigarettes for profit.

The scenario involved four truck drivers:

- One mala fide traveler, who tried to smuggle hidden cigarettes into Poland.
- Two bona fide travelers.
- One neutral traveler.

3.3 TRESSPASS Technology (Components/Sensors, Infrastructure)

The list of the components applied in the land pilot consisted of: TRA, WI, C2, SPA, ICTS Document Scanner/Module, TCSS, IAS/TLC, MMCAT, LSI, DRAS, DFA and DMS.

3.3.1 TRA

The Traveler Registration Application (TRA) enabled traveler pre-registration with TRESSPASS system, initial risk level assessment by other TRESSPASS applications, and continuous interaction with BCP authorities. Using this simple-to-understand application, the travelers, following their consent, were able to enter and update their personal information, enter travel related document information (such as VISA, passport) and triple information (tickets, hotel reservation, vehicles data etc.). All information collected during the pre-registration phase was retrieved by the system to speed up the border control process.

The TRA application was used by 13 participants in the land pilot.

The assessment, lessons learned, and experience gained from TRA are summarized below.

- This component was positively verified during pilot activities. No technical issues observed, while the application was tested with three different types of devices (laptop, tablet, desktop).
- This component was evaluated in terms of understandability, learnability, usability, interoperability, integration, usefulness, information, installation, performance, and legal/ethical.
- All participants (100.00%) managed to complete the account creation phase in less than 10minutes, while 84.61% of the participants needed less than 10 minutes to create a new trip in the TRA application. More specifically, in terms of user registration time, half of the respondents took between 2-5 minutes to register in the system and the other half took between 6-10 minutes.

- In general, the users could easily insert information in the application. All of them found easy to select destination countries, to select transit countries, to provide vehicle information, and to provide travel document information without difficulty.
- Most of the participants (92.31%) could easily navigate using the user interface of TRA
- Most participants (84.62%) stated that they would use TRA application to accelerate the border crossing, 69.23% stated that they believe that TRA application would improve security, while 69.23% stated that they would use TRA regularly
- The average score was 4.11 out of 5.
- Understandability in terms of concept and functionality was rated the highest at 4.92 points.
- The lowest rating was given to the ability of the component to work with current systems 3.25 points.
- The stable Internet connection is pivotal for proper operation of the TRA.
- Question presenting the statement "The pre-registration process was efficient" received the highest rating averaging 4.54.
- The lowest rating was for question presenting the statement "I believe the pre-registration system will improve border security."

3.3.2 WI

WI facilitated the utilization of Facebook though data crawling, collection, and analysis in order to contribute to the calculation of risk indicators. It was applied only to one traveler (Ivan Waronka) whose risk indicators had been identified in previous check stages. The risk indicator calculated by WI in the Polish scenario was RI-103.

The assessment, lessons learned, and experience gained from WI are summarized below.

- This component was positively verified during pilot activities.
- This component was evaluated in terms of understandability, usefulness, information, integration, reliability, performance, maintainability, efficiency, and legal/ethical issues.
- The average score was 4.33 out of 5.
- The highest score was the sense of speed and responsiveness in the system and ease of configuration, maintenance, and future support - 4.67 points.
- The lowest rating was given to integration with other components, reliability of operation, and generation of results - 4.0 points.
- The stable Internet connection is pivotal for proper operation of the WI.

3.3.3 C2

C2 consisted of two parts: the back end and the front-end web application which interacted with the border guard officers at the BCP. The C2 back-end allowed front-end components and sensors to connect and send information through DMS receiving also accumulated information in DMS coming from other TRESSPASS sub-systems which were used for further analysis, presentation, and dissemination purposes. The C2 front end web app was used by the border guard officers providing infrastructure maps management, front-end interface with other data sources and sensors, alerts visualization, and the risk assessment results for travelers.

The assessment, lessons learned, and experience gained from C2 are summarized below.

- This component was positively verified during pilot activities.
- This component was evaluated in terms of understandability, usefulness, information, integration, reliability, performance, maintainability, efficiency, and legal/ethical issues.
- The average score was 3.9.
- The highest score was for lack of problems with legal/ethical issues related to data privacy, protection, and security - 4.25 points.
- The lowest rating was given for ease of configuration, maintenance and future support, and integration with other components - 3.58 points.
- The stable Internet connection is pivotal for proper operation of the C2.

3.3.4 SPA

Security Personnel Application was a mobile application, which was used by the first line officers within the BCPs to receive information from C2 or report to C2 abnormal behavior of travelers. SPA produced risk report which was forwarded to the C2 back-end component.

The assessment, lessons learned, and experience gained from SPA are summarized below.

- This component was positively verified during pilot activities.
- It worked well and helped the first line officers to see DRAS assessment results and create the risk reports.
- This component was evaluated in terms of understandability, learnability, usability, interoperability, integration, usefulness, information, installation, performance, and legal/ethical.
- The average score was 4.38.
- Understandability in terms of concept and functionality was rated the highest at 4.83 points.
- The lowest rating was given for the ability of the component to work with current systems 3.83 points.
- Bigger letters, buttons, and numbers would be desired. One should take into account the resolution of the screen.

3.3.5 Document Scanner/Module

The Document Module was composed of Regula passport scanners connected with laptops equipped with special software for passport scanning. The passport scanners and laptops have been provided by WAT while the software and configuration were provided by ICTS. The document scanner module communicated with the TRESSPASS system using Kafka messaging system.

During the pilot, several mockup passports, provided by ICTS, have been used.

The assessment, lessons learned, and experience gained from Document Module are summarized below.

- The stable Internet connection is pivotal for proper operation of the TCSS.
- No passport reading issues have been identified during pilot.
- Scanning time was at acceptable level <10s.
- Document Module should be further integrated into TRESSPASS C2 component.

3.3.6 TCSS

Thermal Counter Spoofing Sensor (TCSS) aimed to detect attempts to falsify facial appearance. Its purpose was to detect attempts to steal an identity in the form of changes in the appearance of the face. Identity theft detection technology was used to determine if the face presented in front of the camera had not been changed by wearing a face mask or presenting a picture with the image of another person.

The assessment, lessons learned, and experience gained from TCSS are summarized below.

- Significant effort was put to develop the system as a non-stop running service, which was positively verified during pilot activities The component provided very high accuracy of 99.3%, which means that almost all bona fide subjects were classified as bona fide and subjects wearing masks were classified as mala fide (attacks).
- The stable Internet connection is pivotal for proper operation of the TCSS.
- The main lesson learnt is that the subject has to be aware of where to look during checking.
- This information can be provided by any visual information (i.e., blinking light, displayed image or textual information).

3.3.7 IAS

IAS enabled countries inside and outside of the TRESSPASS network to send alerts regarding traveler risk information. The component was controlled through C2 inside the TRESSPASS network whereas TLC was used to control IAS from countries outside the TRESSPASS network.

The assessment, lessons learned, and experience gained from IAS are summarized below.

- This component was positively verified during pilot activities.
- This component was evaluated in terms of understandability, usefulness, information, integration, reliability, performance, maintainability, efficiency, and legal/ethical issues.
- The average score was 4.06.
- As many as five categories were rated highest - 4.5 points.
- The lowest rating was that the component functions and produces results reliably - 4.0 points.
- The stable Internet connection is pivotal for proper operation of the IAS.

3.3.8 TLC

TLC provided the ability to third countries (e.g., Belarus) to report information about threats and alerts through the IAS module. This component provided following functionalities: entering data about risky traveler and selection of BCPs which should be informed as well as receiving and presenting data about risky traveler from other BCPs.

The assessment, lessons learned, and experience gained from TLC are summarized below.

- This component was positively verified during pilot activities.
- This component was evaluated in terms of understandability, usefulness, information, integration, reliability, performance, maintainability, efficiency, and legal/ethical issues.
- Average score was 4.9. All items scored in the range 4-5.
- The stable Internet connection is pivotal for proper operation of the TLC.

3.3.9 MMCAT

The multimodal communication analysis tool (MMCAT) was an interview support system, providing technology for a security/authority officer ('Interviewer') conducting an interview. Its application was in interviews with suspected travelers (Aleksey Latuszka and Ivan Waronka) as part of the second line checks.

For the evaluation of MMCAT questionnaires with closed and open questions have been provided at the end of the Polish Pilot by the Border Guard officers interviewing volunteers with the support of MMCAT in the 2nd line checks.

The assessment, lessons learned, and experience gained from MMCAT are summarized below. Detailed Q&As and assessment scorings which were extracted from the answers provided in the questionnaires, are listed in Annex B.

- The scoring of the closed questions showed that the MMCAT tool is considered a helpful tool (4/5.0) and the training material was sufficient to use the tool (4.0/5.0). The graphical user interface is already rather good (3.5/5.0), but it could be improved by providing more information (4.5/5.0). The overall average score was 3.43/5.
- This component was positively verified during pilot activities.
- It worked well and helped the second line officers to conduct the interview and to find mala fide travelers.
- The strengths of the MMCAT system were indicated as efficiency of operation and facial analysis recognizing face signals not observed by the interviewer.
- The weaknesses were lack of automation and problem in passenger reaction assessment.
- Intuitiveness of operation was rated as medium.
- Respondents were satisfied with the GUI interface and training materials.
- They expected more detailed information from the MMCAT system.
- The stable Internet connection is pivotal for proper operation of the MMCAT.
- Suggested improvements of the GUI were discussed with Vicar Vision and TNO. This valuable user feedback allowed the implementation of these improvements in a MMCAT update for the Greek pilot.
- GUI was rather raw. Some further elaboration would be needed.
- Entering login required some time. Automatic login would be desired.
- Automatic selection of camera would be desired.
- Only central part of camera with the person's face should be visible.
- Minimizing of any additional clicking would be desired.
- MMCAT result form should be always visible (e.g., on the right below the buttons).
- Name of the person instead of number, name should be visible.
- Three buttons (green, orange, red) should be used to click.
- 3 values of confidence: 0, 0.5, 1 (or 4-5, no more).
- Bigger letters, buttons, and numbers would be desired. One should take into account the resolution of the screen.
- MMCAT and C2 should operate together on a single 27" screen.

3.3.10 Back End Components (LSI, DFA, DMS, DRAS)

➢ **LSI**

The LSI component aimed at creating adaptors for accumulating and aggregating information from heterogeneous legacy information systems (databases) such as VIS, SISII, and EES. The interoperability between TRESSPASS and these legacy systems aimed at providing useful information and alerts from passenger profile checks, regarding the construction of risk indicators. LSI implemented a list of simulated legacy information systems that leveraged the extraction of risk information. These simulated Legacy Information Systems were developed as Web Services.

The assessment, lessons learned, and experience gained from LSI are summarized below.

- This component was positively verified during pilot activities.
- This component was evaluated in terms of understandability, usefulness, information, integration, reliability, performance, maintainability, efficiency, and legal/ethical issues.
- The average score was 4.23 points out of 5.
- Understandability in terms of concept and functionality was rated the highest - 4.92 points.
- Ease of configuration, maintenance and future support of the component was rated the lowest - 3.25 points.
- The stable Internet connection is pivotal for proper operation of the LSI.

➢ **DFA**

DFA operated as a back-end process which continuously collected input via the DMS (Distributed Messaging System) provided by each data source and periodically inferred potential risk indicators occurrences along with potential confidence value, timestamp and data sources involved.

The assessment, lessons learned, and experience gained from DFA are summarized below.

- This component was positively verified during pilot activities.
- This component was evaluated in terms of understandability, usefulness, information, integration, reliability, performance, maintainability, efficiency, and legal/ethical issues.
- The average score was 4.28.
- The highest rating was given for the feeling of speed and responsiveness in the system and ease of configuration, maintenance, and future support - 5.0 points.
- Legal/ethical issues related to data privacy, protection, and security were rated lowest with 3.0 points.
- The stable Internet connection is pivotal for proper operation of the DFA.

➢ **DMS**

DMS was used as a means of intercommunication between different components in TRESSPASS. Technical partners created channels that they used to public information and other partners could subscribe to the channels that they were interested in and received the information published.

The assessment, lessons learned, and experience gained from DMS are summarized below.

- This component was positively verified during pilot activities.

- This component was evaluated in terms of understandability, usefulness, information, integration, reliability, performance, maintainability, efficiency, and legal/ethical issues.
- The average score was 4.29 points.
- Understandability in terms of concept and functionality was rated the highest - 4.57 points.
- The lowest score was for ease of configuration, maintenance, and future support - 4.0 points.
- The stable Internet connection is pivotal for proper operation of the DMS.

➢ **DRAS**

DRAS provided real-time risk assessment for each traveler based on continuously receiving traveler risk indicator data from Data Fusion Analytics (DFA) module, calculating traveler risk score per threat and overall risk score, using weight-based formulas, and providing traveler risk profile to the border guards through the Command-and-Control front-end user interface.

The assessment, lessons learned, and experience gained from DRAS are summarized below.

- This component was positively verified during pilot activities.
- The average score was 4.67.
- The remaining scores ranged between 4 and 5 points.
- The stable Internet connection is pivotal for proper operation of the DRAS.

3.4 Data and Risk/Bona Fide Indicators

The assessment, lessons learned, and experience gained from the use of mock-up data and the Risk and Bona fide indicators which were prepared for the Polish pilot are summarized below.

- The mockup data and Risk/Bona fide indicators were positively verified during pilot activities.
- Twelve people took part in the Pilot and held rotational roles of travelers and Border Guard officers. Persons were using “names” represented by letters A - L and each day they played different roles – travelers (numbers 1 – 12) or the officers. This way of organizing the pilot was considered optimal as people could become familiar with both roles and better understand the TRESPASS concept.
- The informed consent form (both in English and Polish) and other relevant information (Land Pilot. Additional safety measures - both in English and Polish) were prepared by PBG. PBG sent the informed consent form to the persons before the tests. Before the persons were asked to do any activity, they first had to agree upon the informed consent form and signed this form. The content of the informed consent form was also explained orally to the persons. The PBG TRESSPASS team members collected the signed informed consent forms. This way of organizing the consent form signing was considered optimal as people could become familiar with their roles and requirements.
- On the other side of the card mimicking the first page of a passport, we prepared the description (in Polish) of the role the person was to play including the most important information, behavior during the tests (nervousness, calm), and answers to questions. This solution was considered by the travelers as very useful and helpful.

- To facilitate the process of registration in TRA, for each traveler a two-page sheet with all necessary data was prepared. Description of fields was provided in English and Polish. If the registration process was successful, the traveler should have seen the last page in the TRA app the same as the page. This solution was considered by the travelers as very useful and helpful.
- In order to facilitate the process of adding alerts to the TLC by the third country LEAs, for both travelers a short table was prepared and printed. Every day, these data were entered to TLC by the selected BG officers. This solution was considered by the BG officers as very useful and helpful.

3.5 Risk Assessment (Concept & Methods)

Risk-based concept which is the key element of the TRESSPASS system has been proposed to operate at various types of border crossing points and in various scenarios. The land border concept involved two scenarios, reflecting two main challenges identified at this specific BCP, where the TRESSPASS Land Border Pilot took place (Terespol, Poland). However, the presented challenges, scenarios and approach may be easily extrapolated to other land BCPs in EU.

One of the main challenges for building a risk-based concept for land border crossing point is to define a proper set of indicators that will be supplemented by continuous information provided by various systems at each travel stage (pre-travel, approaching BCP and at BCP).

Selection of indicators and availability of data, especially knowledge of the pre-travel information is crucial for reducing bottlenecks or any side-effects that may be accidentally introduced to the system.

Multiplicity of components and systems used for Risk Assessment at land border crossing point resulted in very interesting observations and conclusions. The operation of proposed approach was strongly supported by technical partners, especially those operating locally. The technical challenges have been overcome and the overall concept was tested and evaluated.

The final list of Risk Indicators alongside their business logic which defines the conditions of the indicator, the indicators' weights and the creation of subsequent risk/threat profiles greatly affects the effectiveness and accuracy of the system and its ability to currently identify legitimate threats.

3.6 Research Questions and Key Performance Indicators (KPIs)

	#	Research questions
Effectiveness	1	What is the effectiveness of the TRESSPASS RBBM concept (and corresponding available information) in comparison to the current rule-based approach (and corresponding available information)? **A: TRESSPASS RBBM concept, if fully implemented in accordance with the assumptions and if more risk indicators and bona fide indicators were taken into account, it would most likely contribute to increase the effectiveness of the border crossing point operation. In addition, thanks to TRA, a priori information on passengers would contribute to better human resource planning and operational efforts. Moreover, use of TCSS would provide the detection of face masks, which is not analyzed at all now.**

	2	What is the probability of a false positive (i.e. people that are bona fide are send to the manual desk) and a false negative (i.e. people that are mala fide are not send to the manual desk) **A: The probability of a false positive and a false negative could be lower, mostly thanks to use of information from many sources, including the Internet (WI), MMCAT system, and broad spectrum of indicators.**
	3	To what extent can the information of the TRESSPASS system potentially help to better execute the process of mitigating the threats handled by your pilot (irregular migration, smuggling etc)? **A: Analyzing current situation at land border, TRESPASS information can certainly detect a mask on a person's face, which is now not taken into account, and help to select cars/trucks more efficiently for thorough checking.**
Flowrate	4	What is the effect of the TRESSPASS RBBM concept on the flowrate of different categories of travelers in comparison to the current rule-based approach? **A: In the analyzed cases, the border check time should be shortened, as the traveler comes to the counter with a preliminary assessment. This should increase the flow rate of different categories of travelers in comparison to the current rule-based approach. An additional factor is the optimization of human resources and, finally, the reduction of border control costs.**
Efficiency	5	What are the resources needed to perform the RBBM-concept to achieve a desired level of flow rate and/or effectiveness? **A: Taking into account the land border scenarios, to achieve a higher level of flow rate and/or effectiveness using the RBBM-concept, one requires at least installation of the TRESSPASS components in majority of BCPs and some points abroad (IAS) as well as training a lot of border guards and customs officers.**
Acceptance	6	To what extent perceive the participating border guards the TRESSPASS RBBM concept and system as usable? **A: In general, the participating border guards positively validated the issue of application of RBBM concept to the current land border. The RBBM could be integrated with existing border control system or the newer modernized one and will be another useful source of information and travelers' check-in would certainly be more complete.**
	7	To what extent perceive the participating border guards the TRESSPASS RBBM concept and system as trustworthy? **A: The participating border guards perceive the TRESSPASS RBBM concept and system as trustworthy to a great/medium extent.**
	8	To what extent perceive the participating RNM-border guards the TRESSPASS RBBM concept and system as efficient? **A: The participating border guards perceive the TRESSPASS RBBM concept and system as trustworthy to a great/medium extent.**

	9	What are the implications of a RBBM concept for the personnel at strategic, tactical and operational level? **A: The implications can be considered on several levels: during border checks, an officer has more information about the traveler and based on the system prompts, can perform border checks faster and easier. Officers responsible for planning and organizing work at border crossings can more easily plan use of human resources and operational efforts and supervise the work of subordinate officers. Finally, strategic level officers have greater system awareness.**
Balancing	10	What is the tradeoff between the effectiveness, flowrate and efficiency of the TRESSPASS RBBM concept? **A: This is not an easy question at this stage of the work. It seems that the most important is the detection rate (effectiveness), which may take place at the expense of the flow rate and efficiency.**

3.7 Training

The land pilot started with a training session – a presentation given to all participants on May 25, 2021. The presentation covered the following issues: project overview, TRESSPASS concept, a risk-based approach, Risk-Based Border Management (RBBM), TRESSPASS Risk Assessment Methodology (TRAM), TRESSPASS System overview, main system components: C2, SPA, TRA, and further components: MMCAT, TCSS, TLC.

The main learning objectives of this presentation were to gain an overview of the TRESSPASS project and its main contents, to learn what the TRESSPASS concept is, and what the idea behind risk-based border management (RBBM) is, to understand the basics of the TRESSPASS system, to familiarize yourself with the components C2, SPA and TRA, the presentation was initially prepared by CASRA in English, and then adapted and translated into Polish by WAT.

At the end of the training session, its participants filled out an evaluation questionnaire. The purpose of the questionnaire was to evaluate the previously conducted TRESSPASS Pilot training. The questionnaire consisted of 18 closed questions. The questions required a response on a scale from 1 to 5, where 1 was the lowest rating (Do not agree at all) and 5 was the highest rating (Fully agree). Other ratings were: 2 - Do not agree, 3 - Undecided, 4 – Agree. Answering each question was optional. 14 respondents completed the questionnaire.

General impression of training was evaluated: did you like it, what was the atmosphere like, was the training useful. Next, content of training materials was evaluated: their comprehensibility, whether they introduced new contents extending knowledge, whether sufficient time was devoted to the discussed topics and whether the discussed contents were supported by examples. At the end respondents were asked to assess if they were able to use acquired knowledge in practice, if they were aware of learning objectives and if they could practice what they have learned and contribute to training.

The average score was 4.1. The highest marks were given for the atmosphere at the training - 4.7 points, general impression of the training and possibility to practice the newly acquired knowledge - 4.4 points. The lowest score was given for awareness of learning objectives - 3.5 points, new topics will be remembered and applied to the role in the TRESSPASS pilot - 3.8 points.

The atmosphere at training courses was crucial to encourage learning new things and it was very important to explain the purpose of learning thoroughly at the beginning.

At the end of the training, participants completed a training evaluation questionnaire, consisting of 19 closed questions, which had to be rated on a Likert scale (Do not agree at all to Fully agree). The questionnaire is based on the Training Evaluation Inventory (TEI), an approach to training evaluation that aims at ensuring that investments have the highest possible degree of efficiency (Ritzmann, Hagemann, & Kluge, 2014). It stands out from other evaluations in that it is a valid and reliable training evaluation, which is theoretically and empirically founded. The TEI assesses training outcome dimensions (subjective enjoyment, perceived usefulness, perceived difficulty, subjective knowledge gain, and attitude towards training) as well as training design dimensions (problem-based learning, activation, demonstration, application, and integration) on ten scales (1=does not apply at all, 5=fully applies).

In total 14 questionnaires were returned complete. Overall, the participants liked the training and considered the training atmosphere pleasant (training outcome dimension "subjective enjoyment"). Most of them also thought the training was useful for their role in the pilot (2/14 undecided) (dimension "perceived usefulness"). All participants agreed or fully agreed that the training contents were comprehensible (Figure 8: Polish Pilot Training Evaluation Questionnaire Answers 1/3

) (dimension "perceived difficulty").

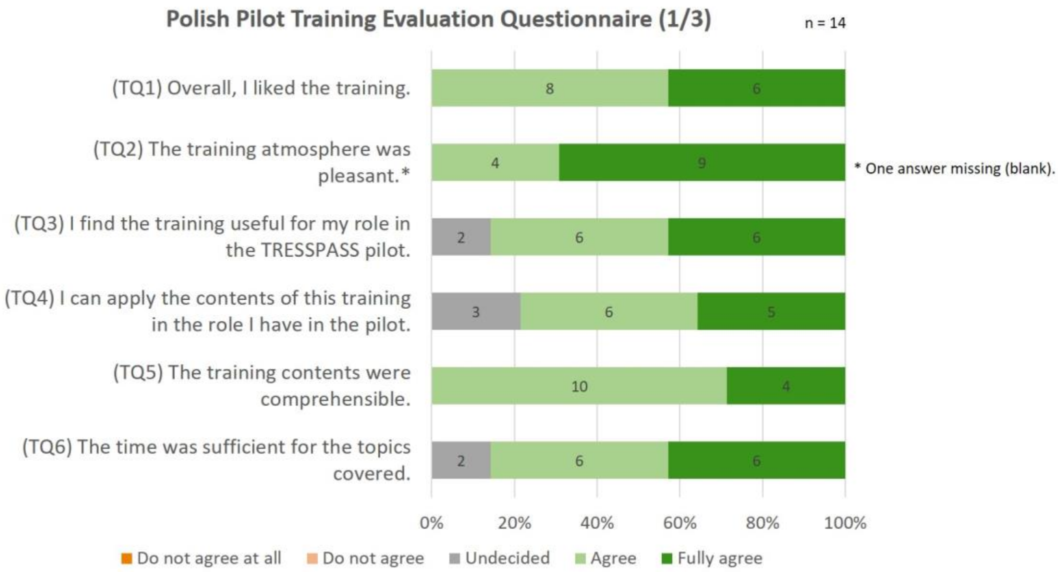


FIGURE 8: POLISH PILOT TRAINING EVALUATION QUESTIONNAIRE ANSWERS 1/3

Regarding the subjective knowledge gain dimension, answers were somewhat more diverging (Figure 9: Polish Pilot Training Evaluation Questionnaire Answers 2/3

). Two participants fully agreed that they have the impression that their knowledge has expanded on a long-term basis, a majority agreed on that, and two participants disagreed. For items TQ8, TQ9 and TQ10, the same participant has given the answer "do not agree". However, the majority of participants agreed that they will be able to remember the new topics, and they will apply them in their role in the TRESSPASS pilot. It appears that two participants did not have the impression that they were invited to bring their previous knowledge and experience into the training context, and two were undecided on this topic.

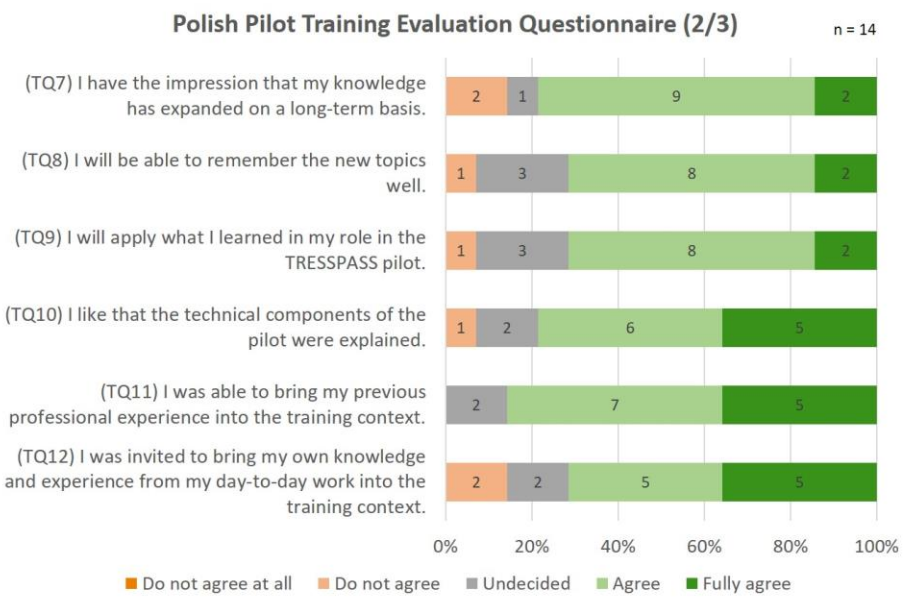


FIGURE 9: POLISH PILOT TRAINING EVALUATION QUESTIONNAIRE ANSWERS 2/3

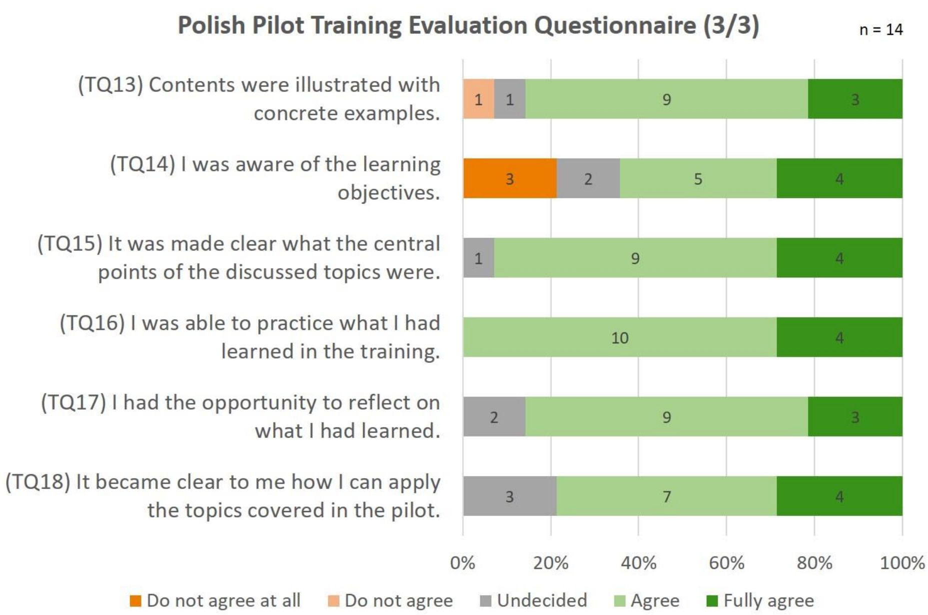


FIGURE 10: POLISH PILOT TRAINING EVALUATION QUESTIONNAIRE ANSWERS 3/3

It appears that three out of 14 participants were not aware of the learning objectives, two were undecided on it, and the rest was aware of the learning objectives. All participants agreed that they were able to practice what they had learned in the training. However, some participants were undecided on whether or not it was clear to them how they can apply the topics covered in the TRESSPASS pilot (Figure 10: Polish Pilot Training Evaluation Questionnaire Answers 3/3

).

In order to improve the training and training contents, the following points could be considered:

- There seems to be some uncertainty regarding the application of the contents of the training to the participants' role in the pilot (see TQ4 and TQ18). Thus, the connection of the training contents to practice could be emphasized more.
- Previous work experience from participants could be taken (more) into consideration, in order to establish a better connection between previous knowledge and new training contents.
- As the learning objectives were not clear to all participants, a more straight forward and concise introduction of these should be considered.

3.8 Simulations

3.8.1 Results from FHG Simulator in the Polish Pilot

This section describes the different scenarios used in the FHG land pilot simulations, the process flows in each scenario, the performance areas and the corresponding key performance indicators observed under multiple simulation runs executed using different configurations in the Fraunhofer's simulator.

Scenarios: The scenarios used in the land simulations explored the risk-based inbound Smuggling (SMUG) and the Cross Border Crime (CBC) threat scenarios evaluating and comparing them versus the corresponding rule-based scenarios having 4 scenarios in total as follows:

1. Risk-Based Inbound Smuggling scenario
2. Rule-Based Inbound Smuggling scenario
3. Risk-Based Inbound Cross Border Crime scenario
4. Rule-Based Inbound Cross Border Crime scenario

For each of the above scenarios different configurations were used expanding the 4 scenarios into multiple scenarios depending on the input parameters used in the simulation tool. A total of 25 configurations were run based on the inputs provided by the end-users.

Process Flows: For each set of the above Risk/Rule-based cases one set of flow process diagrams was drawn as depicted in the two figures below. The flow process diagrams consist of stages, namely the pre-travel, at-bcp and post-bcp travel phases.
Each stage consists of component groups which represent the checking and screening points in the process flow of the sea pilot.
Each component group in turn includes, the TRESSPASS components and the human resources involved at each component group (Border Guards etc.) presented as Border components.
The flow diagrams also include the different paths a traveler can take depending on the travelers' risk level detected at each component group/checkpoint.

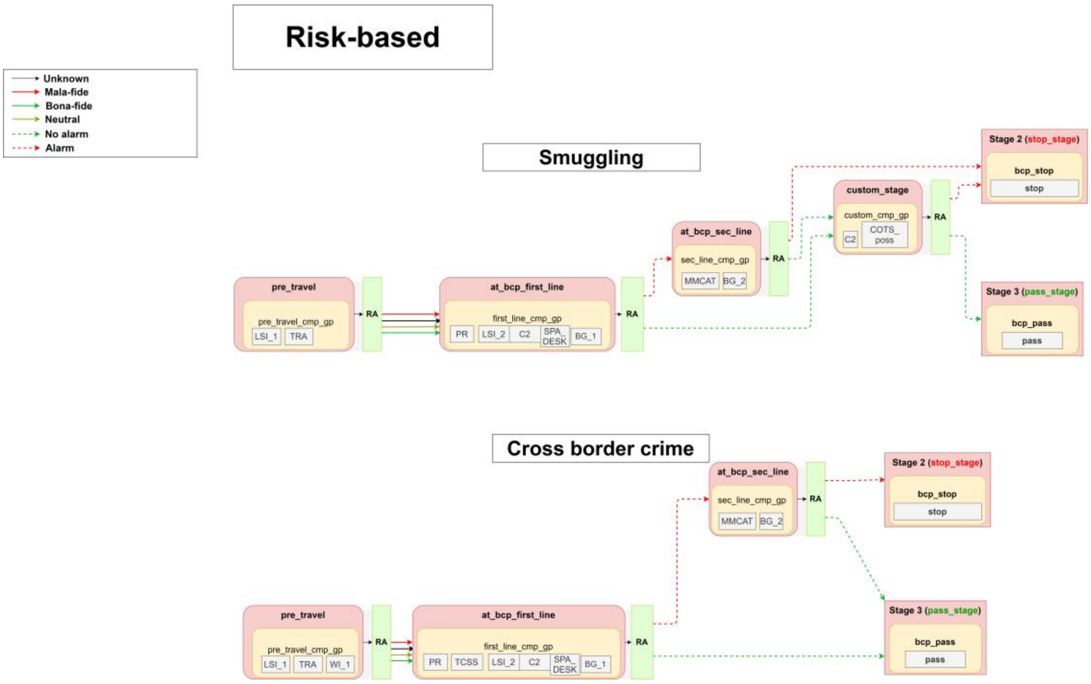


FIGURE 11: LAND PILOT FHG SIMULATION: RISK BASED PROCESS FLOW DIAGRAMS

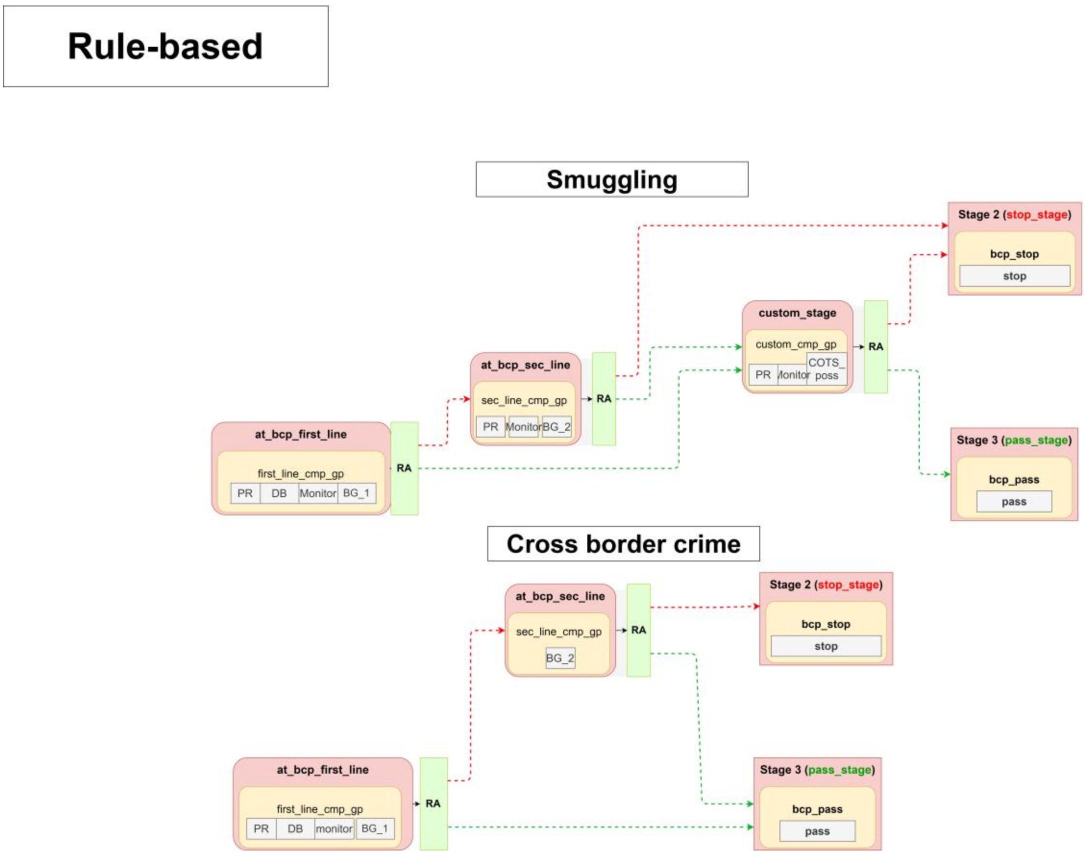


FIGURE 12: LAND PILOT FHG SIMULATION: RULE BASED PROCESS FLOW DIAGRAMS

Performance Areas/Indicators – Observations: For the land pilot, a total of 52 configurations were evaluated as provided by the end-users.

In the simulations the following performance areas were observed associated with the corresponding key performance indicators i.e., flowrate and effectiveness:

1. Flow related (flowrate)
2. Access and egress control (effectiveness)

In the following, the observations from the simulations are listed:

1. When the arrival rate of the travelers was increased (200->550->825), both rule-based and risk-based BCP configurations register a higher waiting time for the travelers.

	Rule-based		Risk-based	
Arrival rate (traveler/hour)	Throughput (traveler/min)	Waiting time (Min)	Throughput (traveler/min)	Waiting time (Min)
825	13.3	16.18187523	13.58743247	0.463045179
550	8.88109106	12.58697957	9.137491301	0.017852565
200	3.359937028	0.03887891	3.369152293	0.357711949

2. For the risk-based configurations, the average waiting time of the travelers is considerably lower thereby increasing the overall throughput of the BCP.
3. No major variations were observed in Detection Rate (DR) and False Alarm Rate (FAR). The possible ways to improve it are to re-adjust the TRAM configurations, accurate traveler profiles, and adapt the BCP layout as per the risk-based design.

Rule-based CBC		Rule-based SMUG		Risk-based CBC		Risk-based SMUG	
DR	FAR	DR	FAR	DR	FAR	DR	FAR
0.95951417	8.08E-05	1	0.000605791	0.96484375	0.00010103	1	0.000667
0.975510204	0.000161584	1	0.000565679	0.96473029	4.04E-05	1	0.000384
0.967512187	0.000121192	1	0.000585735	0.96478702	7.0717E-05	1	0.000525

For a detailed analysis of the results, one should also need to invest efforts in observing the queuing behavior within the BCP and how does it change when changing the service points. Changing the number of service points will also increase the cost of used resources. All these results are available within the simulator output files but for a detailed analysis of the results, the domain knowledge of the border operation is needed. Although the simulator was developed with close collaboration with the end-users, the analysis still highly relies on a thorough understanding of border operations and thus could be improved by the end-users with the support of their technical partners.

3.8.2 End Users' Lessons Learnt from simulations in the Polish Pilot

Border control in a BCP is usually a compiled system consisting of many interconnected parts and is not readily available for actual testing due to ethical, technical, and operational issues. Therefore, a simulation tool that mimics the operation of the BCP and generates the required data becomes a better solution for assessing the performance of the BCP. The FHG Simulator seems to be a very useful and effective tool for the design of risk based BCPs (but also the rule based), their evaluation in terms of traveller flow, effectiveness, and ethical performance without setting up any physical experiment.

We found it very productive to start using the simulator from the hands-on training (30th Sept 2021), which was very efficient and provided a good foundation to understand the simulation basics and starts own work. The subsequent direct contact with the system designers later proved to be very helpful and ensured effective learning and testing.

The Land Pilot team (PBG and WAT) conducted several tests of the FHG simulator to find out how it works, how to build a simulation and logical flow for risk- and rule-based border control, what inputs are required, and how to analyse the results.

Regarding the simulation and logical flow, we found this issue interesting and worth further studies. The idea of grouping the components, that are placed sequentially and can have several service lanes, is very similar to the real operation of the systems. Also, the stage, which represents one of the travel processes of agents within the whole journey and is mainly relevant to define the overall flow direction for different categories of travellers within the system, corresponds well to reality.

As per input parameters, some of them were quite easy to enter the simulation, e.g., mean service time and arrival time, which are precisely defined for all BCPs in Poland. The border guard is described by performance parameters indicating the quality of his/her decisions. These parameters are detection rates and false alarm rates. However, PBG doesn't know exact numbers on detection rate and false alarm rate of BG at first/second line checks (rule-based). Even the bigger problem was with determination of the main parameters of BG in risk-based approach: TDR1, TDR2, FAR1, and FAR2[15]. It is clear, that these parameters are essential to obtain the accurate results and some further research should be conducted to determine them in the methodological and reliable way.

The next group of parameters that we dealt with embraced traveller profiles. Also, it was found quite hard to access distributions of all properties for different border crime travellers as well as for non-malicious persons[16]. This area also should be further studied to provide the reasonable input data for the simulation.

The FHG Simulator generates several outputs, in the form of plots (pdf. files) and numerical values (csv. files), which can be used to analyse the BCP performance. A stage loading plot, which shows the number of malicious and non-malicious travellers evaluated at each stage, seems to be very helpful in understanding how different categories of travellers are moving in the BCP. The arrival and departure plots, which presents the time evolution of arrival and departure of travellers from the BCP, are important to notice if all travellers exit the BCP in the right way.

[15] See D7.8 Fast Simulation and Analysis Algorithms for Risk Based Border Control Systems, section 4.1.3.

[16] See D7.8 Fast Simulation and Analysis Algorithms for Risk Based Border Control Systems, section 4.4.

A csv file with the performance indicators incudes all the crucial information needed for the assessing the performance of the BCP, which corresponds to KPIs and PIs[17], like overall detection probability, false alarm rate, average service time, and throughput.

After running the set of simulation, the Land Pilot team had a better understanding of the RBBM concept, its benefits, and requirements. The simulator in its current form requires the participation of the FHG team because all input data should be provided to FHG in excel files, and 19 output files are sent to end users. We find this procedure tedious which cannot provide any real-time feeling. However, one can easily design a kind of dashboard that allows you to design an experiment, set parameters, and visualize the results efficiently.

The Land Pilot team found out that the simulator generates realistic and relevant results given the input data is of good quality. We have learned that it cannot only be used to mimic different BCP scenarios, visualize performance and conduct performance evaluation, but also to support BCP design and modelling at different levels, as well as to assist in training and optimizing BCP performance. However, it is believed that there is room for further improvement, e.g., by increasing number of traveller profiles, i.e., features of different categories of travellers.

3.9 Overall Pilot Lessons, Limitations and Obstacles

The questionnaire on ‘Qualitative evaluation at project/conceptual level’ was prepared in English by the relevant partners and then translated into Polish. The participants filled them in in Polish and their answers were translated into English.

The questionnaire contained 10 open questions on general topics at the project concept level. The questionnaire was completed at the end of the Pilot on May 27, 2021, by 15 PBG officers. Thirteen officers described their role in the organization as tactical, one as operational, and one as strategic. Providing an answer to each question was optional.

Analyzing the answers given to the question what the expectations from the Pilot were and whether they were met, it can be said that the opinions were divided into two groups. The first group did not have any expectations because e.g.: “the assumptions of the Pilot were not known to me”. The other group expected to learn about new systems facilitating border clearance and their expectations were met. Both groups agreed that the goals of the Trespass project were clearly stated and explained at the beginning of the pilot study.

The main strengths and weaknesses of the TRESSPASS project were identified as:

- Strengths:
  - Using technology for additional third country nationals’ analysis and profiling to help the officer at the border.
  - Gathering material on potential travelers.
- Weaknesses:
  - Extension of border checks.
  - Use of analysis against third country nationals crossing the border several times a day.
  - Lack of integrity with border control systems.
  - Too many activities for a BG officer to perform.

[17] See D7.8 Fast Simulation and Analysis Algorithms for Risk Based Border Control Systems, section 5.9.

Respondents noted the potential positive values of the TRESPASS project for current control and operational procedures. They indicated that these include the potential for counterterrorism, detection of behaviors that indicate travelers are lying, increased data, speed of information flow, and increased importance of human behavioral factors.

Regarding the application of the RBBM concept of the TRESSPASS project to current Border Management systems, the respondents were mostly in agreement. They responded that the concept can be applied but that it must be integrated into current border clearance systems, otherwise it cannot be used.

Summarizing the overall impression of the project, **respondents expressed approval of the project concept**, see its potential, and the need to continually improve the check-in process.

4 GREEK PILOT: LESSONS LEARNT

This chapter breaks down in sections each aspect of the Greek Pilot assessment, the lessons learned, the experience gained, and the obstacles observed by the pilot's stakeholders upon the completion of the pilot's' execution, the collection and processing of the external and internal feedback via the evaluation tools (questionnaires etc.) used and the relevant pilot deliverable report (D8.4).

4.1 Operations

4.1.1 Recap of CONOPS As-Is

As mentioned in the previous sections (2.1.1 and 3.1.1) for the respective air and land BCP pilots, the TRESSPASS CONOPS is concerned with the operational environment of the BCP as it moves to integrate the new TRESSPASS system. Supported by an activity theory framework, the CONOPS detailed the current BCP operational picture at Piraeus Sea port and described the system updates that would potentially take place with the introduction of the TRESSPASS architecture and the operational findings that occurred during the pilot testing phase, from a human factors perspective.

The sea pilot was carried out at the Piraeus Sea Port (Greece). This port includes two cruise terminals (Terminal A and Terminal B) and is one of the largest ports in the world regarding passenger traffic, with a throughput of approximately 20 million passengers per year. Additionally, Piraeus Port was recorded as having a growth of 12.7% between 2017-2018, with regards to cargo and freight, placing it in the top 20 cargo ports in Europe. The cruise vessels usually stay in the port for between 24 and 48 hours. Currently, up to 1.000 passengers are processed per hour: as a third terminal is under construction, the number is expected to increase up to 5.000 travelers per hour.

Multiple authorities are involved in border control activities at Piraeus, including Customs, the Hellenic Police (including immigration), Port Police, PPA security staff, the Hellenic Coast Guard, and staff from the cruise operators (admin and security staff). Representatives from each of these end-user authorities were approached and asked to share any relevant information about the requirements, challenges, and expectations of their BCP and of TRESSPASS system, respectively. This was done to ensure that the TRESSPASS system would be useful for their daily operations.[18]

The current BCP operating system at Piraeus port consist of rule-based checks; however, Customs authority are allowed to perform additional controls whenever they are deemed necessary. These checks are time-consuming for both staff and passengers and this was mentioned as one of the core challengers of the BCP. The introduction of the TRESSPASS risk-based border management system is hoped to unify and enhance the security process across the port and between Border Protection and Customs Authorities. The technological

[18] TRESSPASS. *D6.2 Evolving CONOPS framework* (1st version)

architecture aims to support the identification and the generation of risk alerts, relying on historic data analysis, behavior analysis, improved and unified cross-border communication and the integration and sharing of information.[19]

The process diagram below illustrates the departures process for Piraeus Port to non-Schengen countries for the current as-is system, as reported in the first version of the evolving CONOPS guide, D6.2.

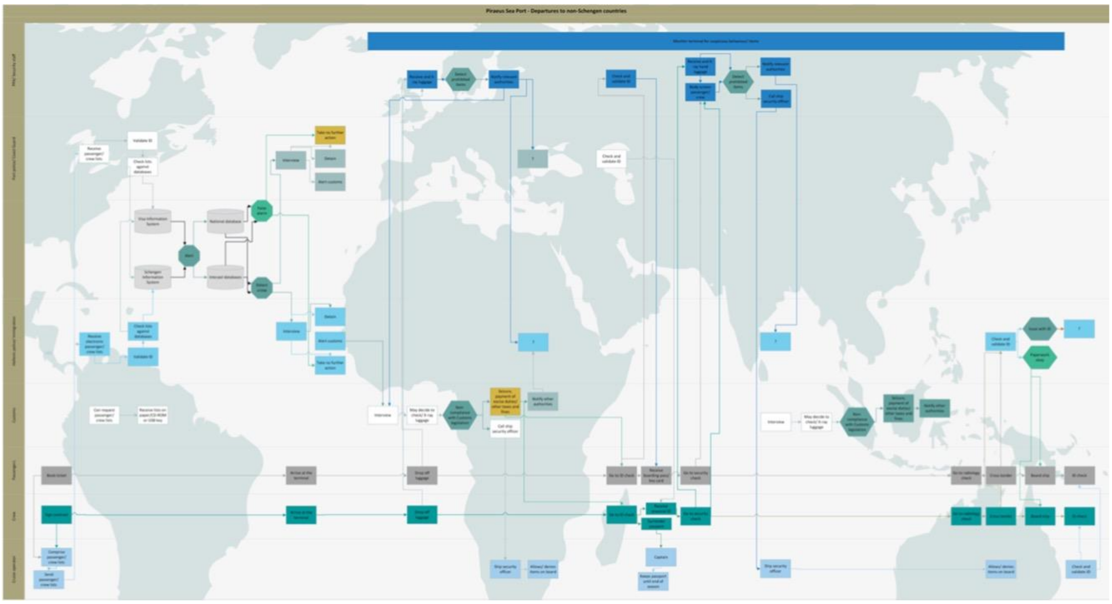


FIGURE 13: DEPARTURES FROM PIRAEUS TO NON-SCHENGEN COUNTRIES

Throughout the course of interactions with the sea BCP end-users, the following core challenges in daily border control activities were listed:

- Heavy workload and fatigue for staff.
- bottlenecks and delays in control procedures.
- lack of practical collaboration amongst multiple authorities.
- lack of a system facilitating the detection of suspicious passengers.
- inconsistency in the availability of information, as customs personnel does not have access to the same databases (i.e., VIS and SIS) as the police and port police.
- challenges in hosting the embarking and disembarking services of many cruise ships simultaneously.[20]

Based on these challenges, end-users indicated the following anticipated outcomes:

- Better use of existing infrastructures and facilities, allowing to speed up border and custom control processes with limited resources and without delay to passengers; the implementation of a none stop point control for security and border control (SBC) would be desirable.
- Access to shared data set from reliable sources.

[19] TRESSPASS D6.5 *Evolving CONOPS framework* (2nd Version)

[20] TRESSPASS. *D6.5*; TRESSPASS website.

- Availability of an IT system to facilitate information sharing and greater interoperability between the system and the databases.
- Enhanced cooperation amongst the different agencies in charge of border control management.
- Compliancy of TRESSPASS Technologies to GDPR requirements.
- Overall reduction of operational costs, while maintaining/increasing security, reducing time for checks and workload for border control operators.

Taking the core challenges of their BCP into consideration and the anticipated improvements and/or changes that the introduction of the TRESSPASS system will bring to their operational environment, the team at Piraeus port BCP aimed to cover all three threat categories, i.e., National Security, Cross Border Crime, and Irregular Migration in their sea pilot test as well as test the majority of the TRESSPASS system components.

The overall goals of piloting the TRESSPASS system at Piraeus port were:

- to reduce long waiting times, while subsequently increasing the efficiency of the BCP by improving the security checks,
- to introduce Maritime-PNR, integrate it into existing systems and share these data with other system dependencies (e.g., Customs)
- to enhance multi-agency cooperation and information exchange and to introduce innovative concepts and integrate them with current security processes, in compliance with current policies and protocols.

For a more detailed description of the sea pilot teams' objectives, please see D8.1: Pilot Planning and Training, section 4.5.

4.1.2 CONOPS consideration post- pilot

The main operational findings, from a technological, procedural and organization perspective, that emerged post-piloting phase for the Piraeus port BCP were as follows.

I. Technological: A type of risk pre-assessment is already carried out in the current system at the sea BCP, when certain specified risk indicators are identified, searches are done to EU, International and National Databases through a single interface prior to the passenger's arrival. These databases are considered Legacy Systems within TRESSPASS. Border control operators mentioned that it would be of greater benefit to them if the list of risk indicators within TRESSPASS was shown with a detailed description on the central database (C2). In that way they would have more information available to them to know who to direct to further checks and who to not. Having the risk information located in one centralized component was already considered beneficial and if the risk information was more detailed it would decrease the difficulty of the risk assessment task for the border guards, subsequently reducing their workload and pressure, whilst also providing more secure risk information.

II. Procedural: From an operational perspective, end-users found that the TRESSPASS system would provide them with more security information, therefore increasing the effectiveness of the security checking procedure, however, the procedurally, the duration and time consumption of the security check was considered to be the time with the introduction of the TRESSPASS system.

III. Organizational: The current system in operation at Piraeus BCP is complex with varied activities carried out by several BCP actors. Greek pilot team end-users emphasized the fact that current procedures are in compliance with the Schengen. Therefore, TRESSPASS system was seen as providing additional useful information for the security checking procedure but would not replace the current system procedures from an organizational perspective due to this Schengen governing protocol.

The introduction of the innovative TRESSPASS technologies is deemed to positively impact the security check process at Piraeus port BCP by providing additional risk information about passengers at a prior period in time. The passenger flow and queue time crossing the BCP was predicted by end users to remain the same but customer/passenger satisfaction and waiting times were not considered to be the highest priority for the BCP, rather the effectiveness of the security checking at the port BCP.[21]

4.1.3 CONOPS Conclusions

Piraeus Port Sea BCP currently operates a complex system with multiple authorities engaging in many activities. The sea end-users team wished to address challenges of heavy workload, bottlenecks at the port, introduce maritime PNR and increase multiagency communication and cooperation. The sea pilot test aimed to address several threats to the port BCP. Operational findings showed that the introduction of the TRESSPASS system was considered to assist the workload of the BCP operators by providing additional risk information and hence would improve the security effectiveness of the BCP. The passenger flow across the BCP may not be improved, however this was considered to be the loss for a gain of a more secure BCP.

4.1.4 End Users' views

The TRESSPASS system defines a pre-set of default procedures and rules, adaptable each time to the necessary risk indicators regarding the threat category of interest.

Border guards made use of specific components and in total the majority of the system's components.

Knowledge through information gathering at a pre-travel stage enhances the quality of border controls and guarantees at the same time the effectiveness of passing through, in proportionality to the measures in effect.

From the operations point of view, TRESSPASS system introduces many new and conceptually innovative and future technologies. While all these new technologies need some more development, the introduction of a system like TRESSPASS to the classic border control, is feasible.

However, the introduction of system like TRESSPASS should be done at the same time from all EU countries. Large scale systems as such could not work if the implementation is done only from few countries. In addition, the introduction of such system would not make much difference in the during border checks at BCPs, but the most important gain would be the availability of information and more secure screenings and checks prior to the travelers' arrival at a BCP so that Border Authorities focus on mala fide travelers and enhance the flow of legitimate travelers.

Furthermore, efforts should be invested -first and foremost- to the legal and ethical compliance of the system, as well as, to the development of the components that may act without a human intervention to ensure ethical and legal compliance.

[21] *Ibidem.*

From the terminal operator point of view, the introduction of new processes during embarkation and disembarkation of cruise travelers creates some changes in the operation that need further exploitation for the consequences in the operational efficiency. On the other hand, the creation of a (more) secure and modern terminal can be promoted and offset possible delays. For example, the enrollment phase is new for the terminal operations. The way that it has been introduced during the pilot operation, most probably cannot be applied in the everyday practices. A study is needed to assess the most suitable location of the kiosks and how many kiosks are needed to serve the flows. Furthermore, collaboration is needed with the cruise operators who are responsible to tag the luggage to be delivered on the ship i.e. where the TRESSPASS tags will be installed. Furthermore, the presence of border guards with the mobile terminals, requires attention on the way it will be performed (randomly based on suspicious behaviors or to all travelers). Considering that the cruise travel is for leisure and in some cases quite expensive, the presence of border guards all around terminal may create a discomfort feeling. Otherwise, the remaining TRESSPASS control points and actions are in line with the current practices.

4.2 Threat Scenarios

A great number of people that do not meet the entry standards into the Schengen area, naturally view the geographical location of Greece as a potential entry point to the territory of EU. The challenges for Greece in terms of border protection are magnified by numerous factors, such as the country's extended borderline, its numerous islands, the variety of geological features, as well as its geographical location in the crossroads of Europe, Asia and Africa. Thus, all the current threat categories are considered of great importance to be included in the scenarios.

The scenarios for the Greek Pilot focused on the threat categories of terrorism, irregular migration, and smuggling. These were tested through 46 different traveler scenarios/stories. TRESSPASS through a dedicated component for the risk assessment screened the travelers on specific risk profiles created from the combination of various risk indicators. The scenarios offered a wide variety of every day occasions activating as many as possible of TRESSPASS 's components.

The creation of the threat scenarios was a quite demanding process as the Greek pilot has used all TRESSPASS components and the threat categories. In this context the scenarios had to include threats as close to reality but in a way the TRESSPASS components could handle. Through an iterative process during pilot preparation activities and in collaboration with the technical partners the required improvements and adjustments took place (both from the scenarios perspective as well from the functions of the components). During pilot operations, the scenarios have been demonstrated as planned and the outputs were the expected. Thanks to the well organisation and the training of the volunteers' participants the scenarios have been easily implemented. Another interesting finding is that in the very few cases in which the scenario has been altered, the TRESSPASS system has provided correct outputs. In this context, there is a confidence about the TRESSPASS components operations.

The scenarios implemented are presented in detail in the Greek Pilot Report (D8.4)

4.3 TRESSPASS Technology (Components/Sensors, Infrastructure)

The list of the components applied in the sea pilot consisted of almost all the TRESSPASS components and sensors, namely: TRA, VTC/RTBA, WI, C2, SPA, ICTS Document Scanner/Module, IAS/TLC, TLTP, TCA, MMCAT, LSI, DRAS, DFA and, DMS.

However, the main experience gained was from the front-end components and sensors with which traveler and border actors had interactions with them, whereas back-end components such as LSI, DFA, DRAS and DMS were mostly transparent to the users.

From the overall Hellenic Police point of view regarding the above components and sensors, it could be noted that the TRA and the SPA are friendly, easy to use applications. The MMCAT is a good tool for a novice or an intermediate user. However, this is not the case for an experienced Border Guard. Furthermore, the TLTP component runs perfectly without any disruptions. It provides a friendly UI and requires no special effort to learn.

4.3.1 TRA

The TRA application was used twice by 26 participants in the sea pilot, once used for the departure scenarios and once for the arrival scenarios:

- No technical issues observed, while the application was tested with three different types of devices (laptop, mobile phone, desktop).
- 92.31% of the participants manage to complete the account creation phase in less than 10minutes, while 96.15% of the participants needed less than 10 minutes to insert trip details.
- In general, the users could easily insert information in the application. The information input was realized and tested through android, iOS smartphones and laptops with Windows OS, all of which ran the application with no issues. The participants found easy to select destination countries, transit countries, and provide travel document information without difficulty.
- Most of the participants (96.16%) could easily navigate using the user interface of TRA
- Most participants (96.16%) stated that they would use TRA application to accelerate the border crossing, 88.46% stated that they believe that TRA application would improve security, while 92.30% stated that they would use TRA regularly for their trips.

4.3.2 VTC/RTBA

At the sea pilot the VTC component has been applied collecting the data from 4 cameras during the execution of the departure scenarios. The data have been recorded in the PAL video format (704 x 576 pixels) instead of the recommended full HD format (1080 x 1920 pixels) needed for optimal VTC performance, as specified in D8.1 pilot planning and training, page 191. Before applying VTC, all videos have been resized from the PAL video format to the full HD format. This took place as the HD videos had an extremely large size making the downloading and uploading almost impossible.

Regarding the Detection Rate: in the first two camera views all travelers have been detected. These two cameras had 100% detection rate. The third camera missed a number of travelers, because the last part of the video was missing. Only one traveler entered the view of the fourth camera. Therefore, videos from the 3rd and 4th cameras were not taken into consideration for the analysis of the detection rate. It was concluded that for the Greek pilot VTC detection of travelers was perfect.

Regarding the Re-Identification Rate: For each traveller a video was generated containing all his/her detections and inspected for Re-Identification errors. The Re-Identification Error Rate was estimated based on these 23 traveller videos.

From the technical point of view the lessons learned and recommendations are as follows:

- Using a lower resolution than the recommended HD resolution should be avoided.
- People wearing similar clothes are very difficult to distinguish.
- VTC Re-Identification Error Rate can be reduced by disallowing, that a person detection at the same place, by the same camera, at the same time, is assigned to different person identities.
- VTC Re-Identification Error Rate can be reduced by disallowing, that two person detections are assigned to two different cameras with non-overlapping views.

Based on the feedback provided by the Border operators through a questionnaire, the following 10 performance indicators were evaluated in a quantitative scoring way as follows:

- Understandability: The component was understandable regarding its concept and functionality (4,0/5)
- Learnability: The Component was easy and fast to learn. (3,7/5)
- Usability: The component was easy to use and interact since the users are involved only in the installation and setup process and in the collection of the video data. (4,0)
- Interoperability: The component can easily work with the current systems provided that the online connection with the dependent component RTBA is present on the field. (3,3/5)
- Integration: The component is well integrated with the other components e.g., with the C2 regarding the matching of the traveler with the Unique Traveler ID provided by C2 at the registration/enrollment point. However, the integration of VTC is dependent on the capability of the RTBA's integration. (4,0/5,0)
- Usefulness: The component will be useful, simplify the workload and enhance the quality of the operations (3,0/5,0).
- Information: The component would increase the amount and quality of the information. (3,3/5)
- Installation: The component can be easily installed, configured and setup. (3,7/5)
- Performance: The component seems to run and respond fast in the system. (3,7/5)
- Legal/Ethical: No legal/ethical issues related to data privacy, protection and security seemed to exist in pilot. (3,7/5)

Although the scorings for understandability, usability, and integration of the VTC component seemed to be high (score 4.0) its usefulness is doubted (score 3.0). This may be due to the fact that the operator cannot interact with VTC. To increase the usefulness of VTC "a human in the loop" would be a valuable addition to the VTC component, but that may be not allowed due to privacy restrictions.

Interoperability and Information were also doubted (scores 3.3) since the there was no online real time connection with RTBA and the results and the risk indicators expected to be raised were not available during the pilot since RTBA, as in the Dutch pilot, processed the VTC outcomes offline and after quite long time after the pilot.

Regarding the installation requirements, there were clear instructions on the data that should be provided from the site, in order to have the calibration. Several issues have been taken into account and tested in order to capture the videos in the appropriate analysis. E.g., distance of the human from the cameras, area of coverage etc. The recalibration process, as some things have changed during pilot preparation (e.g., the placement of the enrollment kiosk) took place in an acceptable time period.

The quantitative evaluation results and the re-identification rates can be found in Annex A.

4.3.3 WI

Web Intelligence was deployed to facilitate the information gathering of the involved passengers. For the needs of the scenario execution, the pilot owner facilitated the creation of synthetic social media (Facebook) profiles, which were populated according to the profile of the participating mala fide passenger personas. As soon as the travel details and social media information of each of those passengers was entered in TRA, WI was triggered to assess their digital footprint by calculating the risk indicators raised for their profiles.

WI in the Greek pilot was triggered only when the classification of the traveler from the risk assessment of the system was mala fide. The risk indicators calculated by WI in this scenario were SEA-RI-007, SEA-RI-008, SEA-RI-012, SEA-RI-095, SEA-RI-112, SEA-RI-118, SEA-RI-119.

Due to the fact that WI makes searches to social media, clear and dark web, in practice this takes a lot of time for the collection and processing of the data and therefore it is more suitable to run in the background at the pre-travel phase rather than at an ad-hoc real-time situation. This processing at the pre-travel stage enables the information to be available to the Border Authorities in the next travel phases (approach and at-bcp phases).

Issues though regarding the legal/ethical compliance may arise, and thus measures should be taken into deep consideration to mitigate any side effects.

4.3.4 C2

The C2 is the core command and control application run in the BCP. It enables the management and monitoring of all security-related data at a BCP, such as risk assessment results. In addition, operators can receive alerts (and observations reported in the SPA) and send alerts.

C2 provides a user-friendly interface, allowing border guards to better their effectiveness. The fields are searchable and at the same time filter options minimize the results to the best result. Applications such as C2 are useful as it provides information all-in-one about the ship-passengers-possible hits or alerts, without the need to change applications. The variety of this information enhances the overall awareness level, and thus border controls become more efficient.

The C2 component was demonstrated to the participants during the training sessions. C2 was accessed online and presented to the Police Officers with the intention to show them the latest version of C2 with data from the Greek Pilot. The component was tested against its ability to present information of upcoming cruises, registered passenger information, risk indicators for each passenger and related alerts. Moreover, the component was able to display specific alerts for passengers sent through the International Alert System (part of TRESSPASS). The C2 component can be adjusted to operate for different types of BCPs, while also providing a map of the BCP area. At the time of the pilot most but not all options from the main menu were functional, such as the display of Reports or the Simulator. Overall, most of the functions of the C2 worked as intended with passenger information inserted through and automatically drawn from the Traveler Companion Application (TRA) at the time of the Pilot.

Based on the feedback provided by the Border operators through a questionnaire C2 received a score 4 out of 5 in almost all the 10 evaluation parameters (understandability, learnability, usability, interoperability, integration, usefulness etc.)

4.3.5 SPA

SPA as an application it is quite easy and simple to use. Passport scanning was quick and alerts easy to enter. Hence, It worked successfully during the test.

One of the critical observations was how to locate the passenger details without scanning the travel document so that they will not be noticed by the traveler. This is something which was pointed out during but also before the tests, during the pilot planning phase. Especially in a real environment with hundreds of people in the waiting area, BG having identified a person with strange behavior, he is not able to enter risk information without knowing how the identification details of the passenger. To overcome this issue during the pilot all passengers were scanned one by one so that BG can pass a risk indicator to a traveler for whom an abnormal behavior was identified. Furthermore, risk indicators in SPA must be categorized and more risk indicators with a short description should be added as there was only one sample in the test.

Based on the feedback provided by the Border operators through a questionnaire SPA received an average score 4 out of 5 in the 10 evaluation parameters with the understandability and learnability having the top grade (5).

4.3.6 DOCUMENT SCANNER/MODULE

Although some minor wireless networking issues were experienced at the beginning of the test phases documents scanning worked properly. The module displayed the details of the passenger normally. The laptop which was connected to the scanner could also run the C2 front-end application which enables the BG to use multiple applications from one single workstation without the need to have different devices. However, the full functions of the scanners were not tested since the mockup passports used were plain papers printed with MRZ lines.

Based on the feedback provided by the Border operators Doc. Scanner received a score 5 in all the 10 evaluation parameters.

4.3.7 IAS/TLC

The TRESSPASS Lite Client was successfully demonstrated and tested during training, regarding its ability to display alerts about high-risk travelers, to create and send alerts to the C2 through the International Alert System.

TLC is a useful tool for BG but to be applicable there must be agreements in place to exchange information with the third countries. It may be better applied at land cross-border points between neighboring BCPs as at the air and sea cases an airplane or a ship may come from a distant third country, which makes it difficult for an agreement of information exchange to exist.

Furthermore, messages coming from TLC, should be categorized, so as there would be a common comprehension with the other part included. However, a variety of cultural differences or legislations might not produce the expected results.

From the feedback received through a dedicated questionnaire for TLC, the end users evaluated TLC as easy to use, easy to navigate and create alerts and easy to learn providing scores higher than 4 in the scale (1-5). They also believe the TLC system will improve border security and will be efficient to be used in third country BCPs as part of an international alert system.

4.3.8 TLTP

TLTP provides tracking of carry-on luggage and indirect tracking of travelers through the usage of Ultra-High Frequency (UHF) RFID technology.

TLTP was tested through the pilot via RFID tags which were placed onto the traveler's data at the BCP kiosk prior entering the interior area of the BCP and the check-in desks. RFID reader devices were also placed along the conveyor belt inside the BCP. Upon placing the RFID tags, the luggage was loaded onto the conveyor belts and the bags would be scanned at different points while onto the conveyor belt inside the BCP. At the same time, the baggage would be correlated with their traveler who would be passing through the interior of the BCP starting with the check-in desk, and then walking towards the security checks. The component worked as intended. Moreover, the scenario where a luggage would move along the conveyor belt and into the BCP without the correlated traveler was also tested and the component responded as intended, by provided the respective risk indicator which indicated unattended luggage inside the BCP.

In case of real-life implementation of TLTP collaboration is needed with the cruise operators for the RFID tags. Otherwise, the process of tagging of luggage can remain the same. According to current functionality there are specific steps to be followed in order TLTP to trigger (or not) some indicators. There is space for improvements so the traveler and the luggage flow to proceed independently. For the purposes of the pilot TLTP functioned as specified.

Further development of the TLTP is needed, to avoid any extra effort at the expense of border guards. TLTP might occupy no personnel (border guards) unless there is a reason for one to accompany a traveler to the second line check (mala fide travelers' cases). Thus, the controls taking place at that point might be applicable in a more automated way.

Overall, TLTP worked as it should. It is a very useful tool in terms of handling a large number of passengers and luggage. It is no coincidence that many new airports apply the RFID to separate the luggage of each flight, as it reduces the number of lost or unloaded luggage on a wrong flight.

Based on the feedback provided by the Border operators TLTP received a score 5 in almost all the 10 evaluation parameters.

4.3.9 TCA

The Traveler Companion Application (TCA) is a mobile application for the travelers, which provides useful information about their trip and the BCP, such as schedule data, boarding information, route/ navigation instructions, BCP area map details, alerts, and notifications.

The Traveler Companion Application (TCA) was presented to the participants during the training sessions of the pilot through a power point presentation and hand-on experience which included the demonstration of the application on iOS smartphones (iPhones). The aspects of the application that were tested included its interface, the ability to display traveler's information, trip information and the display of the travelers' boarding pass. The application was downloaded to four (4) iPhones and a traveler profile (from the Greek scenario travelers) was pre-created within the application. During the demonstration, the iPhones were passed around the travelers so they could see the function of the application, the provided options and information which in the Greek Pilot consisted of Trip information and the Traveler's digital Boarding pass.

Additionally, the use of the application for scanning the passengers' passport was successfully tested during the Pilot.

Regarding the intended operation of the application, TCA's strong feature was in the arrival Extra-Schengen scenarios where the notification of a bona fide traveler enables the passenger to move to a fast lane without waiting in the normal queue for passport controls. However, certain features of the application such as the display of the Pilot's Maritime BCP's map or the ability to produce location data of the passenger (inside the BCP) to the C2 were not available at the time of the Pilot.

Overall, TCA is a useful application since the trend is to replace the printed tickets and boarding passes with digital ones on mobile devices as well to include and provide prompt information and notifications to the travelers about the trip. It was very useful to test such an application in the maritime case, as similar applications are already used by many airlines to reduce the waiting time for passengers either to check in or to queue-up for passport controls in the arrivals.

4.3.10 MMCAT

For the evaluation of MMCAT, questionnaires with closed and open questions have been provided at the end of the Greek Pilot to be filled by the Border Guard officers interviewing volunteer actors in the role of mala fide travellers with the support of MMCAT in the 2nd line checks,

The assessment, lessons learned, and experience gained from MMCAT are summarized below. Detailed assessment results from the answers, which were provided in the questionnaires, are listed in Annex B.

- The scoring of the closed questions shows that the MMCAT tool is considered a helpful tool (4.7/5.0) and the training material was sufficient to use the tool (5.0/5.0). The graphical user interface is already rather good (3.7/5.0), but it could be improved by providing more information (4.7/5.0). The overall average score was 4,14/5 and comparing it with the corresponding total average score of the Polish Pilot (3,43) it is obvious that the provided feedback by the Polish end users contributed significantly to the improvements made to the tool in the meantime between the Polish and the Greek pilot.
- The end users expected the MMCAT tool will help them to make the work easier, to obtain better information and make decisions with greater confidence.
- The expectations have been met for some interviewers but not for all. One interviewer initially expected a clearer decision but realized that it was already explained as an assisting tool for the BG to help him/her take a decision. Another interviewer gave two suggestions: firstly, to present results not as continuous analysis but as analysis by question, and secondly, to add audio (in offline/remote analysis).

- The main strengths are the reliability, the speed and the relative ease of use. It is considered a strength that the video analysis can highlight something that the BG has not noticed.
- The main weaknesses are that (1) it takes too long to do all the procedures while the BG does not have that time on the field, (2) it should display the details of the passenger and not only the ID number, and (3) the graphic display analysis should be more understandable for the BG. The first two points could easily be improved while the third point would be harder.
- The suggestions for improvements in the GUI are:
  - Addition of passenger name besides the passenger number
  - The option to add notes
  - The option to add more risk indicators
  - Audio in the playback of a recording. Note that the audio was intentionally not added in the tool because it was intended as a real-time supporting tool during a live interview, however audio is off-course very valuable for a remote connection or offline processing.
- The interviewers requested the option that the BG can add new risk indicators. The list of risk indicators was already editable with a configuration file, but this was not easily accessible in the GUI.
- Although the training material was sufficient to use the tool, more time for practical training would have been appreciated to get acquainted with the indicators.
- A final suggestion indicated that it would be preferred if MMCAT would be triggered directly from C2 to open MM

4.3.11 Back End Components (LSI, DFA, DMS, DRAS)

Border guards, feel more confident conducting their duties, if they are aware of the structure and functionality of the system's core modules running at the background. Thus, TRESSPASS gradually will act no more as a black box and any outcome of the system with the proper notifications could produce the best decision.

4.4 Data and Risk/Bona Fide Indicators

TRESSPASS system collects all the necessary data from its components and sensors regarding the travelers, so that certain risk and bona fide indicators (RIs/BIs) should apply. However, when assigning Risk indicators to travelers, there is the risk that the system raises false alarms to "bona fide" travelers. Moreover, the application of the system in real-life operations involves the use of personal data which naturally raises issues of legal/ethical compliance. Furthermore, the whole infrastructure of the system should adhere to the highest security and privacy standards.

Additionally, it could be noted that most of the mockup data (used for the purpose of the pilot) were prepared in correspondence to real data. That means representative cruise trips with departure and visiting countries, type of passports, countries, addresses etc. resembled existing ones, to improve the realism of the pilot.

Regarding RIs/BIs, data was collected, and a relevant base of RIs/ Bis regarding the threats examined at the maritime BCP was created for the pilot test purpose but also for future use. However, the preparation of indicators using the TRAM tool was a difficult task and took a long time for the end users. Beside this, the categorization of the RIs/BIs according to the threat categories (smuggling, immigration, terrorism) was not considered as the most suitable or the appropriate way to define and work with the RIS and sometimes was not correct. End users believe that the indicators should be configured independently of threat categories since threats are derived as a result from the raised indicators and not as input which should be associated with the indicators in advance.

Generally speaking, the preparation of mockup data for the pilot in order from one hand to simulate representative and realistic data, implementing the newly introduced, and on the other hand to associate the data with the travelers' scenarios and indicators was a difficult and time-consuming task for the end users. It was proved that the team members who had to prepare the mockup data should have been involved also in the design of the indicators. The input of each data field should be done with cautious manner, since even small mistake in a data field could change the whole scenario and the expected results.

Finally, data consistency was very important to be ensured during the preparation phase, since many of the mockup data were used in different data base tables of the TRESSPASS system, especially in the Legacy Information Systems (e.g., TRA, Marine PNR, VIS, EES etc.).

4.5 Risk Assessment (Concept & Methods)

Risk-based concept is at the heart of the TRESSPASS system and as such, if it was introduced by EU Institutions, it could be applied with a certain amount of effort. Risk assessment is a continual process that takes place in each travel stage (pre-travel, approaching BCP and at BCP), collecting all the necessary data so that certain risk indicators, if any, match a profile that point to a threat category. Furthermore, knowledge at a pre-travel stage, allows border guards to act beforehand. Mitigating any side effect (for example bottlenecks at BCPs), in proportionality to the risk acceptance, may produce an overall positive outcome.

The accuracy of a risk assessment strongly depends on well-chosen indicators, the available data, and how reliable are the systems to provide the required data. Furthermore, and in addition to those mentioned in section 4.4, the preparation of RACS Excel-tool (Risk Assessment Configuration Support-tool) was quite complicated from the end user point of view. The support of the technical partners was required and with collaboration has been prepared. The demonstrated concept proved to be efficient, and the expected results were valid. To apply the system in real operations, an extensive impact assessment on indicators must be conducted adding new ones and optimizing the configuration, even after some trial and error.

Considering this, the processes of conducting a list of applicable RIs and the decision regarding whether Bona fide Indicators will be used for real application of such a system or not, is a thorough and long process which involves input and effort from various stakeholders such as BCP operatives, analysts, IT specialists. In addition to that, the final list of Risk Indicators alongside their business logic which defines the conditions of the indicator, the indicators' weights and the creation of subsequent risk/threat profiles greatly affects the effectiveness and accuracy of the system and its ability to currently identify legitimate threats.

4.6 Research Questions and Key Performance Indicators (KPIs)

	#	Research questions
Effectiveness	1	**Q:** What is the effectiveness of the TRESSPASS RBBM concept (and corresponding available information) in comparison to the current rule-based approach (and corresponding available information)? **A: TRESSPASS RBBM concept involves more and risk-related information than the rule-based system which can be available to the border authorities before a traveler arrives at a BCP. This could potentially overhaul the quality of the border controls, identify, and focus on risky travelers thus improving the effectiveness.** **Taking into account the limited checks of intra Schengen cruises, it could be noted that by using the TRESSPASS system, travelers from threat categories can be identified, who otherwise would not have.**
	2	**Q:** What is the probability of a false positive (i.e., people that are bona fide are sent to the manual desk) and a false negative (i.e., people that are mala fide are not send to the manual desk) **A: The probability of a false positive and a false negative could be considered relatively low. Some changes in the DRAS configuration can lower the probability.**
	3	**Q:** To what extent can the information of the TRESSPASS system potentially help to better execute the process of mitigating the threats handled by your pilot (irregular migration, smuggling etc.)? **A: As TRESSPASS can provide different types of information, as collected, and processed from the various components, it can be analyzed how it can be used and at which stages (e.g., for strategic planning of border control management).**
Flowrate	4	**Q:** What is the effect of the TRESSPASS RBBM concept on the flowrate of different categories of travelers in comparison to the current rule-based approach? **A: The TRESSPASS RBBM concept, could potentially have a positive impact on the flow rates without compromises to the security levels. Another positive impact of automation is the potential minimization of human resources needs and cut down of the relevant costs.**
Efficiency	5	**Q:** What are the resources needed to perform the RBBM-concept to achieve a desired level of flowrate and/or effectiveness? **A: Knowledge at a pre-travel and approach-to-BCP stage contributes to a better planning of human resources and operational efforts.** **Regarding the ICT resources, taking into account that cruise ships carry huge numbers of travelers, the infrastructure (physical components and software / network components) must be high end with very low latency and minimum lags.**

Acceptance	6	**Q:** To what extent perceive the participating border guards the TRESSPASS RBBM concept and system as usable? **A: The sea pilot was performed on the field of the cruise terminal under very realistic operational conditions and therefore the participating border operators with the direct interaction they had with the TRESSPASS components perceived the TRESSPASS RBBM concept and system as usable to a great extent.**
	7	**Q:** To what extent perceive the participating border guards the TRESSPASS RBBM concept and system as trustworthy? **A: Comparing the expected results with the actual results produced by the system during the different test phases of the pilot, the participating border operators perceived the TRESSPASS RBBM concept and system as trustworthy to a great extent.**
	8	**Q:** To what extent perceive the participating RNM-border guards the TRESSPASS RBBM concept and system as efficient? **A: The participating border guards perceive the TRESSPASS RBBM concept and system as efficient to a great extent.**
	9	**Q:** What are the implications of a RBBM concept for the personnel at strategic, tactical, and operational level? **A: Though the final decision belongs to the official in the fields, the risk acceptance level is of utmost importance.**
Balancing	10	**Q:** What is the tradeoff between the effectiveness, flowrate, and efficiency of the TRESSPASS RBBM concept? **A: Among these factors, risk acceptance is the tradeoff.**

4.7 Training

During the pilot, the training was carried out through analytical presentations, making use of a mixture of audiovisual material and hard copies in the form of manuals which presented informational content about the TRESSPASS project overview, the RBBM concept and the front-end components used in the Greek pilot with a step-by-step guide on how they are used.

The presentations were based on the material included in previously developed e-learning modules, developed by the TRESSPASS partner, CASRA. specifically for the project's training sessions. However, the information presented in the e-learning modules appeared too technical at times and needed to be adjusted to the Greek Pilot's target audience which included end-users from the Hellenic Police, who were not familiar with the RBBM concept and the associated new technologies. Some components, like MMCAT and C2 were updated till the last day of the pilot, so some training material had to be continuously updated.

Moreover, hands-on training was provided to both Border Guards and Traveler actors for each respective system component that was part of the Pilot scenarios. The informative step-by-step training and hands-on training complimented each other and facilitated the assimilation of training material. Furthermore, between the sessions and at the end of the training, the trainers discussed with the trainees to resolve and remaining questions and make sure that gaps in trainees' knowledge were covered before the execution of the Pilot scenarios.

At the end of the training, participants completed a training evaluation questionnaire, consisting of 19 closed questions, which had to be rated on a Likert scale (Do not agree at all to Fully agree). The questionnaire is based on the Training Evaluation Inventory (TEI), which was described in the above.

In total, 27 questionnaires were handed in completed. The questions as well as the answers given are summarized in the following. In general, participants liked the training and considered the atmosphere pleasant. The training was also perceived comprehensible and useful for their role in the pilot (Figure 14: Greek Pilot Training Evaluation Questionnaire Answers 1/4

). The majority of participants (25/27) stated that their knowledge has expanded on a long-term basis, however, a small number of participants (2/27) indicated that they will not be able to remember the new topics well (Figure 15: Greek Pilot Training Evaluation Questionnaire Answers 2/4

).

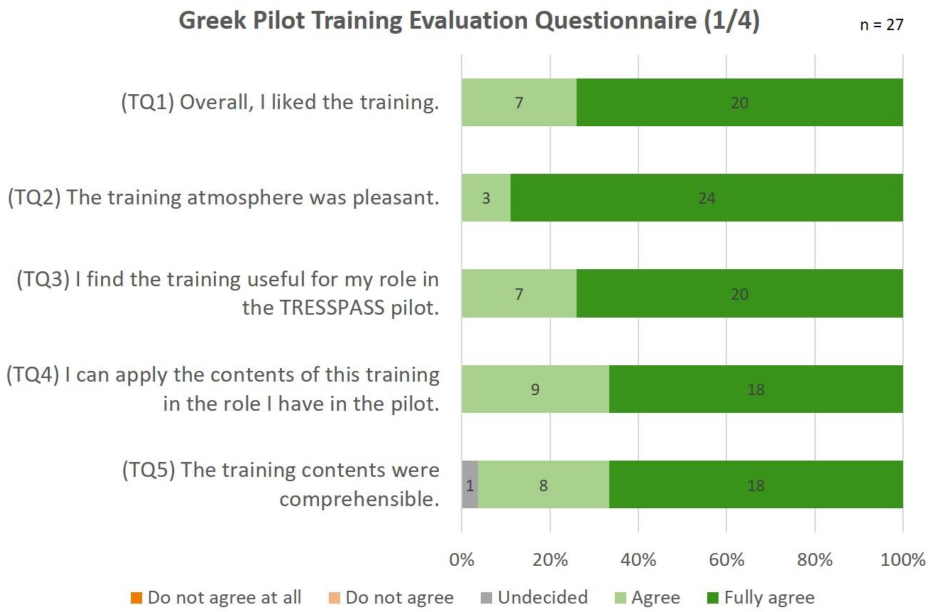


FIGURE 14: GREEK PILOT TRAINING EVALUATION QUESTIONNAIRE ANSWERS 1/4

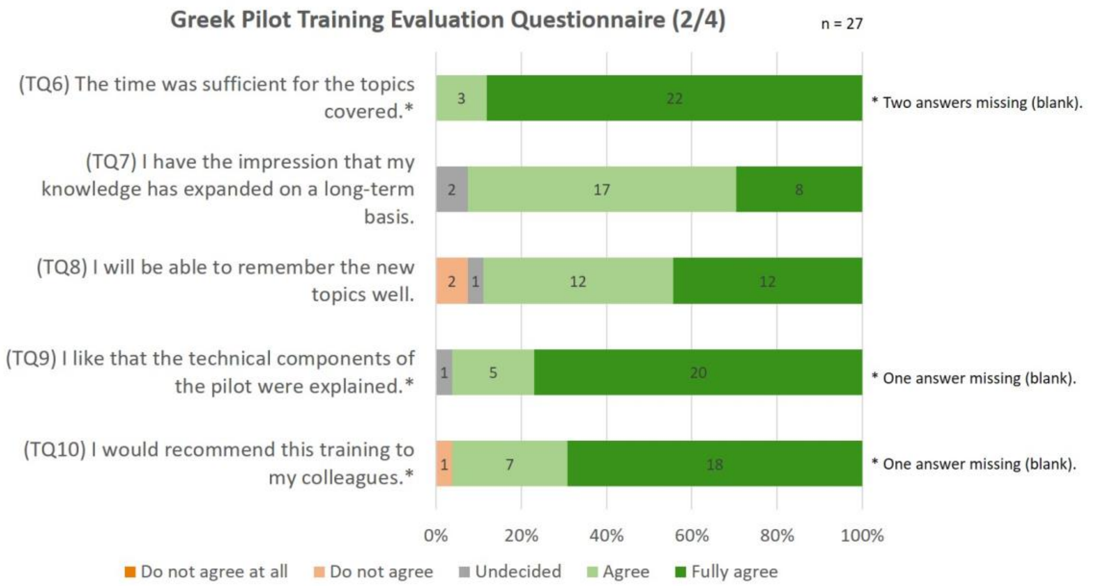


FIGURE 15: GREEK PILOT TRAINING EVALUATION QUESTIONNAIRE ANSWERS 2/4

For items 11 and 12 of the evaluation questionnaires (Figure 16: Greek Pilot Training Evaluation Questionnaire Answers 3/4

), training participants gave rather diverging answers. While most participants were able to bring previous professional experience into the training context, a considerable amount was undecided on this topic (6/27) or did not feel like they were able to use their previous professional experience (4/27). When asked, if they were invited to bring their own knowledge from day-to-day work into the training context, the majority agreed on this (17/27); however, 5 people disagreed (strongly) and another 5 were undecided.

As far as the training contents and learning objectives are concerned, participants agreed that contents were illustrated with examples, central points were made clear, and the learning objectives were achieved.

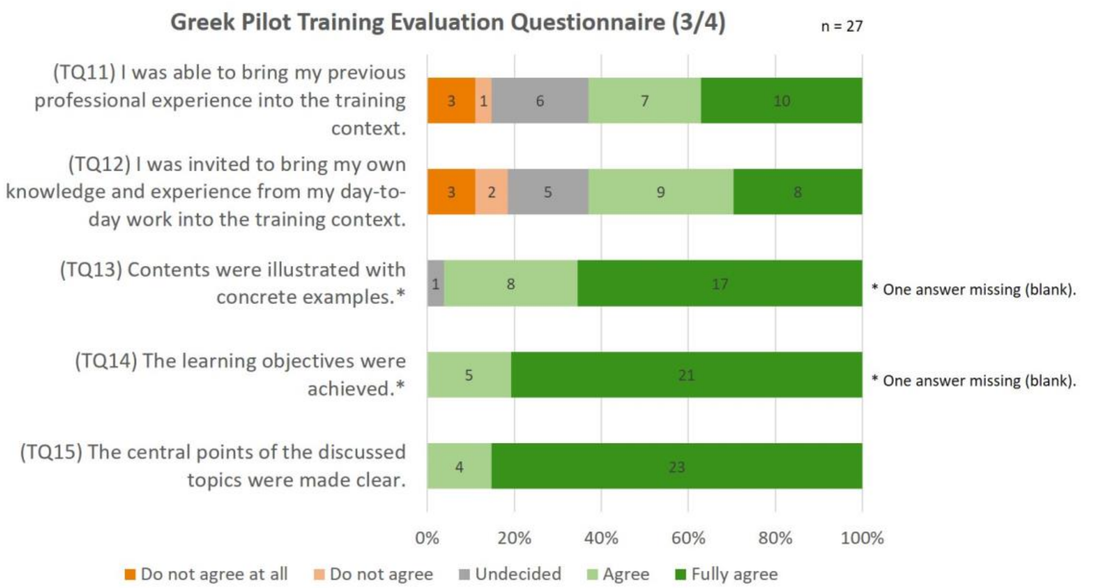


FIGURE 16: GREEK PILOT TRAINING EVALUATION QUESTIONNAIRE ANSWERS 3/4

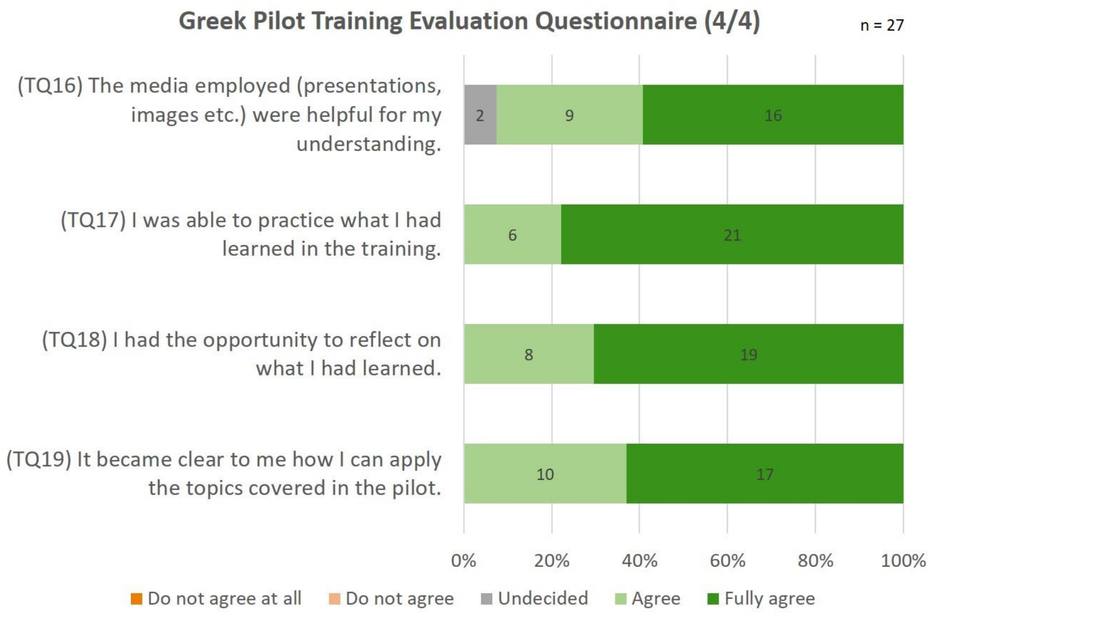


FIGURE 17: GREEK PILOT TRAINING EVALUATION QUESTIONNAIRE ANSWERS 4/4

Additionally, participants rated the media employed in the training sessions as helpful for the overall understanding, and agreed that they were able to practice what they have learned in the training, as well as that they had the opportunity to reflect on the learning contents (Figure 17: Greek Pilot Training Evaluation Questionnaire Answers 4/4

).

Finally, through discussion with the participants after the training session, the Police Cadets who had no previous knowledge about the TRESSPASS system pointed out that the training included a lot of new information sometimes difficult to comprehend. This is a consequence of the complexity of the TRESSPASS system and the many components it consists of. Additionally, unlikely the experienced Police Officers, most of the Police Cadets did not have any previous hands-on experience with border checks procedures. This means that their understanding of the TRESSPASS system, processes and its added value to current operation was somewhat limited and dependent on solely on their theoretical knowledge of procedures at a BCP.

In spite this, all the participants expressed their positive experience and admitted that the training provided important knowledge for future consideration.

Considering the above, for training improvement, previous work experience from participants should be taken (more) into consideration, to establish a better connection between previous knowledge and new training contents.

4.8 Simulations

4.8.1 Results from FHG Simulator in the Greek Pilot

This section describes the different scenarios used in the FHG sea pilot simulations, the process flows in each scenario, the performance areas and the corresponding key performance indicators observed under multiple simulation runs executed using different configurations in the Fraunhofer's simulator.

Scenarios: While in the experimental part of the sea pilot two major risk-based cruise travel scenarios were used: (Departure Intra-Schengen and Arrival Extra-Schengen) in the simulated part we focused on the cruise departure scenarios as the most complicated scenario in terms of processes and components, exploring both the Intra and Extra Schengen cases and evaluating and comparing them versus the corresponding rule-based scenarios having 4 scenarios in total as follows:

1. Risk-Based Cruise Departure Intra-Schengen scenario
2. Rule-Based Cruise Departure Intra-Schengen scenario
3. Risk-Based Cruise Departure Extra-Schengen scenario
4. Rule-Based Cruise Departure Extra-Schengen scenario

For each of the above scenarios different configurations were used expanding the 4 scenarios into multiple scenarios depending on the input parameters used in the simulation tool. A total of 25 configurations were run based on the inputs provided by the end-users.

Process Flows: For each set of Intra/Extra Schengen - Risk/Rule-based cases one set of flow process diagrams was drawn as depicted in the two figures below. The flow process diagrams consist of stages, namely the pre-travel, approach-to-bcp, at-bcp and post-bcp travel phases.
Each stage consists of component groups which represent the checking and screening points in the process flow of the sea pilot.
Each component group in turn includes, the service time, the TRESSPASS components and the human resources involved at each component group (Border Guards, Security operators etc.) presented as Border components.
The flow diagrams also include the distances between the component groups, and the different paths a traveler can take depending on the travelers' risk level detected at each component group/checkpoint.

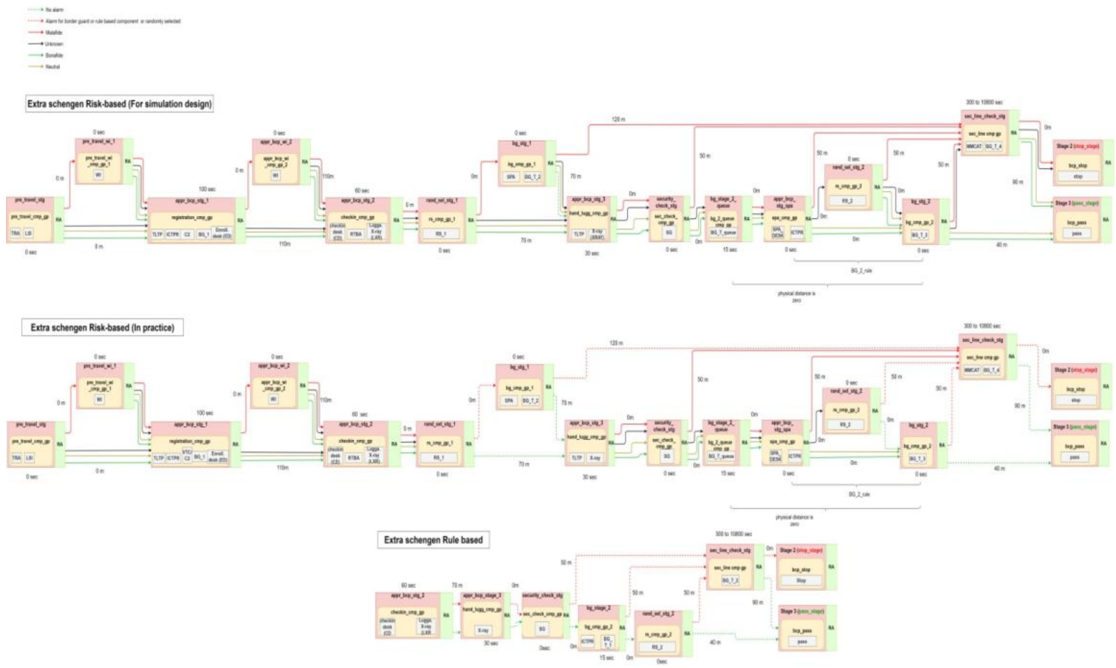


FIGURE 18: SEA PILOT FHG SIMULATION: EXTRA-SCHENGEN RULE & RISK BASED PROCESS FLOW DIAGRAMS

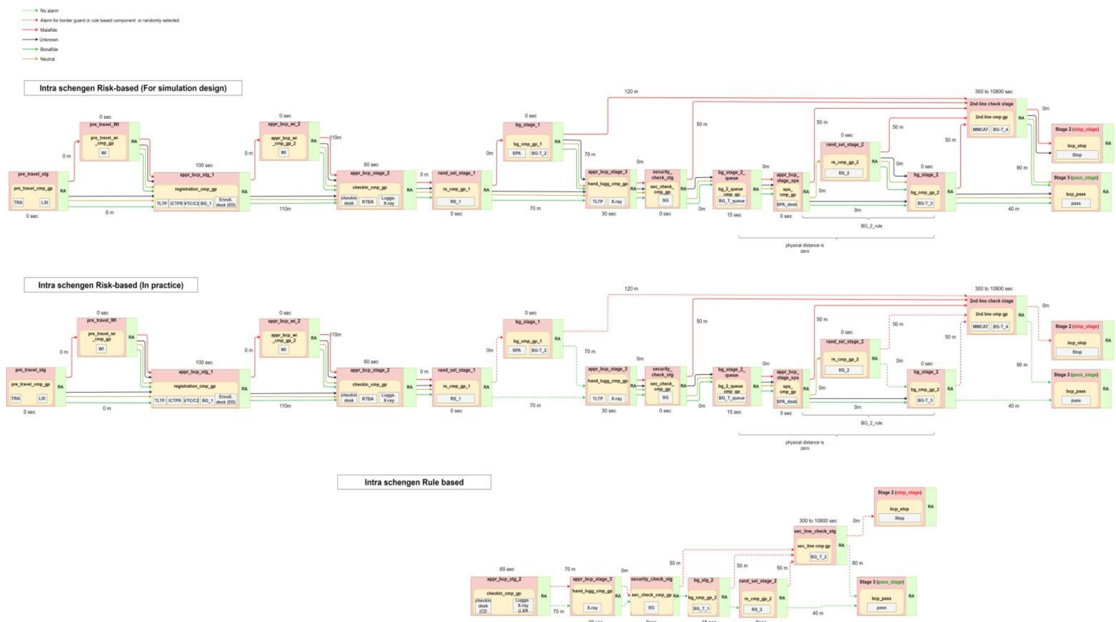


FIGURE 19: SEA PILOT FHG SIMULATION: INTRA-SCHENGEN RULE & RISK BASED PROCESS FLOW DIAGRAMS

<u>Performance Areas/Indicators – Observations:</u> In the simulations, the following performance areas were observed, associated with the three corresponding key performance indicators: flowrate, effectiveness and efficiency.

1. Flow related (Flow rate)
2. Access and egress control (Effectiveness)
3. Resources usage (Efficiency)

In the following, the observations from the simulations are listed:

1. <u>For the Flow and Resources related simulation</u>, the research question used was: *"What are the optimized number of resources for the risk-based and rule-based scenario while keeping the average waiting time of the travelers within an acceptable range?".*

   With the term 'resources' used in the above question we are referring to the number of 'service points', which in the simulation tool are represented by the 'component groups' in the chain of the above process flow diagrams, where each component group (xxxxxxx_cmp_gp) corresponds to one check point.
   The number of service points determine in turn the human resources needed, thus the question above tries to answer a combination of both the flow rate and the efficiency aspects.
   i.e keeping the travelers' waiting time to minimum, what are the optimum resources needed for an optimum arrival rate and what is the optimum arrival rate for an optimum number of resources that can be available at certain component groups/checkpoints.

   To evaluate these behaviors, several simulation runs were performed by Fraunhofer with different arrival rates and number of service points: one set for the risk-based and one for the rule-based scenario in order to have the comparisons between the two sets.

A summary of different arrival rates and service points needed within the BCP for the Extra-Schengen departure scenario is as follows:

Extra-Schengen	Rule-based			Risk-based		
Arrival rate (travelers/hour)	**360**	**720**	**126 (optimized)**	**360**	**720**	**144 (optimized)**
COMPONENT GROUPS & SERVICE POINTS						
registration_cmp_gp				11	21	4
checkin_cmp_gp	20	38	7	7	13	3
hand_lugg_cmp_gp	4	7	4	4	7	3
bg_2_queue_cmp_gp	5	5	2	2	4	1
sec_line_cmp_gp	5	8	2	3	5	2
THROUGHPUTS & WAITING TIMES						
Throughput (travelers/min)	6,114	11,317	2,121	5,7732717	10,22422	2,368143
Average waiting time (Min)	1,165	1,589	6,386	1,9395767	1,778641	15,33501

FIGURE 20: SEA PILOT FHG SIMULATION: OBSERVATIONS IN THE FLOW AND EFFICIENCY/RESOURCES RELATED PERFORMANCE AREAS

From the above results it was observed that in the risk-based case most of the resources needed to establish will be the enrollment desks at the registration component group/check point, where 11 desks will be needed for an arrival rate of 360 travelers/hour, 21 desks for 720 arrival rates whereas for an optimal 4 enrolment desks the optimum arrival rate would be 144 travelers/hour.
In other words, if the BCP is operated on peak arrival rates (360 and 720 travelers/hour) for a very long duration, a lot of resources are needed to keep the waiting time of the travelers' minimum.
Therefore, for a risk based BCP the bottleneck is identified to be at the registration desk (registration_cmp_gp) which has an average service time of 100 sec. But for a rule based BCP this bottleneck is identified at the check in desk (checkin_cmp_gp).

Simulation tools are very efficient to optimize the resources needed at a BCP.
<u>For the risk-based scenario:</u> in the Greek BCP, the maximum number of registration desks that can be installed are four (4). Because of this, longer queue appears for arrival rates higher than 144 travelers/hour. Although there are more service points available for check in desk, hand luggage check and border guard booth but they won't provide any added benefit of increased throughput because the overall bottleneck of the system is at registration desk.
<u>For the rule-based scenario:</u> using a similar strategy, the optimized arrival rate is also determined. It can be seen that in order to maintain minimum queuing at an arrival rate of 126 travelers/hour, a rule-based BCP needs a more check in-desk, hand luggage check, border guard booth to keep minimum queuing within the BCP. For the rule based BCP the bottleneck is identified at the second line check. This bottleneck is very sensitive to the number of malafide travelers and the false alarm rates of the border guard.

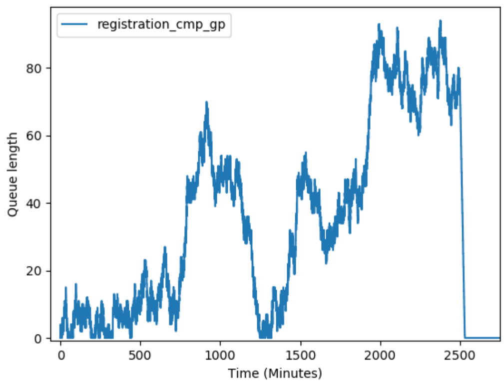


FIGURE 21: EVOLUTION OF QUEUE LENGTH IN RISK-BASED AT REGISTRATION DESK FOR THE ARRIVAL RATE OF 144 TRAVELERS/HOUR. IT REPRESENTS OVERALL QUEUE AT ALL THE 4 SERVICE POINTS.

4. <u>For the effectiveness related simulation</u>, two main parameters were evaluated:

   a) <u>The True Detection Rates (TDRs)</u> which were observed that in both intra and extra-Schengen scenarios the risk-based configuration has overall a higher detection rate (95.7% and 95.4%) than the corresponding rule-based (88.9% and 92,8%) configurations respectively.
   b) <u>The False Alarm Rates</u> (FARs), which apart from the effectiveness are also related to the flow rate, were also observed that in both intra and extra-Schengen scenarios were a bit higher for the risk-based scenario.

Departure Scenarios	Rule-based		Risk-based	
	Detection rate	False alarm rate	Detection rate	False alarm rate
Intra schengen	0.888888889	0.013088346	0.957300275	0.014835605
Extra schengen	0.927858924	0.012185535	0.954554555	0.01407334
	0.908373907	0.012636941	0.955927415	0.014454472

Figure 22: SEA Pilot FHG Simulation: Observations In The Effectiveness Related Performance Areas

For a detailed analysis of the results, one should also need to invest efforts in observing the queuing behavior within the BCP and how does it change when changing the service points. Changing the number of service points will also increase the cost of used resources. All these results are available within the simulator output files but for a detailed analysis of the results, the domain knowledge of the border operation is needed. Although the simulator was developed with close collaboration with the end-users, the analysis still highly relies on a thorough understanding of border operations and thus could be improved by the end-users with the support of their technical partners. More informatin on how to use the simulator to optimize a BCP is provided in D7.8.

4.8.2 Results from iCrowd Simulator in the Greek Pilot

The iCrowd simulator was used to simulate different BCP configurations in the context of the Greek Pilot. A specific BCP-flow for each pilot case was implemented by NCSRD in the form of a Behavior-Tree. A single behavior tree covered all normal and special cases of a passenger going through the BCP, in a dynamic and easily expandable way to allow for last-minute changes. Various parameters regarding the simulated initial DRAS assessments and their changes, service time distributions, and internal mechanisms such as queuing and random sampling by simulated Border Guards, were responsible for guiding the execution of said behavior tree, so that each passenger could have their own personalized flow through the BCP. These input parameters are presented and described in Annex C and a sample of results is shown in Annex D. The parameters were initialized by the configurations (set of values for all required parameters) that the end-users provided.

The end-users were trained to use the iCrowd simulator through 2 simulation workshops that NCSRD organized. During the workshops, all attendees had access to isolated virtual environments where they had the opportunity to start and manipulate the simulator on their own, while being guided by the trainer.

- The first workshop was intended solely for training. It was split into 4 sessions, to enable a hands-on experience for all attendees at the same time and to ensure enough time for the trainer to tend to each trainee as needed. It was noted by both NCSRD and the end-users that the splitting of the workshop into multiple sessions was useful.
- The second workshop was intended for the end-users to run the simulations with their configurations and extract results and was conducted in one session. The end-users had submitted a number of configurations for their respective pilot case prior to the workshop, which had been imported into iCrowd in the form of initialization scripts by NCSRD. The end-users had the time to execute any configuration they wished, adjust them as required, and extract new results. Some end-users suggested that this workshop might have benefited from being split into 2 sessions. NCSRD agreed but considered the workshop successful, nevertheless.

- Complementary sessions were requested by some end-users after the second workshop, in order to run simulations with more configurations and extract new results. In these workshops, the end-users were given access to a virtual environment while a trainer was on-call to provide any assistance.

During the workshops, the end-users had the option to adjust the input parameters and see the simulation in real-time. They also had real-time access to the KPIs that were being measured and collected by iCrowd. These were in the form of numbers or time-series depicted as graphs, and are presented and described in Annex C. The KPIs, along with information regarding the actions of the simulated passengers, were extracted to .csv (semicolon-delimited text) files, and all results that represented time-series were plotted to graphs and saved to .png image files. For each configuration provided by the end-users, an archive containing the .csv and .png files was created and passed onto the end-users for review. A sample of the content of the output files is shown in Annex D. The end-users assessed the format of the results overall positively. There was a recommendation during the second workshop that it would be useful if some of the files contained the initialization parameters in them, to allow for easier organization and post-processing. The recommendation was noted by NCSRD for future development.

Both workshops concluded without any issues, technical or otherwise. The iCrowd simulator and the virtual environments functioned properly, and the end-users were satisfied with the simulation of their respective pilot cases. Feedback was given by all attendees at the end of each session, which was positive about the overall experience and included suggestions regarding additional features that they would have liked to see in the simulator. These mainly included adjustments of- or additions to the Graphical User Interface (GUI), that would have made the usage easier and the overall experience smoother. Functional features were also requested, such as the option to reverse the simulation or reset it while retaining any adjusted settings. Some suggestions from the first workshop were quickly implemented and were available for the second workshop. The rest of the suggestions were assessed by NCSRD and noted for future development. It was also requested that a user manual be available for the trainees, which was distributed to end-users for the second workshop and can be found as Annex attached to the deliverable D8.4 “Training Material”.

4.8.3 End Users’ Lessons Learnt from simulations in the Greek Pilot

The simulators tested (iCrowd and FhG), are tools to run a variety of scenarios in a specific BCP, allowing the user to tune a set of parameters and instantly view the outcome graphically. End users could evaluate their BCP, by understanding the passenger flows for different number of passengers and arrival rates. The simulators may provide tools for optimizing the processes taking place at a BCP, setting the parameters to the desired values.

iCrowd simulator (agent-based behavior) has a very nice 3D visualization and the interface provides a user-friendly menu for some of the parameters’ customization and for the rest of them hard-coding is needed (number of border guards, place etc.). The number of parameters provides the means to create more realistic scenarios, taking into consideration any special measure in effect, such as social distances. Also, many parameters deal with profile values (bona fide, mala fide), physical properties, speed etc. Furthermore, if needed a whole behavioral tree of the passengers may be exported. For an inexperienced user is difficult for him to use it; Training and manual is needed.

FhG simulator (Monte-Carlo block diagram model) provides statistics for a risk-based security system. It has a simple design with text files to set values, although syntax errors may affect the simulation and many parameters to touch reality. It can measure detection probability for different threat scenarios and classify the travelers (mala fide-bona fide) based on distribution of risk indicators. We can verify resources needed to avoid bottlenecks and calculate relevant costs. Finally, statistics can be exported regarding average checking time and waiting time per traveler.

Overall, both simulation tools enabled the end users to move from theory to practice and to identify and realize those parameters which affect the new RBBM system so that to adjust them appropriately and apply them with iterative trials in the practical implementation of such a system in order to achieve the best performance indicators (effectiveness, flowrate, efficiency etc.).

Regarding the involvement in the preparation and the operation of the simulation environments, the end users expressed the following views and recommendations based on the experience they had from the two simulators:

- The configuration and the setup of the simulators started right after the experimental part of the pilot and therefore the time given to the end users to prepare and provide the requested simulation input parameters was relatively short. The end users' involvement could have started long time before the experimental pilot but on the other hand, the finalization of experimental part of the pilot also finalized the pilot's flow of processes, the resources needed and the risk configurations which all contributed significantly to the simulations' set-ups.
- End users should have a user manual long time in advance in order to understand, how the simulations process and software works.
- Training sessions on both simulators were very helpful but hands-on experience was very limited. End users should have a reasonable time to practice with the simulators on their own to fully understand how they work and be able to try different configurations with different input parameters so that at the end to achieve the desired outcomes in the different performance areas and indicators.
- End users due to the limited time didn't have the chance to install the simulation software at their sites and the interaction with the simulators was limited only during the online demonstration and training sessions.
- Simulation output data should be presented in a simple, easy, and understandable way so that the end users be able to interpret and evaluate the results on their own without the aid of the technical partners.
- A close collaboration between end users and technical partners is essential, so that both parties fully understand what is requested and expected from each side.
- During the simulation preparation stage all input/output parameters should be gathered in one file so everyone can have access to them easily and be able to check in each simulation trial run which set of results corresponds to which set of inputs and make the appropriate comparisons.
- It is preferable by the end users each simulator to have a GUI to enter online their input parameters, instead of having them entered in script text files stored in their local folders.

4.9 Overall Pilot Lessons, Limitations and Obstacles

The pilot was organized in an exemplary manner and the execution was successful taking to the account the time limitations and the complexities compared to the other pilots, such as the number of threat scenarios, the number of indicators, the number and the variety of technical components, the variety of cruise departure and arrival scenarios combined with Intra and Extra Schengen cases etc. However, some small constraints and obstacles were observed. We didn't see live the VTC component and the RTBA results, as VTC was in recording mode the day of the pilot and the video footage would be processed by the RTBA component to provide results after the pilot.

Regarding the SPA and MMCAT components, there were many behavioral indicators, which could be inserted via these components, but it would be useful if some of these indicators could be inserted to the system from a security officer handling the CCTV cameras in the control room by clicking the figure of the traveler.

Also, in order, the TLTP component, to function properly, a border guard had to push a button in his tablet every time after a traveler was passing his/her handbag through the x-ray, to update the system. In real situation a border guard can't press a button all the time, this must be done somehow automatically.

From the end users' point of view it was seen positively anything that would help to be effective in the controls, while at the same time it will reduce the waiting time and the inconvenience of the passengers.

As crime knows no borders, international and cross-border co-operation must now be taken for granted, to address threats and perform better border control. An effective communication and exchange of information between border organizations, better cooperative mechanisms between different countries and agencies must be created. If promoted through specific technology - networking activities - specific protocols - specific agreements between countries, cooperation can be improved at local, regional, EU and international level.

At local BCP an analysis tool which takes into equal account flexibility, range of options and user friendliness, so that we carry out more efficient checks at the BCP, to detect sets of different combinations of threats - modus operandi, risk indicators with risk scores calculations and assessment both in the pre-arrival stage and at the stage of BCPs checks, so that we can improve the efficiency of our daily activity.

We need to adopt technologies such as biometric, data fusion and pattern-based threat prediction systems, so that we can improve detection.

we need to employ risk-based border management, so that we can increase the efficiency by focusing the attention on high-risk individuals.

For all these technologies, however, the appropriate training must be carried out, to make the right use of new technologies and to deliver on what they aim for.

In general, with the large and massive movement of passengers such as the cruise, a tool that will make a significant contribution to cross-border control in terms of efficiency and speed must be adopted.

5 SUMMARY & CONCLUSIONS

This chapter summarizes the lessons learnt from all 3 pilots and the overall experience gained throughout the implementation of the new TRESSPASS RBBM concept and system during the project, concluding with the achievement of goals, the feasibility of the technological challenges and the synthesis of results around the four performance areas i.e. the flow rate, effectiveness, efficiency and legal/ethical compliance.

5.1 Goals Achievement

The goal of this task (T8.5) as stated in the introduction section was to present the lessons learnt and the experienced gained through the pilots and throughout the project in general and finally compare and evaluate those against the fulfillment of the following which are analyzed in the sub section below:

- the initial end-user requirements, needs and expectations (T1.1/D1.1)
- the objectives and activities set in the initial pilots' plan (T8.1/D8.1)
- the goals and activities of each individual pilot (T8.2/D8.2, T8.2/D8.3, T8.3/D8.4)
- the overall project's objectives and goals (DoA)

5.1.1 Fulfillment of the end-users' requirements

The initial requirements, needs and expectations have been recorded and described in D1.1 based on the available literature, surveys, and questionnaires with participants from various countries. In general, the end-user requirements that were relevant to the pilots were related to aspects such as operational capacity, risk/ threat identification, risk assessment, simulation tools and capabilities, legal and ethical compliance, threats detection capacity, threat alerts, pre-travel screenings, exploitation of OSINT data, suspicious behavior of travelers, exchange of information with existing systems and databases.
More specifically the main end user requirements from the long list in D1.1 which were assessed during the pilots were the following:

- Detection capabilities to new/improved threat methods & Modus Operandi though new sensing and data fusion technologies
- Detection of risk indicators with risk scores calculations and assessment both in the pre-arrival stage and at the stage of BCPs checks
- To recognize behaviors or intents and raise threats in the form of alerts/notifications.
- International and cross-border Collaboration/co-operation, sharing/exchanging info between different countries and BCPs
- Multiagency cooperation, communication, and collaboration/consolidation of various authorities' functions (cost savings, efficiencies)
- Establish common border control infrastructure and equipment at BCPs
- Employ risk-based border management resources
- Integration of OSINT information such as social media, clear and dark web for further search and analysis of mala fide passengers before their arrival at BCPs.
- Analysis Tool-Flexibility, user friendliness

Overall, the assessment of the above user requirements as defined at the first stage of the project showed that even if they were quite ambitious in principle have been fulfilled during the pilots through the implementation and testing of the TRESSPASS methodologies and the components developed in the project. A characteristic example is the requirements related to traveler identification, behavioral analysis, and any associated risks' detections. TRESSPASS has managed to have a set of tools for traveler identification and risks detection that allow the border authorities to easily know "who is", "what is carrying", the results of the risk assessment etc.

5.1.2 The achievement of the pilots' planned objectives and activities

The planned activities in the pilots' design phase described in D8.1 were all implemented following by the book all the phases and the steps of the Pilots Guidance Methodology (PGM).
The implementation of the plan is described in detail in the pilot deliverables (D8.2, D8.3, D8.4) and include the implementation of the planned threat scenarios, the mockup data and the technical components used, the execution flow of the risk-based processes in all travel phases, the ICT infrastructure and resources deployed, the human resources and the various actors' roles involved, the blended training methodologies and process applied accompanied with the training material published and finally the evaluation approaches and tools used for the assessment of the pilots.

5.1.3 The goals of the individual pilots

- The goal of the Dutch pilot was to test in the air modality the performance and potential of the TRESSPASS RBBM concept and system in terms of the balance between the following objective elements: effectiveness, efficiency, acceptance, flowrate, and feasibility. In other words, the goal was to discover if RBBM is indeed a valid concept for the future when the flow of passengers will increase, and available personnel will be scarce. One of the capacity consuming but not only, threats tested was the inbound irregular migration. With RBBM, the effectiveness of detecting all relevant threats should be equal or better than the current situation and/or adjustable based on varying the risk acceptance upon actual urgency and circumstances.
  To achieve the pilot's goal a set of research questions were formulated per objective element in order to be answered by the experiment and/or simulation part.
  To answer the research questions the TRESSPASS Key Performance Indicators (KPI's) or equivalent were used per objective element for both the current system approach and the future RBBM system approach using and evaluating the feasibility of the TRESSPASS system through the C2 and SPA components and focusing on the assessment of Border Guards' decision making in different conditions. Detailed analysis of the research approaches, the research questions and the associated KPIs per objective element for the experimental part of the pilot can be found in the Dutch pilot report (D8.2) whereas the corresponding outcomes for the simulation part of the pilot are included in D8.5.
- Similarly, the goal of the Polish pilot was to validate the RBBM concept in a land border setting, specifically in the performance areas of effectiveness, flow rate, usability and validity of the operational usefulness and feasibility of using specific innovative technical components in the RBBM land border concept. The goal was achieved with the joint use of web intelligence, behavioral analysis, information from third countries, and data from various sources (Legacy data bases and Open Source from web and social media) which provided additional data for risk-based screenings not only at the BCP check point but before travelers arrive at the BCP. The Polish pilot was focused on three scenarios which were appropriately formulated to detect and tackle attempts of illegal entries of TCN's and cross-border crimes such as smuggling illicit goods to EU carried and hidden in

vehicles with the introduction and tests of new sensors and technologies, not currently available at the BCP. More details on the implementation can be found in the Land pilot report (D8.3) whereas the outcomes for the simulated part of the pilot are included in D8.5.

- Finally, the goal of the Greek-sea pilot was to use Intra and Extra Schengen departure and arrival scenarios with Cruise travelers arriving or departing to and from Piraeus Port with Cruise ships through the Piraeus Port Authority (PPA) which hosts an integrated set of services such as Hellenic Police Border Authorities (HP), Customs Authorities (IAPR), Security Operators (PPA), Check-in desks and other essential services for the passengers. The main concept of the scenarios and the pilot in general was the introduction of a maritime PNR approach for risk-based screening enhanced with risk related information for travelers and luggage identification and verification technologies.
  The implementation of the pilot was based on the planning and training content initially produced in task T8.1 and finally described in detail in D8.4 including the final version of the training material available not only for the sea pilot users but for all the TRESSPASS end users and stakeholders.
  The Greek pilot, apart from the introduction of a unified PNR, focused on aspects related to end users' multiagency cooperation, collaboration, and information/exchange at local/national level between Hellenic Police, Customs, Security Operators etc. and the corresponding aspects at international and cross-border level with non-Schengen countries testing and validating risk-based alerting/notifications features.
  The sea pilot as last pilot in the project implemented the RBBM concept and tested almost all the TRESSPASS components in their final mature state and it was one of the 3 pilots which was performed under the most realistic operational environment and conditions.

5.2 Feasibility of Technological Challenges

Overall, Border operators in all pilots made use almost all the TRESSPASS components covering all the travel phases (pre-travel, approach-to-BCP, at-BCP and post-BCP) in all 3 modalities (air, land, sea).

More specifically, information gathering and availability at a pre-travel stage via the introduction of a unified PNR (Marine, Airline) and the use of TRA facilitate through the LSI and WI components the initiation of passengers' screenings before they arrive at the BCP. This increases the quantity of the current rule-based information and enhances the quality of border controls and at the same time guarantees the effectiveness of passing through, in proportionality to the measures in effect. Prerequisite for the correct screening outcomes is the correct initial setup of the risk and bona fide indicators and profiles in the DRAS configuration system applying the TRAM method and tools used at operational level and designed and developed at strategic and tactical level.

Technologies in the approach-to-BCP phase such as VTC/RTBA, TLC/IAS and SPA with behavior identification and alerting/notifications features can further increase and enhance the quantity and quality of information providing dynamic risk assessment via DFA and DRAS to the Border Authorities before they actually interact directly with the travelers at the BCP phase. However, regarding the on-the-move real-time behavioral identification and analysis through VTC/RTBA, the results are at basic level but with potentialities to be improved if smarter techniques are available for collection and analysis.

Finally, technologies such as TLTP, Document Scanners, C2 front-end application, SPA and MMCAT can easily be implemented at-and-post-BCP phases and in the second-line checks provided that the appropriate ICT infrastructure with the local networking and the remote communications are in place at the BCP and at the remote sites where the back end TRESSPASS components are hosted.

However, while the TRESSPASS system introduced many new innovative technologies, before their final deployment in a real operational environment, they need some improvements, possible customizations per BCP and more tests.

After all, the introduction of a system like TRESSPASS either by introducing or consolidating technical components and sensors in the classic border control operational environment, is very much feasible provided that these are compliant to the current and future Schengen Borders Code and do not neglect the current or the future rule-based system. Selections and variations of technologies can be applied per modality or country according to the specific rules and requirements of the BCPs. We should also note the possibility that the new risk-based system or parts of it may become part of the rule-based system in the future.

Regarding the simulator tools (iCrowd and FhG), they enable end users at strategic and tactical level to setup and run various scenarios in a specific BCP, testing different set of parameters such as number of passengers, arrival rates, percentages of different traveler profile types crossing the borders etc and fine tune them accordingly to optimize the resources needed and achieve the desired values such as service points, waiting times, throughputs etc. Simulation tools act as a kind of resources planning tools and were proved as good means to explore and evaluate the key performance areas (flow rate, effectiveness, and efficiency) of a risk-based system, and finally compare them against the corresponding indicators in the rule-based system, in contrast with the tools used in the experimental part where these performance indicators are not easy to be calculated in practice. Nevertheless, simulation tools need more hands-on experience and a user-friendly GUI environment for the end users to be able to enter the input parameters and get the output results without the support of the technical partners.

5.3 Key Performance Areas and their interactions

This section describes what can be learnt from the three pilots about the main performance areas of RBBM (D1.2, section 4.1). These areas are described in Table 5-1. This section synthesizes the results for each of these areas and provides some insight regarding their main interactions.

TABLE 5-1 BCP PERFORMANCE INDICATORS

Indicators	Description
Effectiveness	Success-rate of stopping unauthorised travellers when they attempt to cross the border at the BCP
Flow-rate	Speed of the flow of travellers as they approach and cross the border at the BCP
Efficiency	Number of resources required at the BCP to achieve a certain degree of effectiveness and/or certain minimal flow-rate
Level of ethical compliance	Extent to which a BCP mitigates negative ethical impact on the travelling public and on the public in general

Each successive pilot learnt and built upon the results and design (e.g., business logic and indicators) of the previous pilot, and added indicators and checks for one extra threat, while simultaneously adapting the BCP design and component-configuration each time to another travel modality: from air, to land, to maritime. In addition, the respective border guard organisations had different tactical goals with their respective BCP design. Table 5-2 describes and summarizes these differences. These variations were necessary to obtain relevant results for a wide spectrum of BCP situations with a small set of pilots but complicate comparing and synthesizing the results.

TABLE 5-2 THREE PILOTS DIFFERED ON TRAVEL MODALITY, TACTICAL GOALS, RANGE OF EFFECTIVENESS AND CONFIGURATIONS.

		1st pilot	2nd pilot	3rd pilot
Travel modality		Air	Land	Maritime
Border guard(s)		RNM & AAS	PBG	HP, IAPR & PPA
Tactical goals is to improve on	Effectiveness	X	X	X
	Flow rate	X	X	X
	Efficiency	X		X
Threats (i.e., range of effectiveness)	Identity fraud[22]	X	X	X
	Irregular migration	X	X	X
	Smuggling		X	X
	Terrorism			X
Identity management		TravelDoc	TravelDoc + TCSS	TravelDoc
Initial screening data sources	Legacy Systems (PNR, API, EES, VIS, SISII, ETIAS, etc.)	X	X	X
	TRA	X	X	X
	SPA	X	X	X
	VTC + RTBA	X		X
	TLTP			X
	TLC/IAS		X	X
Dependent screening data sources	WI	X	X	X
Check altered	First line check – based on	X	X	X

[22] Identity fraud is considered mandatory. It is mandatory by law, and weak identity management can lead to confusion about which screenings belong to which travellers.

	Schengen Article 8, 3a(v)			
	Second line check Supported by ISS (MMCAT)		X	X
	Vehicle inspection		X	

This section provides a limited analysis which focusses on a selection of high-level questions. The TRESSPASS results include several research tools that can be used beyond the TRESSPASS project to perform more in-depth or extensive analysis, most importantly the two simulators developed by FHG and NCSRD.

The TRESSPASS pilots operated at TRL7 because real border guards were used in the pilots in real border crossing point environments. However, no real travellers – or real traveller's data – were used, limiting the representativeness of the outcomes of the pilots[23]. This affects all performance areas but is probably of particular relevance for the performance area "Effectiveness / stopping power". To put this into perspective, this is *not* an analysis about real life effects[24], which should be done at higher TRL levels and can only be done with real travellers.

5.3.1 Anchoring ethical and legal assumptions underpinning the pilots

The three kinds of impelling principles of effectiveness, flow rate and efficiency should always be weighed against the anchoring principle of ethical compliance (D1.2 section 4.1 and D9.8).

RBBM, the TRESSPASS project, and each pilot had as ethical starting point that it is ethically good to make first- and second line checks[25] more proportional to the perceived threat per traveller. This idea was further refined to also make the processing of more invasive data sources dependent on the outcome of earlier screenings which makes it possible to limit the processing of personal data only to high-risk travellers. For example, Table 5-2 shows that the Web Intelligence component was only used if the outcome of the initial screening gave cause to do so.

TRESSPASS used the method of profiling and the concepts of indicators and profiles to methodologically creating screenings. Profiling is inherent to decision making in the face of uncertainty. For example, it could be argued that the current first line check in rule based BCPs where travellers are interviewed about purpose, duration and means (Schengen article 8, 3a(v) is already a type of profiling. The reason is that the border guard does not have the means to verify (check) whether a traveller actually has the purpose and means for an

[23] TRESSPASS D5.1 about the TRESSPASS Architecture and Use cases describes how the consortium explored various ways to use more representative data, perhaps even from real travelers. Ultimately, this proved not feasible due to ethical, legal and risk management reasons.

[24] The Maryland Scientific Methods Scale (MSMS) can be used to describe the methodological quality of studies describing the (criminological) real life effects of interventions, such as the introduction of risk based. It could be argued that even the lowest MSMS scales *by definition* can only start at the highest TRL level – beyond the scope of H2020 Innovation Actions.

[25] First and second line checks are defined in Schengen article 2, definition 13: "*'second line check' means a further check which may be carried out in a special location away from the location at which all persons are checked (first line);*"

authorized travel and is truthful about when asked about this. Ergo, the border guard has to deduce that information, which makes it de facto a type of profiling.

TRESSPASS attempted to build human values into the components and architecture by design. TRESSPASS delivered a modular and highly configurable set of components, and this allowed for detailed customization. This allows for border guards at design-time (i.e., already before runtime) to design a BCP that is as proportional to the actual situation as possible. More specifically, since new data sources (VTC, RTBA, TLTP, TLC/IAS, TCSS and ISS) were not used in all pilots, it was learnt that RBBM is possible with only the legacy information systems (PNR, API, VIS, SISII, EES, ETIAS, etc.)[26] as data sources[27]. The Traveller Registration Application (TRA) is a special case in this regard. In the current (rule based) situation there are several data sources that rely on the collection – based on consent – of information from travellers that could otherwise not be reliably collected[28]. In TRESSPASS we collected those kinds of data sources together and simplified it in the TRA. That means that, for this purpose[29], the TRA is a stand-in application, a proxy, that allowed TRESSPASS in the pilots to emulate such existing forms. And to do this in a generic and easy-to-realize form. This kind of functionality is therefore not new, nor is it mandatory for RBBM, nor is it a replacement for identity checks (see next section). The quality of this kind of traveller-provided data can be an issue, and it can be manipulated by the (mala fide) traveller. This can be accounted for by comparing it with other data sources, using indicators to look for signs of manipulation and properly weighing this information.

TRESSPASS had other tasks and supporting actions, besides the pilots, to explore ethics[30]. Still, the pilots gave some additional insights into the ethics of RBBM[31]. The opinion of the end users of RBBM, the border guards that participated in the pilots, is that from an ethical perspective, risk-based border control is an improvement over rule-based border control

[26] The existing legacy information systems were provided by simulator called the Legacy Systems Interface (LSI). No actual traveler data was used.

[27] TRESSPASS assumed high quality data sources for the legacy information systems. This may be realistically obtainable, but not always representative for the current situation. E.g., PNR is currently in practice notoriously sparsely populated.

[28] Examples of existing forms that collect information from the traveller before, during or directly after his travel are the passenger locator form which is now being used to fight COVID. Or the arrival card regarding immigration: https://en.wikipedia.org/wiki/Arrival_card. Or the declaration of liquid assets: https://www.belastingdienst.nl/wps/wcm/connect/bldcontenten/belastingdienst/individuals/abroad_and_customs/restricted_prohibited_import_export/10_000_or_more/declaration_liquid_assets/filing_a_declaration .

[29] In the TRESSPASS concept, the TRA also had two other (potential) purposes. First, to explore supporting guiding the traveler to the proper next phase of his travel. This guiding functionality can also be organized in the physical environment of the BCP through e.g. signage or staff. Second, the TRA was the means for a traveler to provide his social media references. In the future, that can also be organized as part of an identity management process, or as part of a booking process (via PNR).

[30] TRESSPASS T6.3 has explored the opinion in questionnaires regarding the use of particular technologies, and TRESSPASS sister project H2020 PERSONA has explored the acceptability and ethics of no gate border control solutions.

[31] The FHG simulator has the ability to produce results about ethical effects of a BCP-design. Due to time, budget and capacity constraints, these analyses were not performed in the pilots.

because it provides more accountability, and it seems to them a more proportional way of working. They do warn that a physical layout of a BCP where high-threat travellers are made to wait in view of other travellers, may lead to stigmatization. In the pilots, travellers that were classified as bona fide or not as bona fide, nor as mala fide (i.e., "neutral"), were sometimes allowed to pass through "unseen". Border guards expressed discomfort with this idea, reinforcing the idea that meaningful human control is not just a demand from external (ethical) experts, but also from the end users themselves.

Many more guidelines for decision makers wishing to develop risk-based border crossing points have been found through desk-research (i.e., not by the pilots) and haven been collected in TRESSPASS D9.9.

5.3.2 Effectiveness / Stopping power

The TRESSPASS concept dictates that effectiveness should be determined per type of threat. On top of that, TRESSPASS consortium decided that the threat of identity fraud must not be compromised at all – in fact should be strengthened where possible[32].

The stopping power of a BCP is obtained through two factors. The first is the psychological effect deterrence, which affects the influx of mala fide travellers (see D1.2, section 4.1.3.2). In the pilots, deterrence was obtained by reducing the predictability of a traveller's path through a BCP by randomly selecting travellers to receive more thorough checks than would be required based on the outcome of their respective screening alone. This is a very simple and very robust mechanism; however, it is very difficult to explore the efficacy of this effect, especially at TRL7. Another factor contributing to deterrence is hiding the indicators and weights that are used in the screenings. Deterrence was not explored in the pilots, because no real mala fide travellers were used.

The second factor contributing to stopping power is the detection of mala fide, i.e., unauthorized travellers that, despite deterrence, attempt to cross a border. That is the focus of the remainder of the subsections in this section.

5.3.2.1 RBBM requires and supports strong identity checks

It is currently mandatory to perform an identity check (e.g., Schengen article 8, 3a(i)), but the strength of current identity checks varies over member states and within them, over BCPs. In theory it might be possible to apply the risk-based paradigm also to the strength of identity checks. Although this may have benefits in a situation where the strength of identity checks is already varying, it might make it more difficult to combine the outcomes of screenings to the appropriate traveller. In addition, modern identity authentication processes and systems (including biometrics) can be highly accurate and fast, so it is not clear if there are significant benefits to obtain in terms of improved flowrate, improved efficiency or improved ethical impact by applying a risk-based paradigm to the identity checks. Therefore, TRESSPASS decided to avoid this complexity in their pilots, and not reduce the strength of identity checks.

Consequently, each of the pilots used a BCP design that was at least as effective in detecting and stopping identify fraud as their equivalent rule-based counterparts. This was done by directing all travellers first to existing, modern identity authentication kiosks, using regular

[32] TRESSPASS cooperated with H2020 D4FLY about mitigating identity fraud at borders. NCSRD and TNO are partners in both and have provided a risk-based identity management scenario to D4FLY.

systems for the authentication of their identity with their normal configurations, irrespective of the outcome of any screening. In addition, screening indicators were defined that could provide early signs of identity fraud, thereby still allowing border guards to allocate more time and resources if they wanted to. And the variety of TRESSPASS data sources such as PNR, TRA and the rest of the Legacy Information Systems allows for the cross-checking of the same traveller's data with the various data sources and sensors. This provides the ability to detect errors in the data proactively, which allows the BCP to take appropriate action: correct, look for causes of errors, ask the traveller for an explanation, etc.

5.3.2.2 Stopping other threats

The pilots also gave insights in the stopping power of risk-based BCPs versus that of their equivalent rule-based BCPs[33]. Besides detecting identity fraud, the three pilots focused also on mitigating irregular migration, smuggling, and terrorism.

The pilots showed that if the particular data sources worked properly[34] then the proper indicators regarding these threats "fired" when the respective scenarios (i.e., travellers) were offered. This means that the core of the TRESSPASS risk-based concept, the Data Fusion Analysis (DFA), the Dynamic Risk Assessment System (DRAS) – based on the TRESSPASS Risk Assessment Method (TRAM), the Distributed Message System (DMS) and the Command & Control (C2) all worked as desired. The effect was that the detection rate for mala fide travellers posing a threat, was higher in each pilot for the risk based BCP versus its equivalent rule based BCP. This was also recognized by the operational border guards that participated in the pilots. For example, border guards expressed trust in the detection power of the system for both mala fide and bona fide travellers. And they expressed appropriate hesitancy regarding the idea of letting neutral travellers walk through.

While the pilots used relatively low numbers of travellers, the simulations used larger numbers with different variations. They showed that the detection rate for a risk based BCP that is designed for increased effectiveness by increased detection rate, also provided that increased detection rate. This includes the effects of false negatives. This happened for every pilot, also where screenings and checks for multiple threats were combined.

5.3.3 Flow rate

The air pilot designed a BCP that besides detection rate (i.e., stopping power) was also optimized for flow rate. In order to do this, the air pilot also applied bona fide "trust" indicators[35].

---

[33] The FHG simulator has the ability to produce results about effectiveness of a BCP per threat. Due to time, budget and capacity constraints, these analyses were not performed in TRESSPASS.

[34] The contribution per subcomponent / data sources to the obtained effects follows in a separate subsection.

[35] It is controversial whether this is currently legally allowed. On the one hand, there is no explicit legal basis for relaxing border checks such as those described in Schengen article 8, 3a. On the other hand, there is no specification in that article about the duration of the interview, or the depth of questioning that must be applied, so perhaps these could be made dependent on risk and trust indicators. The legal aspect of this matter is discussed in more depth in the TRESSPASS deliverables D1.4 and D9.x.

The air pilot also varied with the accuracy of the indicators in order to learn how good the flow of travellers must be separated in order to obtain certain effects. It was learnt that with a reasonably accurate set of indicators, a reduced waiting time for all travellers can be obtained – both bona fide and mala fide travellers, while also improving the detection rate of mala fide travellers. This includes the effects of false positives.

The land and maritime pilots decided to optimize for the detection rate, and not for the flow rate. Consequently, they experienced reduced flow rate for several arrival rates of travellers. Even then, using the simulations, it turned to be possible to alter the capacity of specific travel phases in their risk-based BCPs such that optimal variants could be found that also improved the flow rate.

5.3.4 Efficiency

The air and maritime pilots also looked for optimizations in terms of efficiency. This was expressed in terms of the amount of parallel flows and corresponding service points with accompanying staff needed to be arranged in order to have queues of a certain maximum length while supporting the improved detection rates.

Comparing a risk with a rule-based system, the simulations indicate that significant reductions in services points and staff seem possible for higher flow rates with the risk-based BCP. At lower flow rates there was no gain in these pilots. This may be explained having to account for staff waiting idle when there are too few travellers to fully occupy service points.

5.3.5 Interactions between performance areas

Usually, these performance areas relate to each other like communicating vessels, or like a zero-sum game. Improvement in one area, takes a toll in another. It therefore seems counterintuitive to see how it is possible to see improvements in all areas simultaneously.

With risk-based border management, this becomes possible with the help of a refined separation of traveller flows. Using the combination of both risk and trust indicators, and therefore of bona fide, mala fide, and neutral travellers (and travellers for which a screening was not done / finished in time), it becomes possible to optimize for stopping power, flow rate and efficiency at the same time.

Comparing to the rule-based, the “paradox” of seeing improvements in all areas simultaneously can be explained with introducing another performance area: that of the intelligence or complexity of the information processing. The additional “cost” in the risk-based border control, is the additional processing of additional risk-based information such as updating and maintaining a risk assessment configuration system such as DRAS (risk/bona fide indicators and profiles using TRAM), the processing of pre-travel information, such as PNR, API, VIS, SISII, EES, etc., and where desired also additional information obtained through new types of data sources.

5.4 Contributions to RBBM per data source

It has been shown that pre-existing data sources (in TRESSPASS simulated by LSI and TRA) are sufficient to design a risk based BCP. However, TRESSPASS also explored additional, new types of data sources for risk-based BCPs which can optionally be added to a risk-based BCP, see Table 5-3. TRESSPASS developed a highly modular and configurable system of systems for risk based BCPs which facilitated this flexibility. The feasibility and interoperability of these new data sources were already described in section 5.2. Table 5-2 in section 5.3 described which data sources were used in each of the pilots.

TABLE 5-3 TRESSPASS DATA SOURCES AND THEIR (POTENTIAL) FUNCTIONS IN THE TRESSPASS CONCEPT FOR RBBM

Data Sources	For screening and/or for check	Function
Legacy Systems Interface (LSI)	For screening and check	Simulation of pre-existing large scale ICT systems used in the border domain: PNR, API, EES, VIS, SISII,ETIAS.
Traveller Registration Application (TRA)	For screening and check	(1) Pre-registration traveller forms. (Emulation of ETIAS) (2) Allows entry of social media references.
Security Personnel Application (SPA)	For screening (potentially) and check	(1) (Potential) Manual entry of values of pre-defined types of indicators per traveller to be used in screening. (2) Mobile C2 interface for displaying screening output per traveller during checks.
TRESSPASS Light Client (TLC)/International Alert System (IAS)	For screening and check	1) (Potential) Manual entry of values of pre-defined types of indicators per traveller from TCN BCPs to be used in screening. 2) C2 interface for displaying screening output per traveller during checks.
Web Intelligence (WI)	For screening and check	The extraction of information from online open data sources such as the public parts of a traveller’s social media accounts and clear and dark web.
Video Tracking Component (VTC) + Realtime Behaviour Analysis (RTBA)	For screening and check	The registration of information from cameras at the BCP that observe the physical movement of travellers.
Traveller and Luggage Tracking Sensor Platform (TLTP)	For screening and check	The registration and scanning of information from luggage trackers at the BCP, correlation of luggage with a traveller, cross-checking the number of carried luggage with the trip’s length and verify them against those declared and registered for detection of any inconsistencies.
Thermal Counter Spoofing Sensor (TCSS)	For screening and check	The detection of the presence of masks in front of the face of travellers, i.e., of a type of presentation attack.
Interview Support System (ISS)	For check only	The support of interviewers in a first- or second-line interview during a check of high-risk travellers.

In the following subsections, the contributions of each of those new data sources to RBBM are discussed as they were experienced through the pilots[36].

5.4.1 Security Personnel Application

The idea behind the Secure Personnel Application (SPA) as a data source is that professionally trained border guards currently observe travellers, and that in the foreseeable future technology will not come close to their flexibility and accuracy. The implication for RBBM is that there has to be a way for operational border guards to register their observations in a way that they are accessible in the fusion (DFA) and screening (DRAS) process. It must also be done in a way that safeguard ethical and legal use of this functionality, i.e., traceable to the individual border guard, and conforming pre-defined ethically and legally validated indicators. E.g., the SPA may not become a backdoor for biased low-quality information.

The Dutch pilot did not explore this type of functionality of the SPA because the challenges were not very technical, and this decision allowed the Dutch pilot end users to focus on other more technically challenging data sources.

Regarding the sea pilot, mobile border guards used, via SPA, several behavioral risk indicators which were previously checked and approved internally through the TRAM tool as legally/ethically acceptable. Additionally, the SPA in the sea pilot was used in combination with the TLTP component using the 'SPA-TLTP Scan' functionality to cross-check the travelers' luggage and handbags declared and registered against those actually carried and also against the traveler's trip length and/or his duration of stay.

5.4.2 TRESSPASS Light Client (TLC)/International Alert System (IAS)

The idea behind TLC/IAS is that border guards from Third Country BCPs can send risk information about travellers to specific EU BCPs and this information can be available to EU border operators before a traveller arrives at the BCP contributing to the travellers' risk assessment.

However, an harmonization mechanism in the configuration of such an international system is needed between Third Countries BCPs and EU since each EU BCP will need to configure and deliver its own risk indicators. TLC is an easy-to-use application, and it was tested successfully in the sea pilot by border guards who pretended to act as TCN BCP and appreciated a lot. Such an international alert system could possibly be extended to be applied not only between TC and EU but also as a notification and messaging system between EU BCPs.

---

[36] The TRESSPASS simulators can be used to explore the contribution to performance areas on a component basis. This can help to answer a question such as "what is de benefit of Web Intelligence to the flow rate of a maritime BCP?" Due to time, budget and capacity constraints, these analyses were not performed in TRESSPASS.

5.4.3 Web Intelligence

The idea behind Web Intelligence (WI) is that border guards might benefit from the publicly available data that travellers themselves published on the internet[37]. This works by automatically detecting patterns of (combinations of) keywords in open sources. It is conceivable that this kind of technology in the future will also be capable of automatically interpreting images and videos. In order to prevent misuse, WI does not allow border guards to freely browse the public parts of social media accounts. Obviously, the quality of this kind of information currently varies immensely, so it is not a priori clear what the value to RBBM could be. This quality could improve over time when social media themselves develop beyond what is available today, such as verified identities, fact checking, and the potential merging of social media with new types of (self-sovereign) identity systems.

WI contributed to indicators in each of the pilots. In these pilots, WI was only applied conditionally on the outcome of previous screenings and only for high-risk travellers. WI functioned properly in the pilots and contributed to the detection of mala fide travellers that would otherwise not be detected.

Concerning WI, significant operational and ethical questions remain. How easy is it for mala fide actors to evade or spoof WI? Should WI only be applied conditionally, or also unconditionally, i.e., for all travellers in a traveller flow? Should WI only contribute to risk indicators, or also to trust indicators? TRESSPASS “obtained” the travellers social media references through the TRA. In reality, the USA obtains them as part of the visa application process. Which is the best way, or should it be part of (future) identity management systems or of the booking process (e.g., part of PNR)?

5.4.4 Video Tracking Component and Realtime Behaviour Analysis

The idea behind the Video Tracking Component (VTC) and the Realtime Behaviour Analysis (RTBA) is that travellers may show physical behaviour at BCPs that is relevant for the screening and checking in a (risk based) BCP. Descriptions of such behaviour have been elicited in previous research[38] and during the requirements gathering phase of TRESSPASS. A selection of those behaviours (i.e., tracks of travellers) which are provided by modern intelligent camera’s, was implemented in the VTC and RTBA.

The performance of VTC needs to be much higher than for similar other applications[39]. It needs to be highly accurate in order to provide value for the screenings, and to do this for a large number of travellers – some of whom may even actively attempt to hide malicious behaviour, and it must be near real time because the information is immediately used in the screening. Consequently, in TRESSPASS the technical feasibility of this functionality proved to be difficult in the environment of a border crossing point. This has to be taken into account when assessing the usefulness of the VTC and RTBA in the pilots.

---

[37] USA has requested social media references for certain types of travelers as part of the visa application process.

[38] Methods to elicit the specifications of such behaviours are described, for example in Van Rest, J. H. C., Roelofs, M., van Nunen, A., & Don, S. B. (2014). Deviant behaviour-Socially accepted observation of deviant behaviour for security-extended summary. Den Haag: TNO.

[39] Similar other applications of intelligent camera’s are sometimes offline and human assisted, e.g. in forensic analysis.

VTC and RTBA contributed to indicators in the air and maritime pilots. The accuracy of the VTC was such that end users could see its potential but did not really experience it in the pilots.

The VTC and RTBA are both generic tools, and they can be used for a wide range of applications at or near border crossing points. It is to be expected that their quality will further improve, and their uptake will grow, making them potentially very relevant components for the future of RBBM.

5.4.5 Traveller and Luggage Tracking Sensor Platform

The idea behind the Traveller and Luggage Tracking Sensor Platform (TLTP) is that the way travellers handle their luggage can contain information that is relevant for border control. A different, but related challenge is that of lost luggage detection, which may also be mitigated with similar technology.

For the maritime pilot, indicators were defined based on the interaction between a traveller and his carried luggage and handbags. The TLTP worked as specified to detect any potential inconsistencies associated with the length of the trip as well as any inconsistencies between the number of the carried luggage and those declared or registered, mala fide travellers' scenarios were successfully detected in the screening, and the maritime border guards and the security operators appreciated it.

5.4.6 Thermal Counter Spoofing Sensor

The idea behind the Thermal Counter Spoofing Sensor (TCSS) is that presentation attacks such as masks are a strong indication of a mala fide traveller. Modern facial recognition systems are more accurate than professionally trained humans in comparing faces, and modern masks are practically undetectable by humans. So, presentation attacks can be directed at automated biometric systems, but also at human professionals. TRESSPASS did not use biometrics in the pilots, only "human" recognition in the land border pilot. TRESSPASS integrated the TCSS as part of the screening, while obviously, it might be better conceptualized as part of the identity check.

For the land pilot, indicators were defined based on the detection of masks. The TCSS worked as specified, mala fide travellers were successfully detected in the screening, and the land border guards appreciated it.

5.4.7 Interview Support System (ISS) - MMCAT

In RBBM, border checks can be more or less thorough. The idea is that checks for travellers that generate hits on mala fide profiles are more thorough than the checks for other travellers. The thoroughness of an interview for such a check can be altered by changing the duration of the interview. However, there may be better alternatives to make a check more thorough besides only making the interview last longer.

The idea behind the Interview Support System[40] (ISS) is to explore such alternatives for "thorough checks" by making the interviewing border guard operate on a high level of

[40] The Interview Support System (ISS) in TRESSPASS was initially called the 'Multi Modal Communication Analysis Tool (MMCAT)'. However, we use interchangeably the term ISS since the name MMCAT seemed insufficiently descriptive for what the tool actually does.

quality, by facilitating him / her in a second line interview-setting in observing and registering the behaviour of travellers while they are being interviewed – and in providing suggestions of possible explanations of this behaviour with relevant mental processes – such as emotions.

In the TRESSPASS concept, when the screening of a traveller generates a hit on a mala fide profile, then the traveller is directed to the same interview that currently also occurs in a rule-based setting. Depending on the configuration of the particular border crossing point, this can be either a first line interview at the front of a queue, or a second line interview in a separate room.

For that interview, border guards and customs officers currently use interviewing techniques to reliably determine information that is relevant for determining whether a traveller is authorized to cross the border. For this purpose, they typically look at the contents, the semantics, of a statement in relation to other information (e.g., travel origins and route and possessions), and also in relation to the way information is conveyed in the interview. This includes signs of potentially relevant mental processes such as emotions and cognitive load. Signs of these mental processes can be found in the contents of a statement, and also in the body language of a person. While it is undisputed that mental processes can lead to observable physical and physiological behaviour, it is much more difficult to reason the other way, i.e., to explain specific behaviour with (assumptions of) underlying mental processes.

For certain mental processes – such as emotions – this relationship is somewhat better understood than for others. TRESSPASS explored the feasibility and operational usefulness of technology that facilitates the border guard in a second line interview-setting in observing and registering the behaviour of travellers while they are being interviewed – and in providing suggestions of possible explanations in a small selection of relevant mental processes – including emotions[41]. This kind of technology potentially improves e.g., traceability and correctness of the outcomes of such interviews.

The idea of technical support in the interview setting is not new. In criminal investigations (which border checks are not), it has become a good practice to record witness statements to help improve the process of fact finding in criminal investigations. The idea to support the (live) analysis of such data for signs of relevant mental processes with technology is controversial. There is the risk that the technology is not accurate (e.g., is biased), is misused or that sensitive data is created that can be leaked or reused for illegal purposes. In TRESSPASS this kind of technology is not used in combination with automated decision making (so no automatic lie detection). Due to the novelty of this kind of technology, there exist virtually no established value-sensitive design patterns for technology in interview settings. The ISS pushes the conceptualization of what the benefits and risks of this kind of technology could be, and thereby stimulates the (public) debate about the why, how and what of improving interviews at border control.

The ISS was built according to the functional specifications elicited early in TRESSPASS. However, due to the novelty, it is not clear how good those specifications actually were. This has to be taken into account when assessing the usefulness of the ISS in the pilots.

[41] With the ISS, TRESSPASS builds on pre-existing research on this domain, such as has been conducted by our consortium members prior to TRESSPASS, and by external sources (e.g. The Deception Research Society).

The land and the maritime pilots used the ISS. These pilots showed that the ISS worked in the sense that border guards found the ISS useful in the interviews and liked the idea of such technological support in interviews – even in the somewhat rough instantiation of the TRESSPASS ISS. Border guards that experienced the ISS in both pilots provided many suggestions for improving the user interface further. Many of these suggestions were implemented in between the land and maritime pilots, and this was recognized by a higher appreciation in the maritime pilot of the ISS user interface. The maritime border pilot provided the suggestion that the ISS is more useful for novice border guards than for experienced border guards. This may be a reflection of the limited functionality of the ISS in terms of limited types of behaviour and limited types of mental constructs that it assesses. Perhaps a more functionally rich version of the ISS would also be more useful for more experienced border guards. In the meantime, the current version of the ISS might (also) be useful for training purposes.

All end users received training in the purpose and functionality of the ISS. However, after experiencing the ISS in the pilot, one end user expected a clearer decision (from the technology). This instance was used to explain the purpose again (i.e., as a support tool, not as an automated decision-making tool), and this “solved” this misunderstanding. Still, this example must serve as a warning against the potential for future misuse of ISS-type of technology.

5.4.8 The way forward per data source based on the outcome of the pilots

Based on the outcomes of the pilots, it is possible to give recommendations on the way ahead, see Table 5-4. The ethical appreciation of these components is given in D9.8, and the ethical and legal recommendations for border guards that want to proceed are described in D9.9. A big aspect of the way ahead is reserved for standardization, which is described in a dedicated white paper (which is an informal project deliverable). A more holistic perspective on the way ahead is provided in D10.6, the TRESSPASS Roadmap.

TABLE 5-4 RECOMMENDATIONS ON THE WAY AHEAD PER TRESSPASS DATA SOURCE.

Data source	Clarity of required functionality	Technical and operational feasibility	Validation of usefulness by pilots			Recommendation for the way ahead
			Worked as specified	End users experienced and appreciated it	Simulation showed usefulness	
LSI	High	High	Yes	Yes	Yes	Progress to TRL8
TRA	High	High	Yes	Yes	Yes	Progress to TRL8
SPA	High	High	Yes	Yes	N/A	Improve functional specs for more extensive tests at TRL7. SPA's functionality can be extended to apply the border checks in mobile mode instead of in a kiosk.
WI	Medium	Medium	Yes	Yes	N/A	Improve functional specs and quality of test data for more extensive tests at TRL7
TLC/IAS	High	High	Yes	Yes	N/A	Progress to TRL8 in BCPs other than the TCN BCPs (in EU BCPs)
VTC+RTBA	Medium	Low	Yes	Not convincingly	N/A	Improve functional specs, technical and operational feasibility (all through standardization)
TLTP	High	High	Yes	Yes	N/A	Progress to TRL8
TCSS	High	High	Yes	Yes	N/A	Progress to TRL8
ISS (MMCAT)	Low	Medium	Yes	Yes	N/A	Take lessons for next design iteration and testing at TRL6

5.5 Compliance with the principles of Legal/Ethical Framework

Using as reference the conclusions of D9.5 a comprehensive ethics framework was developed in TRESSPASS T9.1 to ensure that the demonstration activities were conducted in compliance with ethical principles and norms, including the requirements set by the EC as outcomes of the Ethics Review, the Ethics Check and Reviews. The relevant aspects for pilots included:

1. ensuring that all volunteers participating into the pilots are extensively informed and participate on the basis on their freely expressed consent.
2. a strategy is elaborated to deal with incidental findings if they occur.
3. safety rules are in place.
4. data are collected and handled in conformity with the GDPR.
5. the ethical impact of the technology tested is minimized.

D9.2-9.5 describe how these requirements were met in detail. The framework developed in T9.1 was fully implemented in the pilots: to minimize the ethical and data protection impact, whenever possible synthetic data were fabricated and used for the pilots' purposes. When it was necessary to rely on real data, these were collected from volunteers after informed consent. Extensive documentation was prepared for ensuring that data collection and protection responsibility were clearly defined, that the safety of participants was guaranteed during the tests and that a proper strategy to deal with incidental findings was in place.
The Data Protection Officers, legal departments and Ethics Committee of the respective organizations were involved into the pilots planning by preparing the relevant legal/ethical documentation.

Although implementation of the framework was successful, it took considerable effort from the side of end-users and the ethics team to prepare and plan the pilots in accordance with it. The necessary documentation was sometimes perceived as an excessive burden. Also, the need to opt for alternative solutions to the use of real data from the public or re-use pre-existing data owned by end-users needed in some cases extensive explanations and negotiations to be accepted by end-users and technology developers, as it was perceived as an unnecessary limitation to the technical and operational possibilities. It was possible, however, at the end of such process to achieve a consensus on the more ethically compliant solutions. The process which led to this result had an additional positive effect, namely increasing awareness of end-users and technical partners regarding ethical aspects of technology development and deployment.

REFERENCES

- Colledanchise, M. and Ogren, P., "How behavior trees modularize robustness and safety in hybrid systems," ¨in [Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on], 1482–1488 (Sept 2014).
- Perez, D., Nicolau, M., O'Neill, M., and Brabazon, A., [Applications of Evolutionary Computation: EvoApplications 2011: EvoCOMPLEX, EvoGAMES, EvoIASP, EvoINTELLIGENCE, EvoNUM, and EvoSTOC, Torino, Italy, April 27-29, 2011, Proceedings, Part I], ch. Evolving Behaviour Trees for the Mario AI Competition Using Grammatical Evolution, 123–132, Springer Berlin Heidelberg, Berlin, Heidelberg (2011).
- Lim, C.-U., Baumgarten, R., and Colton, S., [Applications of Evolutionary Computation: EvoApplicatons 2010: EvoCOMPLEX, EvoGAMES, EvoIASP, EvoINTELLIGENCE, EvoNUM, and EvoSTOC, Istanbul,Turkey, April 7-9, 2010, Proceedings, Part I], ch. Evolving Behaviour Trees for the Commercial Game DEFCON, 100–110, Springer Berlin Heidelberg, Berlin, Heidelberg (2010).
- Isla, D., "Handling Complexity in the Halo 2 AI," (2005).
- Ritzmann, S., Hagemann, V., & Kluge, A. (2014). The Training Evaluation Inventory (TEI) - Evaluation of Training Design and Measurement of Training Outcomes for Predicting Training Success. *Vocations and Learning*, *7*(1), 41–73. https://doi.org/10.1007/s12186-013-9106-4

ANNEX A: VTC EVALUATION RESULTS FROM THE GREEK PILOT

The table below shows the quantitative evaluation results from the questionnaires filled by the Border Operators, for VTC. X, O, and # denote the answers of the three officers. The Border Guard Officers had to judge the quality of the following 10 performance indicators:

Indicator/Grade 1-5	Completely disagree	Somewhat agree	Neither agree nor disagree	Somewhat disagree	Completely agree	Mean Score
Understandability			O	X	#	4.0
Learnability			O	X#		3.7
Usability			O	X	#	4.0
Interoperability			OX	#		3.3
Integration			O	X	#	4.0
Usefulness			OX#			3.0
Information			OX	#		3.3
Installation			O	X#		3.7
Performance			OX		#	3.7
Legal/Ethical			O	X#		3.7
TOTAL AVERAGE						**3.64**

The table below summarizes the duration of the Re-Identification errors and the duration of the videos.

From the table it follows that the Re-Identification Error Rate = 11:32 / 4:06:18 = 692 seconds / 14778 seconds = 0.0468. This high Re-Identification Error Rate is possible due to low VTC resolution in the data provided to VicarVision.

Traveller ID used in Mockup Data	Duration Re-Identification Errors	Duration of video
TRID-1	0:07	14:28
TRID-2	-	7:07
TRID-3	0:50	7:52
TRID-5	-	7:08
TRID-6	-	32:11
TRID-7	1:57	26:54
TRID-8	0:49	18:05
TRID-9	-	8:13
TRID-10	-	5:57

TRID-11	-	8:32
TRID-12	0:13	3:25
TRID-13	-	5:08
TRID-14	5:49	19:18
TRID-15	3:00	12:39
TRID-16	-	3:51
TRID-19	0:09	6:15
TRID-20	-	19:35
TRID-21	-	4:20
TRID-22	-	4:02
TRID-23	-	9:40
TRID-24	-	1:54
TRID-29	0:38	6:24
TRID-31	1:00	13:20
SUM	11:32	4:06:18

Annex B: MMCAT Assessment: Q&As and Scorings

MMCAT – POLISH PILOT Q&As

The answers by the first Polish operator were indicated by **+** and answers by the second polish operator by **o**.

Closed questions	**Disagree**				**Agree**	**Score**
	1	**2**	**3**	**4**	**5**	**1-5**
MMCAT is a helpful tool to support me taking an interview				+ o		4.0
The graphical user interface (GUI) is good			+	o		3.5
The GUI should provide more information				o	+	4.5
The GUI should provide less information	+ o					1.0
The meaning of the features/indicators is clear			+ o			3.0
The features are helpful.				+ o		4.0
The training material was sufficient to use the tool.			+		o	4.0
TOTAL AVERAGE SCORE						**3.43**

Open questions	**Answer**
What were your expectations from MMCAT tool?	(+) I would expect more detailed info concerning personal data. (o) Fully automated system
Have these expectations been met?	(+) Not exactly. More detailed info should be put in the system. (o) Not completely
What are the main strengths of MMCAT?	(+) The face reaction is analysed in the way we (BG officer) cannot see/notice (o) The system seems to be running smoothly
What are the main weaknesses of MMCAT?	(+) Sometimes it was difficult to rate the persons reaction. (o) No automation, average intuitiveness

Do you have suggestions related to the graphical user interface (GUI)?	(+) It was ok (o) No
Do you have suggestions related to the computed features/indicators?	(+) ok (o) No
Do you have suggestions for the training material?	(+) ok (o) No
Do you have other suggestions?	(+) I would appreciate the system to be more detailed in its evaluation (o) No

MMCAT – GREEK PILOT Q&As

The questionnaire with the following set of questions were completed at the end of the Greek Pilot by the Border Guard operators interviewing volunteers playing the role of mala fide traveller scenarios with the support of MMCAT in the 2nd line checks.

Closed questions	Disagree				Agree	Score
	1	2	3	4	5	1-5
MMCAT is a helpful tool to support me taking an interview				1	2	4.7
The graphical user interface (GUI) is good			2		1	3.7
The GUI should provide more information				1	2	4.7
The GUI should provide less information		2			1	3.0
The meaning of the features/indicators is clear			2	1		3.3
The features are helpful.			2		1	3.7
The training material was sufficient to use the tool.					3	5.0
TOTAL AVERAGE SCORE						**4,14**

Comparing the above Greek Pilot's total average score (4,14) with the corresponding Polish Pilot's (3,43) in the previous section of this Annex, It is obvious that the feedback provided by the Polish Pilot end users contributed significantly to the improvements and enhancements of MMCAT made in between the two pilots.

ANNEX C: ICROWD INPUTS AND OUTPUTS

iCrowd Inputs

Name	Type	Description
Timing parameters		
Enrollment service time	Time Distribution	How long it takes for a passenger to complete the enrollment in the enrollment kiosk. **Dutch pilot only.**
ABC Gate service time	Time Distribution	How long it takes for a passenger to scan their passport in the ABC gate and see if the gate opens or not. This must include the time it takes for the TRESSPASS system to respond, as this is currently being simulated instead of using the real systems. **Dutch pilot only.**
Luggage Drop Off / Initial enrollment time	Time Distribution	How long it takes for a passenger to complete the initial enrollment at the entrance (where the passenger will also leave their luggage). **Greek pilot only.**
Second-line interview time	Time Distribution	The duration of a second-line interview. **Greek pilot only.**
Luggage Pickup	Time Distribution	How long a passenger waits for their luggage. **Greek pilot arrivals only.**
Manual Desk service time	Time Distribution per type	How long it takes for a desk BG to find a passenger using C2, read over their information, and make a decision. (send to ABC gate or second-line interview for **dutch pilot**, simple check-in time for **greek pilot**). This must include any delays caused by accessing the TRESSPASS system through C2.
Mobile BG with SPA service time	Time Distribution per type	How long it takes for a mobile BG to use the SPA to find a passenger, read over their information, and make a decision. This applies to BGs after the ABC gate (**dutch pilot**) and wandering BGs (**greek pilot**).
Mobile BG without SPA (redirection) service time	Time Distribution per type	How long it takes for a mobile BG to just redirect a passenger without performing a check (used to tell the Malafide passenger in the ABC gate to go back to the manual desk without performing a check using the SPA). **Dutch pilot only.**
Xray scanner	Time Distribution per	How long it takes for a passenger to leave

entrance delay	type	their carry-on luggages on the Xray Scanner's belt. **Greek pilot only.**
Xray scan delay	Time Distribution per type	How long it takes for a passenger to go through the Xray Scan. **Greek pilot only.**
Xray scanner exit time delay	Time Distribution per type	How long it takes for a passenger to gather their carry-on luggages from the XrayScanner's belt. **Greek pilot only.**
After-xray-scanner BG service time	Time Distribution per type	How long it takes for the BG after the Xray Scan to use the SPA to find a passenger, read over their information, and decide whether the passenger can proceed to Passport Control (Bonafide) or the BG needs to call a different BG to take the passenger to a second-line interview (Malafide). **Greek pilot only.**
Passport control service time	Time Distribution per type	How long it takes for the BG after the Xray Scan to use the SPA to find a passenger, read over their information, and decide whether the passenger can proceed to Passport Control (Bonafide) or the BG needs to call a different BG to take the passenger to a second-line interview (Malafide). **Greek pilot only.**
Probabilities		
Initial Assessments	Probability per type	Probabilities for passengers to be labelled as Bonafide, Malafide, or Neutral in the Enrollment Kiosk (**dutch pilot**) or during the luggage drop off at the entrance of the BCP (**greek departures pilot**), or during initial Passport Control (**greek arrivals pilot)**.
Wandering BG probability to stop a passenger	Probability per type	Probability that a wandering BG will randomly select a passenger of some type (though the BG has no knowledge of the assessment beforehand). This probability will be evaluated for each wandering BG separately, so e.g. a 0.5 probability would mean that each BG has 50% chance to stop a passenger, so if the passenger is not stopped by one BG, they also have a 50% chance of being stopped by the next one they meet. A passenger cannot be stopped more than once. **Greek pilot only.**
Mistake probability	Probability per type	Probability that a passenger will make a mistake regarding ABC Gates and Manual Desks. **Dutch pilot only.**
Random selection	Probability per type	Probability that a passenger will be

for Neutral and Bonafide passengers		instructed by the Enrollment kiosk to go the manual desk. This is the random selection process for Neutral and Bonafide passengers, Malafide and Unknowns are always instructed to go the Manual Desk so their probability should be 1. **Dutch pilot only.**
Percentage of ABC BGs who make checks using SPA	Probability	Percentage of mobile BGs who, when approaching a Malafide passenger in the ABC gate, will check them and may change them to Bonafide and let them through. **Dutch pilot only.**
ABC BG sample rate	Probability	Probability that an ABC BG will randomly choose to check a passenger whose gate opened normally. **Dutch pilot only.**
Assessment changes		
Assessment change probabilities in Manual Desk	Assessment changes	Probabilities that a passenger's assessment will change by a Manual Desk BG. **Dutch pilot only.**
Assessment change probability in Xray Scanner	Assessment changes	Probability that a passenger's assessment will change during the Xray Scan. **Greek pilot only.**
Assessment change probability by wandering BG	Assessment changes	Probabillity that a wandering BG will eventually change the assessment of a stopped passenger. **Greek pilot only.**
Assessment change probability during Passport Control	Assessment changes	Probability that a passenger's assessment will change by a Passport Control BG. **Greek pilot departures only.**
Assessment change probability by ABC BG	Assessment changes	Probability that a passenger's assessment will change by an ABC BG. **Dutch pilot only.**
Assessment change probability after second-line interview	Probability	Probability that a passenger's assessment will change after the second-line interview. Since the current assessment can only be Malafide, 3 out of the 4 rows of an assessment change matrix would be blank. Assuming that the passenger can only change to Bonafide, then the row will have a probability in the Bonafide column, "auto" in the Malafide column, and blank in the Neutral and Unknown columns. So this is just a simple Probability, no need for an assessment change matrix. Definitely applicable in **Greek pilot**, for **dutch pilot** this will probably be 0.

Miscellaneous		
Wandering BG call radius	Meters	The radius in which a wandering BG will look for passengers to randomly stop and check. **Greek pilot only.**
Percentage of passengers with at least one luggage	Probability per type	**Greek pilot only.** For the departures scenario, only one number is required as this is decided before passengers get their initial assessment. In the arrivals scenario, we can have different probabilities for each time, since this is determined after the passenger has gone through initial Passport Control.
Percentage of passengers with at least one carry-on luggage	Probability	**Greek pilot only.**
Standard positions of wandering border guards	List of positions	BGs who wander around the BCP and randomly stop passengers. **Greek pilot only.**
Standard positions of after-xray border guards	List of positions	BGs who scan a passenger's passport after they walk out of the Xray Scanner. **Greek pilot departures only.**
Standard positions of second-line guides	List of positions	BGs who take a Malafide passenger from the Xray Scanner and guide them to a second-line interview room. **Greek pilot only.**
Arrival rate of passengers	Rate (passengers/hour)	The rate at which the passengers are generated near the entrance of the BCP. **Dutch pilot and greek pilot.**
Total number of passengers	Count	The total number of passengers that will be created throughout the simulation. **Dutch pilot and greek pilot.**

Input Types

Type	**Description**
Time Distribution	Pair of 2 values that define a Gaussian distribution: {Mean, Standard deviation}.**Example: {50.0, 10.0}**
Time Distribution per type	Array of 4 pairs of 2 values that define a Gaussian distribution. This can also be a single pair if it's the same for all types. **Example: [Bonafide: {50.0, 10.0}, Malafide: {..., ...}, Neutral: {..., ...}, Unknown: {..., ...}]**
Probability	Single value between 0 and 1 indicating the probability for something to happen.**Example: 0.43 (=43%)**
Probability per type	Array of 4 probabilities: one for each type of passenger at the time of the event (Bonafide, Malafide, Neutral, Unknown). A value can be 0.0. A value can be blank in case the event should never occur to a specific type of passenger. The probabilities are independent, so they

	don't have to sum up to 1. **Example: [Bonafide: 0.2, Malafide: 0.3, Neutral: 0.8, Unknown: -]** *For mistakes, this would mean that Bonafide passengers would make mistakes in 20% of cases, Malafide in 30% of cases, Neutrals in 80% of cases, and Unknowns will never have to make such a decision. The user can adjust all probabilities on-the-fly except the Unknown.*
Probability per type (sum to 1)	Same as "Probability per type", except the probabilities need to sum up to 1, so one of the values needs to be "auto". The "auto" value will be automatically calculated based on the rest of the values to ensure a sum of 1. **Example: [Bonafide: 0.2, Malafide: 0.3, Neutral: auto, Unknown: -]** *For enrollment, this would mean that 20% of passengers will be Bonafide, 30% will be Malafide, and the rest will be Neutral. The user can adjust the Bonafide and Malafide probabilities on-the-fly, but not the Unknown.*
Assessment changes	**4x4 Matrix of probabilities**: a number for each possible combination of types [a,b] that indicates the probability that a passenger of type "a" will change to type "b". Each row corresponds to the current type and each column corresponds to the next type. A row might be blank, if the event should never occur to that specific type of passenger. Each row should have at least one value of "Auto", which would be automatically calculated by the simulator when the user adjusts the rest of them so that the row sum is equal to 1. A probability in a row can be 0.0 which would mean that by default it is deactivated but can be adjusted by the user on-the-fly. A probability can be blank (-) which would mean that a passenger of the type of the row can never transition to the type of the column and cannot be adjusted by the user on-the-fly. The rows are independent from each other.

iCrowd Outputs

Name	**Type**	**Description**
	Global	
Total served passengers	Count	Reported globally and for each assessment (Unknown, Bonafide, Neutral, Malafide)
Total time spent in BCP	Time	Reported for each assessment (Unknown, Bonafide, Neutral, Malafide)
Average time spent in BCP	Time per passenger	Reported for each assessment (Unknown, Bonafide, Neutral, Malafide)
Served passengers (exited) over time	Graph: N over time	Shows how many passengers exited the BCP between measurements. Measurements are taken in intervals of 10 seconds (can be adjusted)
Served passengers (exited) cumulative	Graph: N over time	Shows the total of exited passengers over time. Measurements are taken in intervals of

		10 seconds (can be adjusted)
Number of passengers in BCP	Count	Makes sense only in real time
Passengers in BCP over time	Graph: N over time	Measurements are taken in intervals of 10 seconds (can be adjusted)
Passengers entered over time	Graph: N over time	Measurements are taken in intervals of 10 seconds (can be adjusted). Min/Avg/Max are also displayed.
Passengers in type of checkpoint over time	Graph: N over time	Reported separately for each type of checkpoint such as Enrollement kiosks, manual desks, xray scanners, etc. Avg/Max are also displayed.
Total utilization time for each type of checkpoint	Time	Reported separately for each type of checkpoint
Average utilization time for each type of checkpoint	Time per passenger	Reported separately for each type of checkpoint
Total served passengers for each type of checkpoint	Count	Reported separately for each type of checkpoint
Number of travellers that were stopped by wandering BGs.	Count	Reported separately for Bonafide and Neutral/Malafide passengers
Number of travellers whose assessments were changed by wandering BGs.	Count	Reported separately for each possible change (Bonafide->Malafide, Neutral->Bonafide, etc.)
Number of travellers whose assessments were changed by Manual Desk BGs	Count	Reported separately for each possible change (Bonafide->Malafide, Neutral->Bonafide, etc.)
Number of travellers whose assessments were changed by a BG after the ABC gate	Count	Reported separately for Malafide->Bonafide for whom the BG made the check instead of sending them to a Manual Desk, and for Bonafide->Malafide who were randomly selected.
Number of Bonafide travellers who went to Manual Desk instead of ABC Gate	Count	
Number of Malafide travellers who went to ABC Gate instead of Manual Desk	Count	
Number of travellers whose assessments were changed during the Xray Scan	Count	

Per checkpoint		
Number of agents in queue	Count	Makes sense only in real time
Total served passengers	Count	Reported in total and separately for each assessment (how many passengers exited the checkpoint as Bonafide, Malafide, etc)
Total utilization time	Time	
Average utilization time	Time per passenger	
Queue times	Graph: Time over passengers	Shows how long each passenger had to wait in queue. The graph is updated every time a passenger finishes waiting in queue and starts being serviced. Min/Avg/Max are also displayed.
Queue size over time	Graph: N over time	Avg/Max are also displayed
Service times	Graph: Time over passengers	Shows how long each passenger spent being serviced. The graph is updated every time a passenger finished their service. Min/Avg/Max are also displayed.
Served passengers over time	Graph: N over time	Shows how many passengers were serviced between measurements. Measurements are taken in intervals of 10 seconds (can be adjusted). Max is also displayed.
Served passengers cumulative	Graph: N over time	Shows the total of served passengers over time. Measurements are taken in intervals of 10 seconds (can be adjusted).
Per Border Guard		
Number of served passengers	Count	
Queue sizes over time	Graph: N over time	Shows how many passengers are waiting to be serviced by this BG over time. Avg/Max are also displayed.
Service times	Graph: Time over passengers	Shows how long the BG spent servicing each passenger. The graph is updated every time the BG finishes servicing a passenger. Min/Avg/Max are also displayed.
Per Passenger		
Initial assessment	DRAS assessment	
Final assessment	DRAS assessment	
Stopped by a wandering BG	Boolean	

Serviced by a Manual Desk	Boolean	
Serviced by a BG after the ABC gate	Boolean	
Did second-line interview	Boolean	
Went through passport control	Boolean	
Went through Xray Scanner	Boolean	
Made mistake regarding ABC-gate and Manual Desk	Boolean	
Enrollment Random Selection	Enum	Takes values of "(ABC|MD) as (Neutral|Bonafide)"
Total walking time	Time	Time spent walking from checkpoint to checkpoint. Does not include walking in queues.
Total time spent in queues	Time	Time spent walking and waiting in queues. Includes waiting for a BG to arrive after checkpoints such as Xray Scanner or ABC gate.
Total time spent in service	Time	Time spent being serviced either by a checkpoint (enrollement kiosk, ABC gate) or by BGs (wandering BG, after Xray Scanner, after ABC Gate, in Manual Desk, etc)

ANNEX D: ICROWD SIMULATION RESULTS

The following results were extracted from iCrowd after simulating a Greek departures scenario.

Inputs

Parameter	**Values**	
Timing parameters		
Luggage Drop Off / Initial enrollment time	Mean: 10 Standard deviation: 5	
Second-line interview time	Mean: 450 Standard deviation: 150	
Manual Desk service time	Bonafide	Mean: 85 sec. Standard deviation: 5 sec
	Neutral	*Mean: 95 sec.* *Standard deviation: 5 sec*
	Malafide	*Mean: 105 sec.* *Standard deviation: 5 sec*
	Unknown	*Mean: 105 sec.* *Standard deviation: 5 sec*
Mobile BG with SPA service time	Bonafide	Mean: 12 sec. Standard deviation: 2 sec
	Neutral	*Mean: 17 sec.* *Standard deviation: 2 sec*
	Malafide	*Mean: 22sec.* *Standard deviation: 2 sec*
	Unknown	*Mean: 22sec.* *Standard deviation: 2 sec*
Xray scanner entrance delay	Bonafide	Mean: 9 sec. Standard deviation: 1 sec
	Neutral	Mean: 9 sec. Standard deviation: 1 sec
	Malafide	Mean: 9 sec. Standard deviation: 1 sec
	Unknown	Mean: 9 sec. Standard deviation: 1 sec
Xray scan delay	Bonafide	Mean: 18 sec. Standard deviation: 3 sec
	Neutral	Mean: 18 sec. Standard deviation: 3 sec

	Malafide	*Mean: 21 sec.* *Standard deviation: 3 sec*
	Unknown	*Mean: 21 sec.* *Standard deviation: 3 sec*
Xray scanner exit time delay	Bonafide	Mean: 8 sec. Standard deviation: 1 sec
	Neutral	Mean: 8 sec. Standard deviation: 1 sec
	Malafide	Mean: 8 sec. Standard deviation: 1 sec
	Unknown	Mean: 8 sec. Standard deviation: 1 sec
After-xray-scanner BG service time	Bonafide	Mean: 5 sec. Standard deviation: 2 sec
	Neutral	Mean: 10 sec. Standard deviation: 5 sec
	Malafide	Mean: 10 sec. Standard deviation: 5 sec
	Unknown	Mean: 10 sec. Standard deviation: 5 sec
Passport control service time	Bonafide	Mean: 6 sec. Standard deviation: 1 sec
	Neutral	Mean: 11 sec. Standard deviation: 1 sec
	Malafide	Mean: 17 sec. Standard deviation: 2 sec
	Unknown	Mean: 17 sec. Standard deviation: 2 sec
Assessments		
Initial Assessments	Bonafide	20%
	Neutral	75%
	Malafide	5%
	Unknown	auto
Wandering BG probability to stop a passenger	Bonafide	2%
	Neutral	7.5%
	Malafide	0.5%
	Unknown	auto

Assessment change probability in Xray Scanner:

	Bonafide	**Neutral**	**Malafide**	**Unknown**
Bonafide	Auto	-	0.1	-
Neutral	-	Auto	0.1	-
Malafide	0.0	0.0	Auto	-

	Unknown	0.01	0.01	0.01	Auto
Assessment change probability by wandering BG		**Bonafide**	**Neutral**	**Malafide**	**Unknown**
	Bonafide	Auto	0.3	0.007	-
	Neutral	0.1	Auto	0.007	-
	Malafide	0.1	0.2	Auto	-
	Unknown	0.1	0.3	0.007	Auto
Assessment change probability during Passport Control		**Bonafide**	**Neutral**	**Malafide**	**Unknown**
	Bonafide	Auto	0.2	0.007	-
	Neutral	0.2	Auto	0.007	-
	Malafide	-	-	-	-
	Unknown	-	-	-	-
Assessment change probability after second-line interview	0.95				
Miscellaneous					
Wandering BG call radius	10.0				
Percentage of passengers with at least one luggage	0.99				
Percentage of passengers with at least one carry-on luggage	0.99				
Arrival rate of passengers	150 passengers / hour				
Total number of passengers	1000				

Outputs

Generic settings

Setting	**Times applied**
Initial Unknown assessments	0
Initial Malafide assessments	53
Initial Neutral assessments	758
Initial Bonafide assessments	189
Xray Scanner Bonafide -> Malafide	4
Xray Scanner Neutral -> Malafide	6
Wandering MBG Bonafide -> Malafide	1
Wandering MBG Bonafide -> Neutral	8
Wandering MBG Neutral -> Bonafide	16
Wandering MBG Neutral -> Malafide	2
Passport control Bonafide -> Malafide	1

Passport control Bonafide -> Neutral	47
Passport control Neutral -> Bonafide	148
Passport control Neutral -> Malafide	4
Second-Line Interview Malafide -> Bonafide	68
Total served passengers	1000

Served passengers

Passenger type	Count	Total time	Average time
Unknown	0	0	n/a
Malafide	3	498.318	166.106
Neutral	637	123787.0	194.328
Bonafide	360	71298.2	198.051

Passengers

ID	Name	Exit time	Initial DRAS	Final DRAS	Total walk time	Total queue time	Total service time
33	passenger_0	300.517	Neutral	Neutral	160.983	146.667	137.667
184	passenger_3	388.683	Neutral	Bonafide	176.633	534.998	134.25
182	passenger_1	392.517	Neutral	Bonafide	159.383	158.167	174.683
183	passenger_2	399.517	Neutral	Neutral	216.583	108.331	113.65
187	passenger_4	407.017	Neutral	Bonafide	167.733	209.998	140.633
191	passenger_7	460.35	Neutral	Neutral	164.533	983.307	126.35
189	passenger_5	491.517	Neutral	Bonafide	200.633	1.25	147.233
190	passenger_6	523.183	Neutral	Neutral	176.817	50	130.617
193	passenger_8	526.017	Bonafide	Neutral	194.15	143.332	137.867
196	passenger_10	553.35	Neutral	Neutral	189.3	120.001	122.467

Border Guards

ID	Name	Type	Using SPA	Number of served passengers	Total service time	Average service time
34	mobile_bg_1	Wandering BG	1	7	112.565	160.807
35	mobile_bg_2	Wandering BG	1	42	701.9	167.119
36	mobile_bg_3	Wandering BG	1	36	575.098	159.749

37	mobile_bg_4	Wandering BG	1	41	684.684	166.996
38	mobile_bg_5	Wandering BG	1	60	1006.28	167.714
39	mobile_bg_6	Wandering BG	1	33	559.107	169.426
40	xray_mbg_1	After-Xray BG	1	158	1385.16	876.686
41	xray_mbg_2	After-Xray BG	1	350	3154.39	901.255
42	xray_mbg_3	After-Xray BG	1	0	0	-
43	xray_mbg_4	After-Xray BG	1	491	4466.48	9.0967
44	xray_mbg_5	After-Xray BG	1	1	5.0498	5.0498
45	xray_mbg_6	After-Xray BG	1	0	0	-
46	secondline_mbg_1	Second-line interview guide	0	9	158.569	176.188
47	secondline_mbg_2	Second-line interview guide	0	62	1223.12	197.278

Checkpoints by type

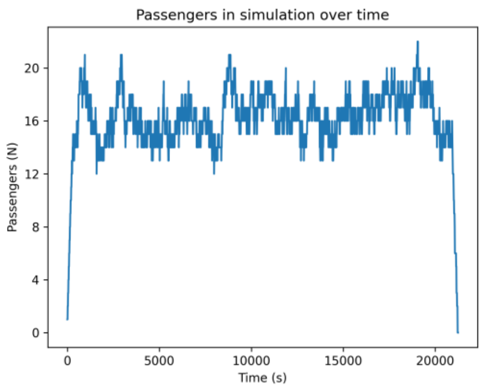
Type	Total utilization time	Total served passengers	Average utilization time
LuggageDropOffPoint	9955.62	1000	4977.81
ManualDesk	93704.7	1066	2928.27
PassportControlPoint	18094.5	1002	9047.26
SecondLineRoom	33614.3	71	8403.58
XrayScanner	9175.5	1008	1835.1

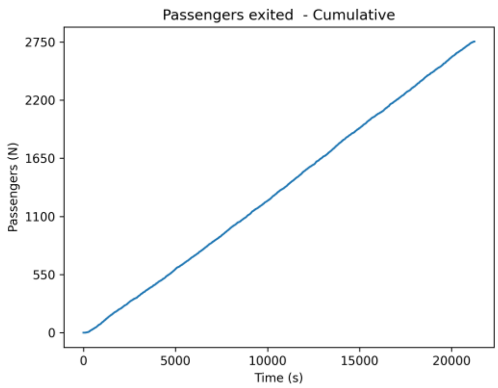
Checkpoints

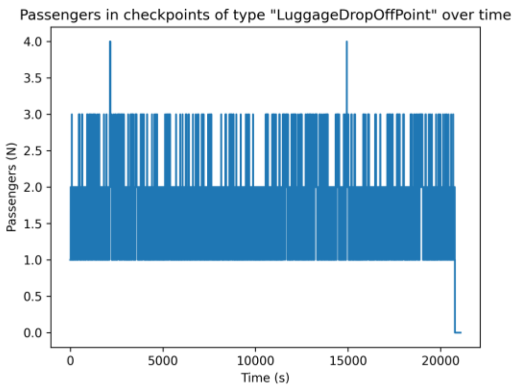
ID	Type	Total served passengers	Total utilization time	Average utilization time	Bonafides	Neutrals	Malafides	Unknowns
167	ManualDesk	245	22636.7	923.946	53	175	17	0
165	ManualDesk	195	17140.6	879.007	35	151	9	0
163	ManualDesk	176	15541.7	883.053	38	129	9	0
161	ManualDesk	1	110.016	110.016	0	0	1	0
141	ManualDesk	22	1770.26	804.665	4	17	1	0
139	ManualDesk	192	16184.1	842.924	32	150	10	0
137	ManualDesk	235	20321.2	864.734	35	192	8	0
135	XrayScanner	1	1.8999	1.8999	0	1	0	0
134	XrayScanner	495	4663.47	942.115	102	365	28	0

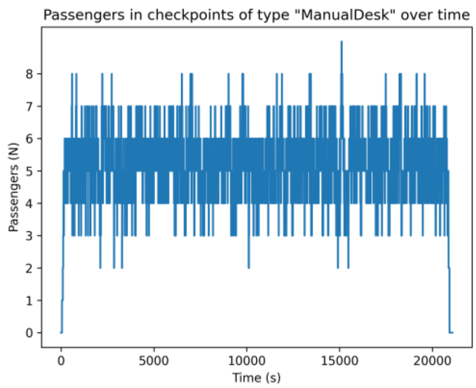
133	XrayScanner	352	2969.72	8.4367	59	268	25	0
132	XrayScanner	160	1540.41	962.755	32	115	13	0
131	SecondLineRoom	10	4746.32	474.632	10	0	0	0
130	SecondLineRoom	12	5460.03	455.002	12	0	0	0
129	SecondLineRoom	25	11268.6	450.745	22	0	3	0
128	SecondLineRoom	24	12139.3	505.806	24	0	0	0
127	LuggageDropOff	998	9935.66	995.557	189	756	53	0
126	LuggageDropOff	2	199.668	9.9834	0	2	0	0
125	PassportControl	612	11566.6	188.996	200	410	2	0
124	PassportControl	390	6527.94	167.383	160	227	3	0

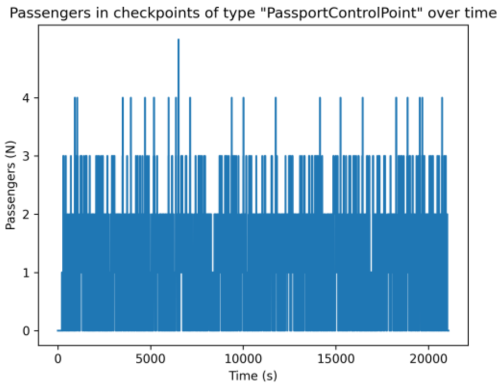
Graphs

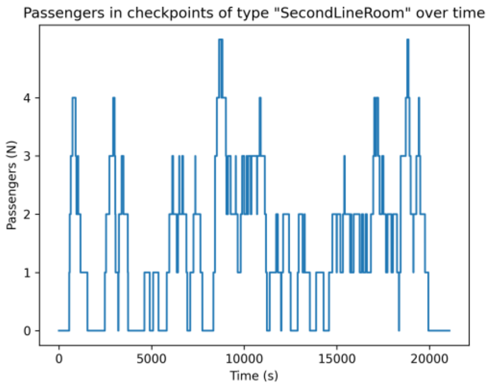


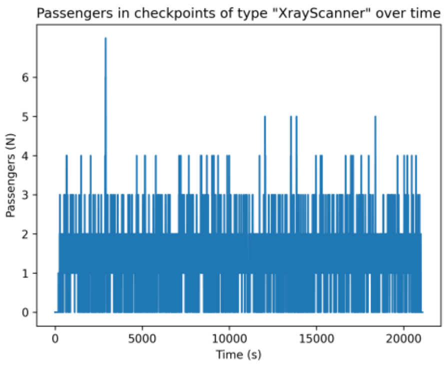


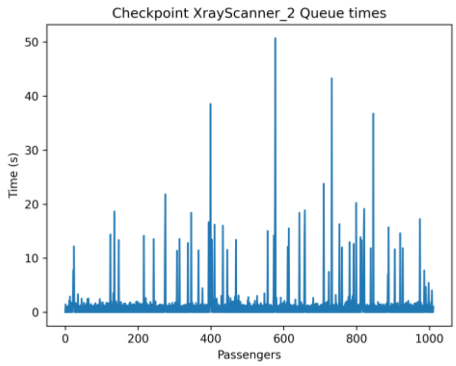





Passengers in checkpoints of type "PassportControlPoint" over time
Passengers (N)
Time (s)


Passengers in checkpoints of type "SecondLineRoom" over time
Passengers (N)
Time (s)


Passengers in checkpoints of type "XrayScanner" over time
Passengers (N)
Time (s)


Checkpoint XrayScanner_2 Queue times
Time (s)
Passengers