D9.7 Framework for assessing direct ethical, legal and societal impact of risk based border screening concepts

Document Date: 30/01/2020

Work Package 9: Ethics and Data Protection

Document Dissemination Level: Public

TRESSPASS - robusT Risk basEd Screening and alert System for Passengers and luggage
Funded by the Horizon H2020 Framework Programme of the European Union for Research and Innovation, Grant Agreement number: 787120

ABSTRACT

The goal of the TRESSPASS project is to develop, demonstrate and validate a single cohesive risk-based border management concept for air, maritime and land border crossing points. As part of this goal, the project follows an Ethics and Data Protection by Design (EDPbD) approach that builds on previous ethical research (Volkmann 2017).

Based on the typology of ethical, legal and societal issues of risk-based border management presented in TRESSPASS's deliverable D9.6, this report is a preliminary and partial definition of a framework for impact assessment. The framework aims at allowing a *comparative assessment* of different procedural designs for border checks. It aims at allowing a better understanding of the trade-offs involved in introducing *especially risk-based* border checks as part of a future border management regime. This is done by comparing the effects of the procedural designs of border crossing points along with the twelve types of relevant impact specified in the typology.

The present report describes the adaptation of the method that will be used for comparative impact assessment of border checks along with four-point ordinal scales ("normative measurement"). It then continues to analyse how each of the twelve types of ELSA related impact relates to the TRESSPASS developed enabling technologies for risk-based border management. This way, it is ensured that the framework can adequately reflect the impact of risk-based checks. As part of this analysis, options for value-sensitive design during technology design become apparent (cf. also deliverable D3.1 and D4.3) and preliminary findings on the ethical trade-offs implied by different forms of traveller risk assessment were established.

For two of the twelve types of impact, modes of impact and observable indicators for impact assessment were already defined in this report. In deliverable D9.8, this will also be done for the rest of types of impact defined in the typology along with the definition of the coding and aggregation rules for qualitative evaluation of the impact along four-point ordinal scales. Furthermore, a preliminary outline was presented on how WP6's collaboration between the TRESSPASS and PERSONA projects on traveller acceptance data could be used in relation to the ethical framework. Lastly, the ethical frameworks integration with WP6's overall CONOPS framework and WP7's simulation and evaluation platform has been described in more detail.

The framework presented in this report aims at allowing a better, comparative understanding *of the positive as well as of the negative effects* of introducing risk-based border checks as part of a future border management regime. We hope that this will allow ethically informed and well-balanced decisions about the kind of border checks desired for Europe's external borders (Ethics and Data Protection by design).

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	**D9.7**
Document Title	**Framework for assessing direct ethical, legal and societal impact of risk based border screening concepts**
Work Package reference	Task 9.3
Dissemination Level	Public
Author(s)	**Sebastian Weydner-Volkmann**
Document Review Status	☒ Consortium ☒ WP leader ☒ Technical Manager ☒ Quality and Risk Manager ☒ Ethical Advisory Board ☐ Security Advisory Committee ☒ Project Coordinator

List of Acronyms and Abbreviations

ACRONYM	EXPLANATION
AI	Artificial Intelligence
BCP	Border Crossing Point
BCSEP	Shared Evaluation Platform for Border Crossings
BG	Border Guard
CCTV	Closed-circuit television
CHAT	Cultural-historical activity theory
CONOPS	Concept of Operations
DoW	Description of Work
EC	European Commission
EDPbD	Ethics and Data Protection by Design
ELSA	Ethical, Legal, Societal Aspects
EU	European Union
FP7	Seventh Framework Programme of the EU
H2020	Horizon 2020 Programme
KPI	Key Performance Indicator
LEA	Law Enforcement Agency
MMCAT	Multi Modal Communication Analysis Tool
VTC	Video Tracking Component
NGO	Non-governmental organization
PERSONA	Privacy, Ethical, Regulatory and Social "No-gate crossing point solution" Acceptance (research project)
RBBM	Risk-based border management
RFID	Radio-frequency identification
SOTA	State of the art
SPOT	Screening of Passengers by Observation Techniques
TRESSPASS	robusT Risk basEd Screening and alert System for PASSengers and luggage
XP-DITE	Accelerated Checkpoint Design Integration Test and Evaluation (research project)

Table of Contents

1. INTRODUCTION

1.1. Background

The goal of TRESSPASS is to develop, demonstrate and validate a single cohesive risk-based border management (RBBM) concept for air, maritime and land border crossing points. This innovation action project addresses border control tasks at regular border crossing points, such as customs and smuggling prevention, immigration control, police searches for suspects, as well as cross border crime and terrorism prevention. Under a newly developed single cohesive concept, related threats will be managed as risks tailored to the specific situational needs of individual border crossing points. The project follows an "ethics and data protection by design" approach that builds on previous ethical research as part of the EU FP7 project XP-DITE.

The goal of TRESSPASS is to:

- Develop a single cohesive risk-based border management concept.
- Apply an ethics and data protection "by design" approach.
- Include passenger trust in risk management model and perform sensitivity analysis and optimization.
- Develop three pivoting pilot demonstrators.
- Demonstrate the validity of the single cohesive risk-based border management concept by using red teaming and simulations.
- Prepare for the further development of this concept beyond this project by linking to other known risk-based border management projects (in- and outside EU, within EU research frameworks and on national levels), and describe how their results contribute to a single cohesive risk-based border management concept.

TRESSPASS's ethics and data protection *by design* approach means that ethical considerations are an integral part of the project's efforts. There are four main aspects to the ethical work in the project: (1) The research ethical work (covered in Task 9.1) encompasses everything that deals with ethical standards of responsible research (such as informed consent for volunteers) throughout the duration of the project. (2) The legal framework analysis (covered in Task 1.4, already concluded) deals with the conformity with various regulations on border checks and data protection. Since RBBM is currently not foreseen in the legal framework, there is currently no corresponding legal basis for such practices in an operational context. However, TRESSPASS operates under the assumption that a change in the legal framework may happen in the future. (3) In order to understand the ethical, legal and societal impact of such a future scenario where RBBM would be operational at the borders, we conduct a corresponding ethical impact assessment (covered in T9.2-T9.4). For this, we develop a method for conducting a range of „what-if" scenarios for decision-makers (including BCP designers, legislators and the traveling public) to understand how different designs of risk-based border crossing points would compare to current forms of customs and border checks in regard to ethical, legal and societal aspects – and how one could minimize such impact during the design. (4) Based on this method, contributions are made to the technical work packages with regard to value-sensitive design during component development so as to mitigate or reduce ethical impact by changing how the specific technologies are designed. The present deliverable is part of (3).

1.2. Aim and structure of this document

All forms of surveillance and control imply *some* conflicts between their primary security goals and other fundamental norms (e.g. the protection of privacy). Different variants ("designs") of risk-based border checks will imply different ethical and societal conflicts in varying degrees of severity. This document provides a preliminary and partial version of a methodological framework to assess the severity of such conflicting impact on fundamental norms. Deliverable D9.8 will present the final, complete version.

When thinking about how border checking procedures should look like at specific border crossing points (including, e.g., what data should be collected and processed), the framework presented in this document is intended as an intellectual tool to get a better understanding of what certain design decisions would entail from an ethical point of view. The framework, then, is meant to be a tool for more comprehensively informed decision making on the ethical trade-offs involved – including public and political discussions around what kind of checks should be legally permissible at Europe's external borders, e.g. with regard to differentiated (risk-based) border checks. Hence, this deliverable develops the relevant groundwork for "allowing more comprehensively informed decision making regarding ethical, legal and societal aspects" towards TRESSPASS's main objective O2 and KPI8.

Chapter 2 of this document will describe the method to be used for developing this framework. Based on the typology of relevant ethical, legal and societal aspects of risk-based border checks (deliverable D9.6), Chapter 3 will then further specify the identified types of impact in the context of enabling technologies for RBBM and different risk assessment methods. Chapter 4 will then provide a first set of examples for observable indicators to be taken into account for the ethical impact assessment. The list of observable indicators is to be completed in the final version of the framework (deliverable D9.8), along with the corresponding coding and aggregation rules outlined in chapter 2. Chapter 5 will take a first look at the relationship between the ethics framework and the use of traveller acceptance data as part of the TRESSPASS-PERSONA cooperation. Chapter 6 deals with the integration of the ethics framework into TRESSPASS's overall CONOPS approach (WP6) and its simulation activities (WP7).

1.3. Input to and output of this document

The preliminary framework developed in this document is based on deliverable D9.6's typology of ethical, legal and societal issues of risk-based screening. As consistent with the DoW, the methodology has been adapted from previous work on aviation security as part of the EU FP7 project XP-DITE (Volkmann 2013; 2017). Furthermore, input has been used from work package WP6 on the TRESSPASS-PERSONA cooperation and on the overall CONOPS approach. Input from WP7 was used on the simulation approach as part of the BCSEP.

As part of the contributions to value-sensitive design, results from this deliverable are used in D3.1 on sensors, in WP4 (especially in D4.3 on Real Time Behavioural Analytics). Furthermore, the findings presented in section 3.3 were used in D2.3 to provide an ethical dimension on risk assessment methods. The framework presented in this deliverable forms input to WP6's overall CONOPS approach and contributes to the TRESSPASS-PERSONA cooperation, as well as T7.5's simulation efforts on the evaluation platform. Lastly, this report is a preliminary and incomplete version of TRESSPASS's ethical framework. It will form the basis of the complete version (deliverable D9.8) and for the guidelines for decision-makers.

2. METHODOLOGICAL CONSIDERATIONS

In this section, we describe the method used measuring the different normative concepts that were identified as part of the typology of relevant ethical, legal and societal aspects presented in D9.6. This method is well established for measuring the normative concept of democracy, and it was previously adapted for assessing the ethical impact of passenger screening in aviation security as part of the project XP-DITE (Volkmann 2013; 2017). Sections 2.1.1 to 2.1.3 will present the three methodological steps involved. Section 2.1.4 will briefly cover aspects of impartiality and reproducibility. In section 2.2, we then show how we can adapt this approach for the purpose of assessing the ethical, legal and societal impact of introducing risk-based border checks.

The methodological description presented in this chapter is an adaptation of the corresponding methodological sections in XP-DITE's D7.10 (as consistent with the DoW). Therefore, it shares some of the sections from the relevant deliverables (especially the more general discussion in section 2.1 below).

We will assess such impact for the border crossing point (BCP) as a whole. This means that, instead of assessing ethical, legal and societal aspects (ELSAs) related impact for certain components and sub-systems, our assessment will focus on the system-level of Risk-Based Border Management (RBBM). This is done so as to support decision making with regard to the concept operations (CONOPS) and design of risk based BCPs. This will be outlined in more detail in section 2.2.1 below.

As a help to understand how the methodological steps in normative measurement are related to each other, one can summarize them as follows: At first, a normative concept is differentiated into several semantic aspects of this concept, such as "bodily privacy" and "spatial privacy" as aspects of the normative concept of "privacy". One could think of this as a kind of "semantic differentiation" of an abstract normative concept into several more concrete aspects, which can, in turn, be further differentiated and concretized. Second, these more concrete semantic aspects are measured separately. Then, in a third step, the different evaluation scores generated for the partial aspects on a more concrete level are aggregated into one evaluation score for the abstract normative concept.

We argue from the epistemological position of moral pragmatism. Hence, the method of "normative measurement" we use for developing the ethical framework for evaluation will need to be adapted correspondingly. We do so mostly following John Dewey's version of moral pragmatism (Dewey 1988; Dewey and Tufts 1985) Consequently, the validity of the approach presented here should be seen in its contextual utility and applicability. The framework we present does not make use of deductions from a set of "values or ends in themselves" or from universal moral principles. Rather, we adopt an established method to define a framework that can be used by decision-makers as an intellectual tool for dealing with the moral dilemmas and intricacies involved when designing risk-based BCPs.[1]

2.1. The method of "normative measurement"

A rich methodological debate around normative measurement has unfolded in political science around the problem of assessing *how democratic* different states are in comparison

[1] A much more detailed methodological discussion of such a pragmatist adaptation of moral measurement has been developed in the context of aviation security screening (Weydner-Volkmann 2018).

to each other or across time. The research problem, here, is structurally similar to our own in that complex empirical situations (different states or the same state at different points in time) are to be comparatively analysed with respect to differences in the quality of a normative concept ("democraticness"). Before we address how this method can be adapted for assessing different BCPs with respect to differences in the quality of normative categories such as privacy and data protection, we will briefly introduce the three main steps of normative measurement in its original context, i.e. the measurement of "democraticness".[2]

2.1.1. Contextual differentiation of a concept in a concept tree

The first methodical step consists in the differentiation of an abstract normative concept in different, less abstract aspects that are relevant to the context of what is being analysed. This is necessary, as abstract concepts like "democracy" cannot be observed or assessed directly. The first step of normative measurement, thus, consists in what could be called "semantic differentiation" of the concept. Munck and Verkuilen suggest using a concept tree that first identifies all relevant attributes of democracy and, thus, avoids minimalist or maximalist definitions of the concept, which would either exclude important attributes or include irrelevant ones (Munck and Verkuilen 2002). Those attributes are then further differentiated by identifying the "components of attributes" (Munck and Verkuilen 2002) Using the Democracy Barometer (Bühlmann et al. 2012) as an example, one could identify three attributes: freedom, control, and equality. For each of these attributes, components are identified. By doing so, "semantic components" are identified in a strictly hierarchical way (cf. Figure 1). This process is continued until a level detail is reached that allows the identification of observable indicators for each of the sub-components in the next step. The sub-components on the lowest hierarchical level are called the "leaves of the concept tree" (Munck and Verkuilen 2002).

The assumption evident in the methodological terminology is that the meaning of an abstract normative concept can be "split" into several semantic components. Indicators for such a semantic component are, thus, also indicators for the concept as a whole. Because this is not necessarily true for semantic concepts in natural language, it is crucial that the process of semantic differentiation is conducted in a formal and strictly hierarchical way: On the different levels of the concept tree, the semantic attributes, (sub-) components and concept leaves must be sharply differentiated, i.e. not overlapping or redundant, and at the same time comprehensive, i.e. not leaving out important aspects (Munck and Verkuilen 2002; Jäckle, Wagschal, and Bauschke 2012).

From the position of moral pragmatism, it is important to note, however, that despite the somewhat misleading terminology, this process of semantic differentiation cannot be regarded simply as a kind of "disassembly of semantic concepts" – similar to how one would imagine the physical disassembly of an object into smaller parts. Instead, pragmatism holds that abstract normative concepts such as democracy, security or privacy only become meaningful in an applicatory context (Dewey 1988). Put differently, normative concepts are commonly ambivalent in their meaning and they often denote different things in different situational contexts. It is, therefore, important to specify the applicatory situation for the framework to be developed. The process of semantic differentiation, then, is not a process of "deduction" from the definition of a normative concept. Rather, it is a differentiation of those semantic aspects that are *relevant in the framework's situational context of application*.

[2] Our analysis is mainly based on Munck and Verkeulen (2002) and Jäckle, Wagschal, and Bauschke (2012); for reference and cross-checking, Lauth (2007) as well as Müller and Pickel (2007) were also used.

The process of pragmatic semantic differentiation is also a process of situational concretion: in each step throughout the hierarchical differentiation, the meaning of an abstract normative concept like democracy becomes more and more concrete within this situational context of application. Accordingly, the pragmatic adequacy of a concept tree depends on whether all *contextually relevant* aspects are fully expressed and well differentiated in the concept leaves; aspects of a normative concept that are not relevant for the context of application, on the other hand, should not be part of the concept tree – even though they may be very relevant in other applicatory contexts.

The aim in developing such a concept tree from a pragmatic perspective must, thus, be the attempt to *reconstruct which moral aspects are relevant in a given situation*. It serves the purpose of offering an intellectual tool that "can help to identify the often-unexamined principles, organizational patterns, or customary assumptions underlying behaviours and beliefs … These principles are often latent and hard to articulate, not consciously applied and debated" (Fesmire 2014).

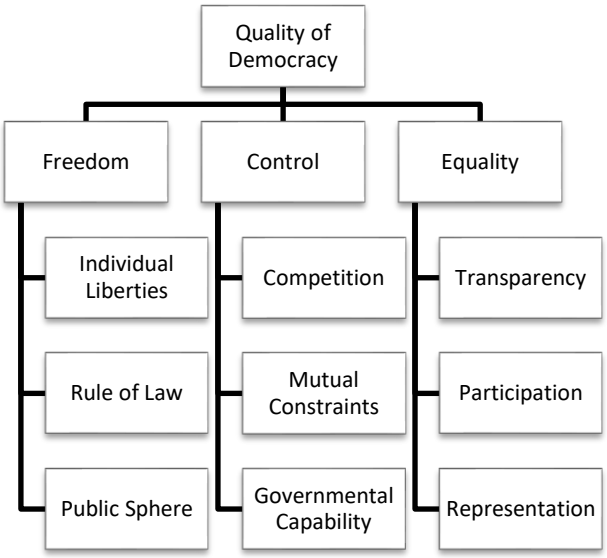

FIGURE 1: DEMOCRACY BAROMETER'S FIRST THREE LEVELS OF THE CONCEPT TREE (BÜHLMANN ET AL. 2012)

2.1.2. Measurement: selection of indicators, coding rules and scales

In the second methodological step of normative measurement, each of the concept leaves is then operationalized for measurement by selecting a set of valid observable indicators. "Because there are no hard and fast rules for choosing valid indicators, this is one of the most elusive goals in social sciences" (Munck and Verkuilen 2002). Therefore, it is crucial to document well, which indicators were chosen for which reasons. The goal is to choose indicators that fit well with the concept leaves, i.e. which actually measure what they are meant to measure.

According to the established methodology (Munck and Verkuilen 2002), indicators should cover only one concept leaf at a time and sets of indicators should cover one concept leaf as comprehensively as possible, in order to avoid a systemic bias by using the same indicator for several concept leaves or not measuring important aspects of the concept leaf at all. At the same time, it is a good rule of thumb to choose only a small set of relatively simple indicators and to not rely on solely one indicator for one concept leaf, especially when statistical indicators are used, as one needs to cope with the unavoidable statistical error (Munck and Verkuilen 2002).

From a pragmatic perspective, however, the main criterion for well-chosen indicators is, again, the adequacy with regard to the applicatory context. The main question, thus, is not whether the indicators "validly" capture the general meaning of the concept leaf in question. Instead, the pragmatic criterion for "contextual validity" refers back to the situations in which the framework is meant to be used and how the concept leaf can specifically be observed in those situations.

The indicators hereby identified will ultimately allow the measurement of the different concept leaves. However, it is important to note that we need to understand "measurement" in a sufficiently broad sense. In a ground laying article, Stanley Smith Stevens proposes to distinguish four "scales of measurement" (Stevens 1946):

1. Commonly we measure beginning from a point of absolute zero (e.g. there is no measurable physical extension of an object at all) in clearly defined intervals (e.g. $^1/_{100}$ part of the standardized extension called a "meter"). This kind of scale is called **ratio scale**.
2. If there is no point of absolute zero, but clearly defined intervals (e.g. one hour today is as long as one hour tomorrow, but there is no point of absolutely zero hours), this is called an **interval scale**.
3. If there are no clearly defined intervals, but a specified hierarchy between different steps on the scale, we speak of an **ordinal scale** (e.g. a scale in a survey indicating "I completely agree", "I agree", "I disagree", and "I completely disagree").
4. Finally, if even a hierarchy cannot be specified, the process of measuring will consist of nothing more than a categorization (e.g. differentiating players into a blue team and a red ream). This is then called a measurement on a **nominal scale** – although this is commonly not considered a form of measuring at all.

The different scales of measurement allow different mathematical operations (Stevens 1946). For the measurement of democracy, ordinal and interval scales are most relevant. The type of indicators that were selected to operationalize the concept leaves determine which type of scale can be used for measurement. For example, if the indicators do not provide the necessary metrics to inform an interval scale, but only qualitative information that allows ranking (e.g. sorting from better to worse), ordinal scales should be used for assessment.[3]

Furthermore, the intervals or steps of the assessment scales, as well as their minimum and maximum value need to be selected wisely so that qualitative differences in "democraticness" between the relevant countries can be mapped adequately. Otherwise, the scales will either measure too roughly or too detailed. This choice needs to be justified depending on what kind of differences the index is supposed to measure (Jäckle, Wagschal, and Bauschke 2012).

Hence, this can be seen as a pragmatic element inherent in the construction of a framework for normative measurement. For example, if all countries of the world should be assessed with the framework, the intervals and minimum and maximum values will need to be chosen differently than if the framework should only assess countries within the EU: the kinds of qualitative variations among democracies that are to be expressed by the framework will differ depending on the purpose of the framework. Thus, the choice of scales used for measurement needs a pragmatic justification of adequateness with regard to the purpose of the framework. In the construction of the assessment scales lies, thus, an *interest in*

[3] Especially in psychology, but also in normative measurement, ordinal scales are often used but subsequently treated as interval scales. However, this "methodological leap" needs to be justified in some way so as to not invalidate the findings.

qualitative differentiation. Well-defined scales make this interest in qualitative differentiation explicit while they also fit to the kinds of indicators used.

Finally, the "gap" between the observable indicators and the assessment scales needs to be bridged. This is done by defining a set of coding rules: Which combinations of indicator values will result in which evaluation on the assessment scales? Again, this is a very elusive process with no clearly defined rules, so it needs to be well documented and justified (Jäckle, Wagschal, and Bauschke 2012).

At this point, the method will ultimately have to rely on some form of common-sense argument that certain combinations of indicator values should result in certain evaluations on the assessment scales in order to express the qualitative differences we are interested in. From a pragmatic perspective, we can say that the more concrete the conceptual differentiation is carried out and the better the qualitative interest in differentiation is specified for the scales, the smaller this gap will be. The common-sense arguments used as coding rules should, thus, firmly rest on both sides of the gap: in the light of the framework's specific context of application and for each of the relevant, situationally concrete aspects of the normative concept "democracy" (expressed in the concept leaves), the coding rules need to reconcile the interest in qualitative differentiation with the possible empirical cases indicated through the indicator values.

2.1.3. Aggregating partial measurements to an overall measurement

While the first step "semantically differentiated" the normative concept of democracy in a concept tree and the subsequent second step developed a concept of measurement for the concept leaves, the third step of the methodology covers how these measurements can be aggregated into an overall evaluation of how democratic different countries are.

The rules according to which this aggregation takes place also need to be justified for each step of the aggregation. Some authors interpret this step as a kind of mirroring of the semantic differentiation, i.e. the assumed relationships between the different levels of the concept tree should match the expression of the aggregation rule (Jäckle, Wagschal, and Bauschke 2012).

Additionally, those authors argue that if different indicators, concept leaves, (sub-) components or attributes vary in their relevance for the concept of democracy, their weighting may need to be adapted accordingly. Thus, each step of aggregation also needs to be justified with respect to the relevance of the aggregated values (Jäckle, Wagschal, and Bauschke 2012).

Finally, those authors hold that it is also crucial to justify that the loss of detail that will necessarily happen with each step of aggregation will not introduce biases for the assessment process and will not lead to ignoring relevant differences (Jäckle, Wagschal, and Bauschke 2012). It may, therefore, make sense to stop the aggregation on a certain level of abstraction in the hierarchy. An example of such a limited aggregation is the Democracy Barometer (Bühlmann et al. 2017), where nine evaluation scores can be presented for each of the countries examined (cf. Figure 2).

However, similarly to what we have said for the construction of the assessment scales and for the definition of the coding rules in section 2.1.2, a pragmatist reading of this step in the methodology would emphasize the qualitative interest in differentiation more strongly. In a certain way, the aggregation of the evaluation scores indeed needs to "mirror" the function of the semantic differentiation of the normative concept "democracy". However, the adequacy of the aggregation rules and the relative importance of the conceptual aspects

expressed in them is determined with respect to the purpose of the framework, i.e. with respect to the researchers' *interest in qualitative differentiation*.

Again, the method will ultimately have to rely on some form of common-sense argument that certain combinations of evaluation scores for the semantic sub-aspects should result in certain evaluation scores in order to express the qualitative differences we are interested in. In the light of the framework's specific context of application and for each step of aggregation, the aggregation rules need to reconcile the interest in qualitative differentiation with the possible empirical cases indicated through the indicator results.

From a pragmatist perspective, it is, therefore, a natural assumption that we should stop the aggregation at a point that serves the purpose of the inquiry, i.e. where the framework still gives us as much detail as necessary, but also reasonably limits the amount of information provided by the evaluation.

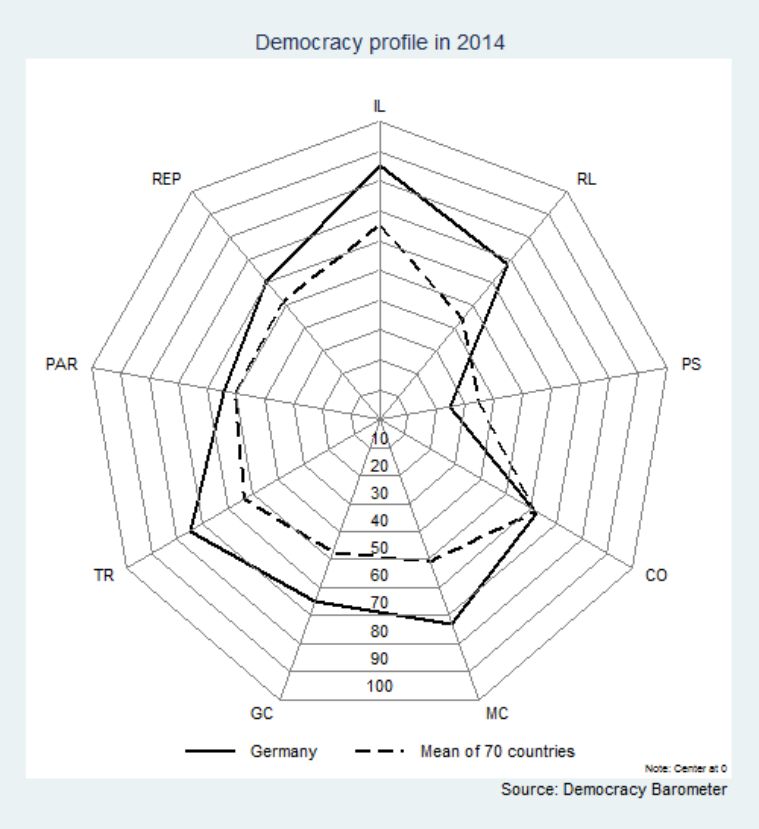

FIGURE 2: DEMOCRACY BAROMETERS COUNTRY EVALUATION FOR GERMANY (BÜHLMANN ET AL. 2017)
LEGEND: (PRINCIPLES AND FUNCTIONS)
PRINCIPLE: FREEDOM: RL (RULE OF LAW), IL (INDIVIDUAL LIBERTIES), PS (PUBLIC SPHERE).
CONTROL: GC (GOVERNMENTAL CAPABILITY), MC (MUTUAL CONSTRAINTS), CO (COMPETITION).
EQUALITY: TR (TRANSPARENCY), PAR (PARTICIPATION), REP (REPRESENTATION).

2.1.4. Impartiality and reproducibility of the measurement

As with all scientific research, neutrality and reproducibility are paramount for determining the validity of a method. For measuring democracy, this means especially that (1) all relevant societal perspectives should be included in the concept tree; (2) there should be no bias for

one particular relevant perspective in selecting and encoding the indicators and in constructing the assessment scales; (3) there should be no bias for one particular relevant perspective in the definition of the aggregation rules.

Naturally, it is never possible to completely neutralize the researchers' subjective perspectives and biases. While, nevertheless, there must be an active effort to include all relevant perspectives in the different steps of measuring normative concepts, this also means that it is especially important to properly document the methodology of the measurement process. Not only does the methodology have to be open to scientific critique, but it must also be documented detailed enough that others, who follow the same procedures, will also come to the same results.

For measuring democracy, this means that (1) all criteria for selecting and measuring the indicators must be precise so that the researchers' perspective of what democracy is and how it can be measured becomes transparent and differing interpretations of the methodology can be minimized; (2) all rules for coding the measured indicator values on the assessment scales must be precise so that all steps of evaluating them can be followed and replicated by others; (3) all thresholds on the assessment scales must be precisely defined so that the same values on the assessment scales will be interpreted or categorized by others in the same way, e.g. as the same "type of democracy."

2.2. Concept for the evaluation of risk based BCP systems

In section 2.2, we illustrate how the approach to measuring democracy can be adapted to our purpose of assessing the ELSA related impact of risk-based border checks. As will become clear in the following sections, we will base the process of semantic differentiation on the typology of ethical, legal and societal issues of risk-based screening (cf. deliverable D9.6). This typology is centred around the *specific procedural arrangements* that travellers are subjected to as part of the border checks.

2.2.1. System-level evaluation of specific paths through the BCP

As mentioned in deliverable D9.6, the logic of border checks can be conceptualized as preforming two main functions: (1) the access and egress control function and (2) the revelatory function. These functions have been conceptualized regarding the BCP *as a whole* (on the "system-level"), not with regard to individual measures. This is important to note, as only such a wholistic perspective will allow a comparison of BCPs with different procedural arrangements. This is due to the fact that the ethical implications of a given border checks measure (e.g. the use of x-ray luggage scanners as part of customs checks) can play out very differently depending on how it is integrated into the overall border checks procedures: are travellers selected completely randomly for this? At what rate? Or is a specific group amongst the travelling public targeted for this measure?

We thus need to define how to evaluate BCPs as a whole (from a system-level perspective). This will also need to be compatible with the overall CONOPS approach developed in work package WP6 (cf. section 6.1 below). For this, we propose an approach to the system-level of border checks that is based on the "paths" that travellers can take through a BCP according to its specified procedural design.

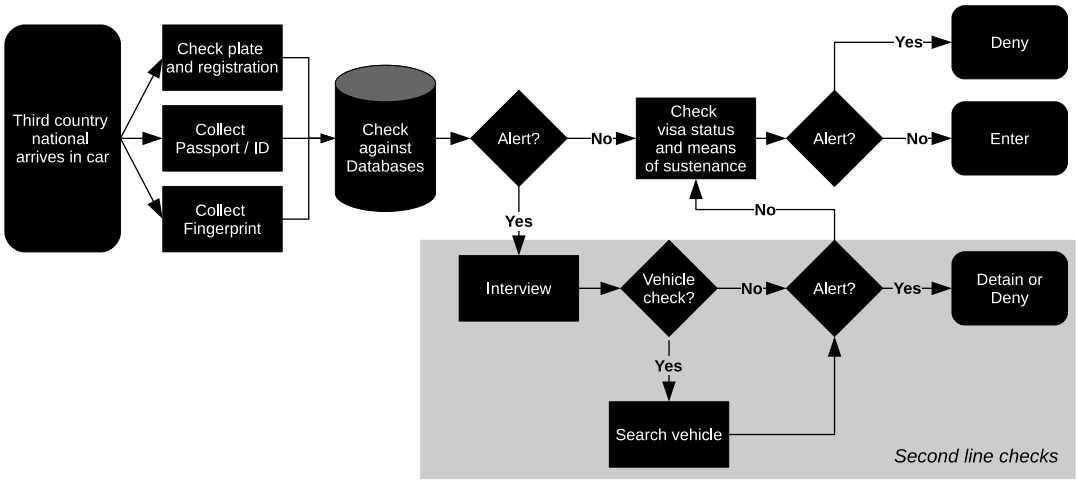

FIGURE 3: SIMPLIFIED ILLUSTRATION OF SEVERAL "PATHS" THROUGH CURRENT BORDER CHECKS PROCEDURES

Figure 3 provides a simplified example for a current procedural design for land border checks for third country nationals that arrive by car (excluding customs checks) based on the Schengen Border Code (EU 2016a). Here, we can see that the specific border checks procedures a traveller is subjected to depend on whether or not previous checks result in an alert. For example, depending on whether the checks against databases such as the Schengen Information System result in an alert, the traveller will be referred to an interview as part of the second line checks. If it doesn't result in an alert the first line checks are continued with checking the visa status and whether the traveller can demonstrate the ability of self-sustenance. Overall, there are eight possible paths through the checkpoint that need to be evaluated regarding the ELSA related impact.[4]

In this example, it also becomes clear that the severity of such impact will differ considerably depending on which path travellers will take: In case a traveller is referred to second line checks, the interview and potentially also the vehicle checks will imply considerably higher ELSA related impact compared to first line checks only. Hence, the different paths through the checkpoint will need to be evaluated separately. And consequently, this also means that there will need to be an additional step of aggregating the evaluation results from the different paths in order to provide a limited set of ELSA scores for a BCP on the system-level.[5]

2.2.2. Semantic differentiation of impact in a concept tree

The evaluation of the different paths through a BCP will be based on deliverable D9.6's typology of ELSA related impact. This typology was developed in a systematic manner to ensure its comprehensiveness and that the different types of impact are sharply differentiated from each other. In addition to that, it is structured in a strictly hierarchical way: The typology identified twelve distinct types of ELSA related impact of border checks, with a special focus

[4] Depending on the complexity of the procedural design of border checks, establishing all possible "paths" through a BCP can become rather complex – especially if we also account for the different status of travellers (Union citizen and persons of similar status vs. third country nationals) and for customs checks. This is why we aim for a shared formal description of procedural designs throughout TRESSPASS, so as to make sure that, once completed for a given BCP, it can be reused across different tasks in the project in an integrated manner (cf. chapter 6).

[5] For example, for the BCP design illustrated Figure 3, the overall privacy and data protection impact will depend on the rate of travellers sent to second line checks.

on risk-based border checks. These twelve types of impact are grouped into three categories (cf. Table 1, deliverable D9.6).

TABLE 1: RELEVANT TYPES OF ETHICAL, LEGAL AND SOCIETAL ASPECTS (ELSAS) FOR RBBM

ELSA category A: Privacy and data protection intrusions	ELSA category B: unfair distribution of impact across different social groups	ELSA category C: restrictions of societal freedoms and liberties
Intrusion into spatial privacy	Disproportionate impact due to infeasibility of standard checks	Accosting travellers
Intrusion into bodily privacy	Disproportionate impact due to accumulation of false alarms	Lack of accountability
Intrusion into private life	Disproportionate impact due to false or incomplete external data	Restriction of self-determination and misuse of data
Disclosure of information	Impact on non-travellers	Lack of transparency

There are different contending definitions, however, of what a semantic concept actually is. For our purposes, it sufficed to use a very broad one: As long as something can be expressed as one reasonably specific and coherent meaning within the situational context of border checks, we considered it as a semantic concept in this sense. This is also true if two concepts are so intimately related that it is difficult to treat them as distinct within the situational context. This is true for the two concepts of privacy and data protection: As we have argued in deliverable D9.6, we follow Ralf Poscher (2017) in considering data protection as a systematic enhancement of other rights, liberties and equality interests. In this sense, we propose treating privacy and those aspects of data protection that aim at *enhancing the right to privacy* as one semantic concept in the context of border checks.

From a methodological perspective, thus, we argue that each of the typology's three ELSA categories can serve as one semantic concept that can be situationally differentiated into a concept tree:

A. Impact related to the intrusions into travellers' privacy and data protection
   - *Semantic concept of "intrusiveness" of border checks*
B. Impact related to discrimination and unfair distribution of impact due to errors
   - *Semantic concept of "unfairness" of border checks*
C. Impact contributing towards the restriction of societal freedoms and liberties
   - *Semantic concept of "restrictiveness" of border checks*

This will mean that we will not aggregate the evaluation of border checks to just one overall "ethics score", but instead provide a few distinct values, similar to Democracy Barometer's approach (cf. section 2.1.3). This will also enable us to reflect in the evaluation the fact that there may be a contradictory relationship between, for example, the intrusiveness of border

checks and the unfairness category, i.e. some BCP designs may better protect the privacy of the majority of travellers at the expense of discriminating against some groups.

As will become clear in chapter 4, we will provide one additional layer of differentiation for each of the types of impact identified in deliverable D9.6. This differentiation will be based on the modes of impact, i.e. on *how the impact becomes manifest for the traveller* with regard to each type of impact. Those modes of impact will methodologically serve as leaves of the three resulting concept trees. This is consistent with the approach taken in XP-DITE (cf. Volkmann 2017).

2.2.3. Measurement: selection of indicators, coding rules and scales

The concept leaves will be contextually specific enough to allow a suitable selection of relevant indicators. In order to support some form of computer-assisted evaluation, only binary values for indicators will be used, e.g. “Is it required to provide a social media account? Yes/No” or “Do border guards touch the traveller’s body? Yes/No”. Regarding the last example, sometimes more detailed information is needed for the ethical evaluation. In such a case, this will be realized through further binary indicators, e.g. “If a border guard touches the traveller’s body, will the border guard touch intimate zones? Yes/No”.

We will then define the coding rules as formalized expressions of which combinations of indicator values will result in which evaluation on the assessment scales. On the one hand, this will allow for adapting the framework relatively easily for a computational model. On the other hand, this will also solve the problem that using several indicators to measure one specific aspect (as we will do in some cases) can result in systematic biases: Since the coding rules are just formalized expressions of conditional evaluation criteria (in the form of yes/no-questions), this will allow us to deal with the indicators in a predominantly *qualitative fashion* and avoid unintended biases resulting from one indicator disproportionately influencing an evaluation as part of a calculus. Furthermore, this will also allow us to document our rationale behind the coding rules more transparently.

We propose that the coding rules will consider all concept leaves of one type of ELSA related impact at once. Again, this is only methodologically valid since we will use qualitative coding rules rather than quantitative assessments. Thus, the hierarchically lowest level of assessment will take place on the level of the types of impact. If we take the risk of “intrusion into bodily privacy” as an example, the coding rules will need to convincingly document, how certain combinations of answers to the yes/no-questions result in evaluating the intrusion into bodily privacy on a certain level of the assessment scale.

Finally, we will construct the assessment scales qualitatively as ordinal scales with four hierarchical steps (cf. Figure 4). In deliverable D9.6, we have presented a set of “rump scales” for each of the types of impact identified in the typology. These “rump scales” consisted of the formulation of a positive and of a negative pole, e.g. for the impact type “intrusion into bodily privacy”, the positive pole was “respecting travellers’ public appearance” and the negative pole was “exposing travellers’ bodies”.

We will adapt these poles for the definition of four hierarchical steps as illustrated for the example of bodily privacy: the most positive step on the scale will be defined as being close to a realistic best-case scenario of respecting the travellers’ public appearance for border checks. Accordingly, the most negative step on the scale will be defined as being close to a realistic worst-case scenario of exposing the travellers’ bodies during the border checking procedures. The two remaining intermediate steps will then be interpreted “more respecting than exposing” and "more exposing than respecting" (cf. Figure 4).

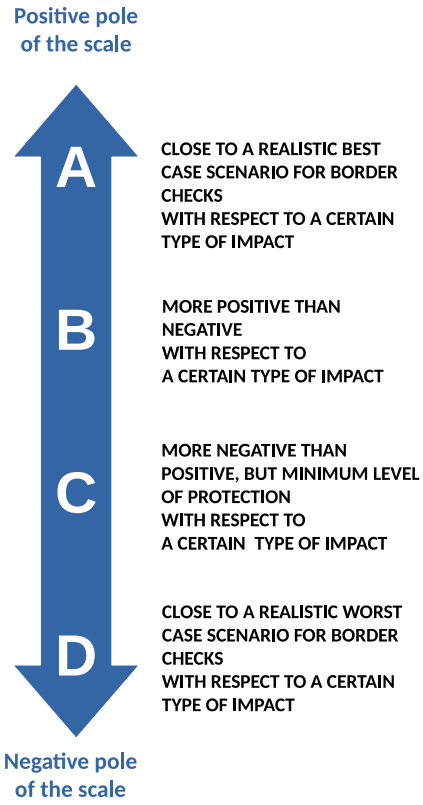

FIGURE 4: SCALES CONCEPT

We use such references to realistic best- and worst-case scenarios in the construction of the scales for the following rationale: As the binary indicators allow only qualitative evaluations, the evaluation scales have to be constructed as ordinal scales. This means that with respect to the different types of impact, they rank checkpoints along a scale from better (ethical impact is less severe) to worse (ethical impact is more severe). In doing so, the scales are not intended to measure some form of (meta-)physical properties of checkpoints. Instead, they are meant to *allow the differentiation of checkpoint designs* with respect to certain types of ethical impact, e.g. with respect to how intrusive they are into bodily privacy. If the most positive or most negative evaluation score will only be given for "unrealistically" good or bad BCPs, then this would result in too few evaluation steps for differentiation to be of any help when comparing real BCPs. The best score does, therefore, not mean "no intrusion at all" (which could only be achieved by no border checks whatsoever), but rather something along the lines of "*if some form of meaningful border checks is implemented* that can realistically contribute to reducing threats at the border, then this form is hardly intrusive *in comparison to other forms of such border checks".*

Thus, in deliverable D9.8, the coding rules will define that a certain combination of "answers" to the yes/no-questions results in classifying a border checks procedure as close to a realistic best-case or worst-case scenario – or as distinctly different from those two scenarios, but somewhere in between and closer to one of them than to the other. For border checks procedures, the evaluation score of "C" regarding the risk of "intrusion into bodily privacy" thus expresses the following: a quality of intrusiveness that is – in comparison to other realistic border checks procedures – distinctly different from a worst-case scenario, but that offers only a relative minimum level of protection for passengers' bodily privacy. By denoting a distinct qualitative difference from a realistic worst-case, the evaluation score of "C" will, thus, play a special role in the framework: it will refer to procedures that offer a minimum level of protection concerning the type of ethical impact in question.

We believe that this will allow us to make convincing common-sense arguments for the rationale of the coding rules. In addition to that, this will not pretend a finer or more detailed granularity of the evaluation than is conceivably possible.[6]

[6] Indeed, for this reason, four point ordinal assessment scales are quite common in applied ethics, as they provide limited but still adequately informative detail. Sometimes such scales are used in the form of "extended traffic lights" ranging from the best-case coded as "green" over "yellow" and "orange" to the worst-case "red".

2.2.4. Aggregating the partial measurements to overall measurements

Despite providing a lot of qualitative detail, twelve distinct ethics scores for each path through a BCP will be too many to be helpful for understanding the overall ethical impact of a BCP design. We, therefore, propose to aggregate the results in two steps.

In the first step, we will aggregate the four scores within each ELSA category to one overall evaluation score of this category. Since there are three ELSA categories in the typology, this will result in three scores for each path through a given BCP design. The second step will then be to aggregate the scores for the different paths. For both steps, aggregation rules will have to be defined.

We propose to conduct the first step of aggregation (aggregating the evaluation of types of impact into the evaluation of risk categories) in a qualitative way: Conditional rules will be formulated that will result in a meaningful categorization of "what kind of path through the BCP" we are dealing with. Figure 5**Error! Reference source not found.** presents an example of an aggregation concept for the privacy and data protection values: For each step, a qualitative description of the path through the checkpoint is presented. Qualitative aggregation rules will then be specified accordingly in a more formalized way.

For the second step of the aggregation (aggregating the results of the different ways through the BCP), some quantitative element will be necessary. For example, evaluating how intrusive a BCP is into the travellers' privacy will depend on how many passengers are going to go on each path through the system. Thus, not all paths will be treated equally. Instead, the evaluation will depend on what percentage of the travellers will face what type of border checks scenario on their specific paths. Figure 6 gives an example for the second aggregate scales concept for ELSA category A: "intrusiveness" regarding privacy and data protection.

How the percentages will be taken into account may be different for the three risk categories: For example, evaluating discriminatory effects of border checks procedures may depend on more than just the number of people using a certain path that potentially has such discriminatory impact, but also on whether there is a process of traveller differentiation in place that has specifically or inadvertently singled out members of a certain group for a more intrusive checking procedure.

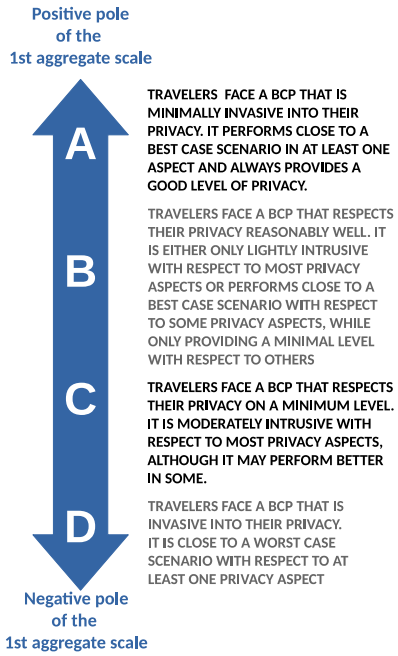

FIGURE 5: 1ST AGGREGATE SCALES CONCEPT (LEFT)

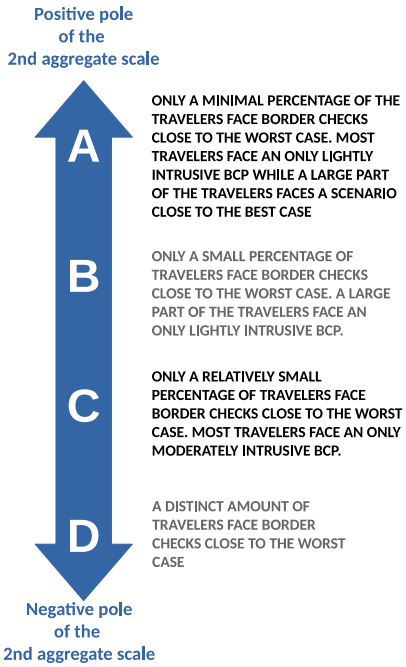

FIGURE 6: 2ND AGGREGATE SCALES CONCEPT (RIGHT)

Figure 7 gives a schematic overview of the three steps involved in our methodological approach for the creation of an evaluation framework for ELSA related impact of BCPs on the system-level.

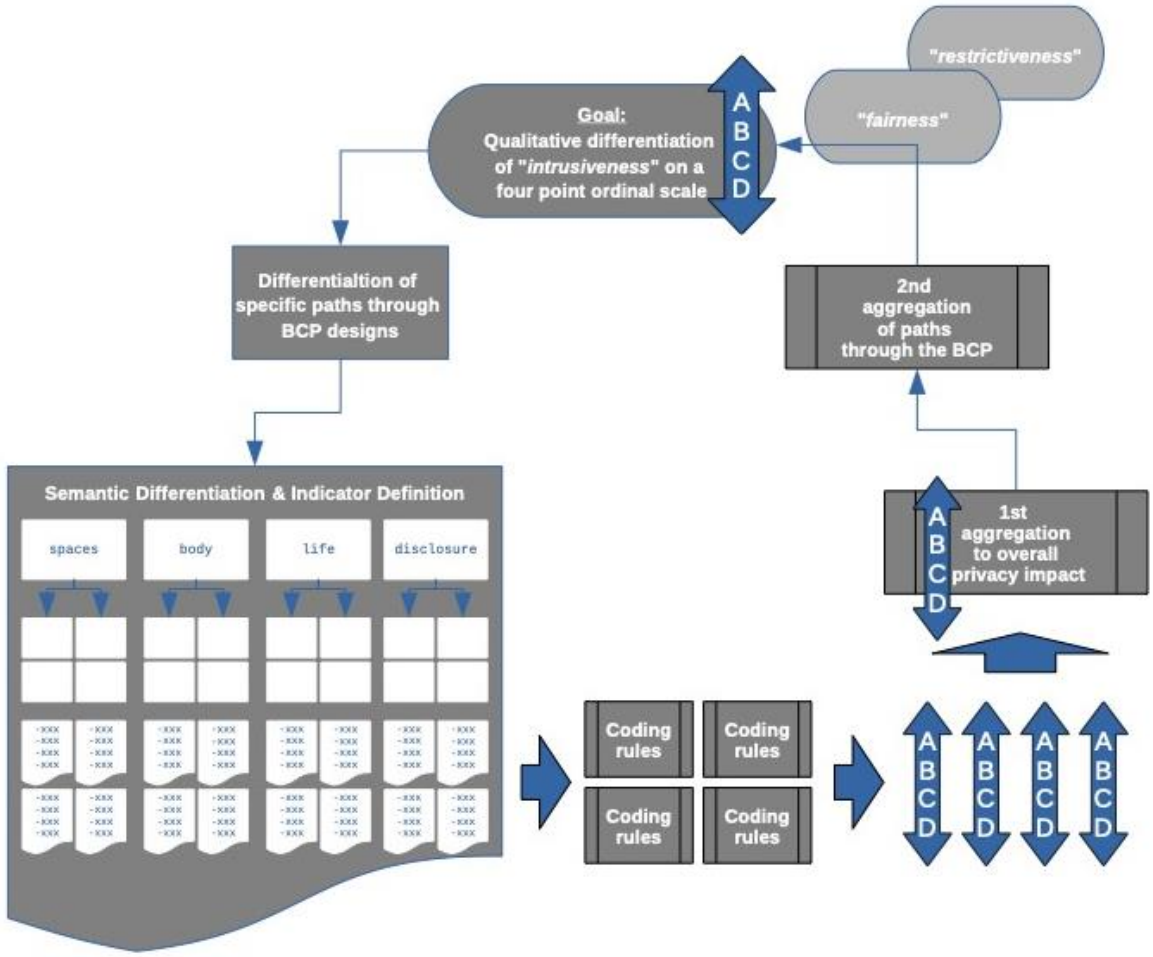

FIGURE 7: OVERVIEW OF THE THREE METHODOLOGICAL STEPS FOR PRIVACY & DATA PROTECTION EVALUATION:

(1) Differentiation of the abstract semantic concept "intrusiveness into travellers' privacy" into less abstract aspects within the situational context of border checks, up to a level of concretion that allows the definition of observable indicators.

(2) Definition of rules how different combinations of indicator values will be coded as certain qualitative scores on the evaluation scales.

(3) Definition of rules how to aggregate the different assessment scores for the less abstract components into one overall score for "intrusiveness into travellers' privacy".

3. ETHICAL, LEGAL AND SOCIETAL ASPECTS OF RBBM

3.1. ELSAs and enabling technologies

As outlined in section 2.2.1 above, the ethical, legal and societal impact of border checks will largely depend on *how sensors and other technologies are used as part of the overall checking procedures*. A meaningful impact assessment can, hence, only be given in the context of a border crossing point's CONOPS. This means that the impact of specific technologies must be assessed in the context of the checking procedures as a whole, i.e. from a comprehensive "system-level perspective" that assesses the BCP as a whole.

From such a system-level perspective, D9.6 has identified and described twelve types of relevant ethical, legal and societal impact and grouped them in three categories. For convenience, Table 2 presents again the overview over this typology of relevant ELSAs.

TABLE 2: RELEVANT TYPES OF ETHICAL, LEGAL AND SOCIETAL ASPECTS (ELSAS) FOR RBBM

ELSA category A: Privacy and data protection	ELSA category B: unfair distribution of impact across different social groups	ELSA category C: restrictions of societal freedoms and liberties
Intrusion into spatial privacy	Disproportionate impact due to infeasibility of standard checks	Accosting travellers
Intrusion into bodily privacy	Disproportionate impact due to accumulation of false alarms	Lack of accountability
Intrusion into private life	Disproportionate impact due to false or incomplete external data	Restriction of self-determination and misuse of data
Disclosure of information	Impact on non-travellers	Lack of transparency

In order to achieve the goal of more comprehensively informed decision making (cf. section 1.2), TRESSPASS's framework for impact assessment needs to adequately reflect how changes and trade-offs will likely manifest for each of the twelve types of impact when introducing risk-based border management. This is no simple task since there are multiple variants of how one may implement risk-based border checks. Furthermore, new technologies enable new concepts and designs for doing RBBM. Hence, the specifics of what the framework's assessments are supposed to "adequately reflect" are very much moving targets.

In order to better understand how RBBM may be implemented as part of a future legal framework and in order to anticipate how developments in technology may *enable* such implementations, we will analyse the new sensor technologies developed as part of TRESSPASS (T3.1), as well as TRESSPASS's novel data analytics technologies (T4.2 and T4.3). Starting from the twelve types of relevant impact, it is possible to describe how certain use cases for the newly developed sensing technologies may mitigate or aggravate such issues.

This analysis also offers input for value-sensitive sensor design as part of the corresponding technology development tasks (T3.1. T4.2 and T4.3), i.e. sensor developers can use this analysis to make efforts in finding technical solutions to mitigate ethical, legal and societal risks as much as possible or promote the minimization of ethical risks. A BCP design that makes use of such improved sensors is then likely to perform better with regard to ethical KPIs.

3.2. Analysis of enabling technologies developed in TRESSPASS

In deliverable D9.6, the logic of border checks at external borders of the EU was conceptualized as performing two main functions: (1) performing an *access and egress control function* with regard to border crossings; and (2) doing so based on *revealing non-obvious, previously unknown aspects about travellers and the goods they bring along with them* (revelatory function). Technologies developed in TRESSPASS aim at providing new capabilities that *enable certain variants of risk-based differentiation* in the performance of these two main functions.

At the current stage, there is no broad consensus on how to differentiate different variants of implementing risk based border checks. It is safe to assume, however, that different forms of differentiation between risk assessment *methods* help highlighting different *characteristics*. As has been shown previously for aviation security (Weydner-Volkmann 2017), *from an ethics perspective*, it is helpful to distinguish three main variants of RBBM based on the kind of information that is used for risk categorization: either (1) only purely situational data is used, or traveller related data is made use of, which can be related to (2) "background information" on the travellers, or to (3) the travellers behaviour during the approach to and at the BCP.[7] This leads to the following three variants of risk assessment methods, which will be discussed in the next sections and analysed with regards to their relevance for the TRESSPASS project:

1. **Situational risk-based checks (cf. section 3.2.1);**
   Only circumstantial data is used to differentiate groups of travellers according to risk. This could be, for example, trends in online drugs trade in certain geographical regions. At an airport BCP, travellers on flights from "higher risk" regions into the EU could then be subject to corresponding increased customs checks, e.g. by increasing the rate of luggage that is searched.

2. **Traveller differentiation based on risk profiles (cf. section 3.2.2);**
   Some form of "background information" is collected on travellers to group them in different risk categories. For example, PNR data or publicly available online information on a traveller could be checked against "risk profiles" that are (un-)associated with certain threats in border security.

3. **Traveller differentiation based on behavioural analysis (cf. section 3.2.3).**
   Some form of analysis of the travellers' behaviour takes place during the approach to or at the BCP. Travellers that display behavioural characteristics that are associated with certain threats in border security may be categorized as higher or lower risk.

[7] These three variants do not neatly map to the different travel stages identified in deliverable D2.3. However, it is clear that the travel stages are relevant in different ways to the three variants; e.g., the pre-travel stage is more relevant to the use of "background information" than it is for an analysis of travellers' behaviour.

Distinguishing between these three variants of risk-based checks will help us better understand *how the enabling technologies developed in TRESSPASS impact each of the twelve types of ethical, legal and societal impact identified in deliverable D9.6*. As a first step towards developing the modes of impact as concept leaves for each of the three ELSA categories, this will be undertaken in the next sub-sections. In section 3.3, as a second step, we will generalize somewhat from the TRESSPASS developed enabling technologies and describe the impact of risk assessment methods with a broader range of technologies in mind.

3.2.1. Enabling situational risk-based border checks

In the first variant, checks are differentiated in intensity and in the use of resources based on contextual factors that do not relate to individual travellers. An example for this could be a situation, in which a large amount of plastic explosives has been stolen in a neighbouring third country. As part of risk-based border checks, one may assume that this leads to a heightened risk that plastic explosives are smuggled across the EU's external border and it may be decided that this heightened risk should be addressed as part of the border checks.

In TRESSPASS, a "dark web crawler" component is developed as part of the Web Intelligence module that facilitates such a form of risk-based differentiation of border checks. Known addresses in the Tor network, including those of known darknet marketplaces, are automatically and continuously crawled to gather intelligence on specific threat objects being illegally traded and on trends of such trades. Based on this information, border guards may change the allocation of resources or what they look for as part of the border checks procedures. This situational risk assessment will change over time as the gathered intelligence on trends and threat items change. Furthermore, the gathered intelligence may allow a more specific assessment of *where* a certain heightened risk may become manifest, e.g. on certain land border crossing points or for flights arriving from a certain geographic region.[8]

One of the main ethical advantages of this variant of risk-based border management is that it is not dependent on processing data on identifiable individuals. Instead, the risk assessment depends on situational factors that will affect larger groups of travellers depending on the where and when they cross the border or on what flight they travel. As with all forms of border checks, implementing situational risk-based checks will imply positive, but also negative forms of impact. Since the aim of this section is to better understand the modes of negative unintended impact specific to data-driven, risk-based border management, the following sub-sections will analyse such aspects in relation to the categories and types of impact identified in deliverable D9.6.

3.2.1.1. ELSA category A: Privacy and data protection

With regard to potential privacy impact, deliverable D9.6's typology highlights new forms of shielding mechanisms for digital data. Such forms include privacy settings in apps and for social media platforms, the use of encryption technologies and password protection, as well as keeping certain data offline on private devices. This "digital shielding" creates protected "virtual spaces", which help individuals to have some form of control over what information is accessible about them from a public vantage point.

A technical distinction has to be made between the much larger "deep web" and the more specific "darknet" that is relevant for TRESSPASS's enabling technologies. The deep web refers to large parts of the internet that are *not* indexed through common search engines like

[8] In aviation security, a similar approach was used to classify flights according to risk criteria as part of the Dutch initiative "SURE!" (van de Wetering 2014).

Google, Bing, Yahoo or Baidu (unlike the so-called "clear web"). Among others, this includes password-protected sites (ranging from personal banking pages to web fora exclusive to club members etc.) as well as websites that are not included in the search engines' index, although they are technically accessible publicly. In the latter case, it may be necessary to *know an address* for this website in advance, as it will not be listed by search engines or other forms of web directories. Cloud services often allow sharing photos and other files *exclusively with a previously known group of people* through this form of digital shielding: a link shared by email allows access to this virtual private space only to those who received it. Technically, however, anyone can gain access if they know where to look for it.

While the darknet is part of the deep web, i.e. while it is not indexed by common search engines, its primary purpose is *not* to create virtual private spaces, but *anonymity*. On a technical level, this is achieved through applications like Tor or I2P, which create several layers of encryption that allow the communication of information from one computer to another (passed along by additional, mediating computers) without either machine knowing the "real" or "unencrypted" network address that would normally reveal identifying information or geographical location. Because anonymity networks like Tor do not offer "encrypted" addresses that are easily readable for humans,[9] the darknet makes it necessary to know the "encrypted" address of a particular service. Thus, the darknet can be used to create virtual private spaces, which are also anonymous (although the address needs still to be communicated in some form at an earlier point).

Darknet marketplaces, on the other hand, are not meant to be private at all; they depend on being publicly known and accessible. Hence, the encrypted addresses of such websites are listed in directories or in the clear web. A list of known addresses in the Tor network can then be indexed to allow some form of darknet search. Because of the way the Tor network handles addresses, however, it is not feasible to index (or "crawl") the entire darknet by going through each possible address; instead, it is necessary to start from Tor sites that have actually been made publicly known and then include further addresses that may be included, there.

Intrusion into spatial privacy

For an ethical assessment of spatial privacy with regard to the enabling technology of the dark web crawler, it is therefore important to assess how severely virtual private spaces are affected by the collection and analysis of data in a specific design of a BCP. A first important factor to consider is whether only darknet sites are included in the analysis, and how sites are added to the list of addresses to be crawled. Furthermore, it is important to assess how information on these sites is indexed – is only information stored that is pertinent to specific security goals of risk-based border management, e.g. by looking for certain keywords, or is the site indexed in its entirety? If so, how long is the data retained?

Unintended negative effects for spatial privacy could, then, be reduced or minimized by limiting the crawling to publicly known websites on the darknet; by further limiting the index to marketplaces and other such places relevant for the *specific security goals* of risk-based border management; by further limiting the collection and analysis through the use of keywords; and by only keeping non-aggregated, fully indexed data for as long as they are available on the darknet.

[9] For example, the Tor address http://sq4lecqyx4izcpkp.onion/ leads to the electronic mailbox of "heise Investigativ". It gives users the technical means to tip the journalists off through reliable means of anonymization.

Intrusion into bodily privacy

For an ethical assessment of bodily privacy, on the other hand, the relevant question is whether information on natural persons' bodies will inadvertently be gathered by indexing known dark web sites. Since the dark web is also used to trade illegally obtained information or images related to (sexual) violence, it is not unlikely that private information especially of the victims of such crimes could be inadvertently collected if corresponding trading sites are indexed.

Again, such additional negative effects for victims' bodily privacy could be reduced by limiting the crawling to specific websites relevant to the security goals of RBBM. This may at least in part be facilitated by administrators of darknet marketplaces, as many of them seem to make an effort to keep trading of at least some of the worst kinds of criminal activity (especially regarding child abuse) off these websites.[10]

Intrusion into private life

For an ethical assessment of the impact on persons private life, a very relevant question is whether this form of web crawling may chill *legitimate use of the dark web* and, thereby, of legitimate and democratically desired activities in general. There is a variety of legitimate reasons why persons may want to host or exchange information in an (almost fully) anonymous fashion. Journalists corresponding with sources may be one aspect – although it is likely that such exchanges will be additionally protected and not publicly accessible. Another aspect may be journalistic reporting or information provided by an NGO that should be accessible by users anonymously and hard to "delete" from the web (e.g. by shutting down the servers where the information is hosted). This is especially relevant for crisis regions and oppressive contexts (either oppression by the state or where the state cannot provide protection against another group's oppressive activities. Since information gathered by a darknet crawler operated in Europe may, in the end, gather and index information from any region in the world, this may pose a problem if there are efforts to analyse this (public) information for clues that may lead to identifying information (e.g. uncommon pseudonyms used not only in the dark web but also in the clear web, etc.).

As before, such chilling effects that may lead to negative impact on persons' private life could be reduced or minimized by limiting the crawling to specific websites and keywords relevant to the security goals of RBBM.

Disclosure of private information

Regarding the disclosure of private information outside of the LEAs, we currently do not consider it likely that the use of enabling technologies like the dark web crawler as part of situational risk-based screening creates a heightened potential for impact. The reason for this is that all information collected by the dark web crawler is *per definition* already publicly available on the dark web. Furthermore, as part of situational risk-based screening, it is used for the purpose of assessing circumstantial risk factors, but not for identifying specific individuals.

Negative impact may become more likely if a lot of information is collected and systematically searched for information that can be linked with identifying information. In case this would

[10] Journalistic articles like Greenberg (2015; 2014) outline some of the differences between bigger dark net market places, including aspects like banning child pornography or limiting illegal trade to what they consider "victimless crimes".

reveal the identity of someone not relevant to border checks (e.g. inadvertently identifying a dissident blogger from an oppressive regime), a negative impact through disclosure of information is possible. The functionality of linking such information from the dark web to real identities, however, is not part of situational risk-based border checks but should be addressed as part of traveller differentiation based on risk profiles.

3.2.1.2. ELSA category B: Unfair distribution of impact across different social groups

Disproportionate impact due to infeasibility of standard checks

For situational risk-based border checks, few disproportionate impacts due to infeasibility of standard checks is expected by enabling technology like the dark web crawler. This is due to the fact that this technology does not aim at gathering information on specific persons, but rather on more general situational aspects regarding identified risks, such as the potential smuggling of certain threat items. Hence, we did not find plausible scenarios where the use of the dark web crawler or similar technology would render the checking procedures infeasible for certain types of travellers (even if identifying information was incidentally collected).

This does not mean, however, that such negative effects need not be assessed for implementations of situational risk-based border management. On the system-level of a specific BCP design, such impact may still very well exist (e.g. because travellers on flights that are categorized as higher risk are then subjected to procedures that are infeasible for travellers with reduced mobility). Rather, it means that we do not foresee that such impact is exacerbated *due to the use of enabling technology* like the dark web crawler when compared to current practices.[11]

Disproportionate impact due to accumulation of false alarms

With regard to disproportionate impact due to the accumulation of false alarms, the use of enabling technologies like the dark web crawler becomes relevant. This is because the risk categories created as part of the situational risk-based border checks may include broad risk indicators that accumulate the privacy impact for large groups of passengers, potentially over a long time. For example, if situational risk data suggests that there is a higher risk of smuggling amongst the travellers on flights departing from a specific third country A, a very broad reaction could be to tighten checks for contraband for all of these travellers. Naturally, citizens of country A would be much more affected by this, as they may travel more often from that country into the EU.

Hence, it is important to not assume the relationship between gathering risk data, creating risk categorizations and implementation of new checking measures to be a direct, almost automatic one. Instead, as with all border checking procedures, proportionality is to be considered by a human in the loop and a special focus should be paid on whether this accumulates costs in privacy, loss of time, etc. for certain groups specifically. This, however, again is an effect that will need to be treated on the system-level of implementation, not at the level of technology or variant of RBBM.

[11] This implies also that, from a value sensitive technology development perspective, we do not expect a substantial contribution to disproportionate impact due to infeasibility of standard checks to be a relevant issue regarding the dark web crawler technology.

Disproportionate impact due to false or incomplete external data

The focus of disproportionate impact due to false or incomplete external data is on the use of external data related to identifiable individuals (e.g. the use of databases on lost and stolen travel documents). However, since enabling technology for situational risk-based screening (like the dark web crawler) does not rely on personal but only circumstantial data, there is no need to collect or use such data from external sources.[12] Hence, recurrent and potentially discriminating effects of classifying specific travellers are unlikely to result from such enabling technology.

However, situational risk assessments may also make use of "external data", e.g. when situational risk data comes from intelligence sources external to the border checking procedures. From this perspective, it can be a matter of definition whether the use of technology like the dark web crawler is part of the checking procedure itself or whether it is considered external input. In the former case, recurrent and potentially discriminating false positives would be considered false alarms; in the latter, they would be considered false or incomplete external input and fall under this type of negative ethical impact.

When considering the use of the dark web crawler external input, then, just as for the previous type of impact, it is important to not assume the relationship between situational risk data gathering, risk categorization and implementation of checking measures to be a direct, almost automatic one. Otherwise, classifying e.g. travellers on flights from certain countries as higher risk for longer periods of time may lead to disproportionate effects, in this case for groups of a certain nationality. Hence, a human should be involved in a meaningful way and special attention should be paid to whether this accumulates costs in privacy, loss of time, etc. for certain groups or recurrently for individual travellers.

Impact on non-travellers

With regard to negative impact on non-travellers, it should be clear that any privacy impact resulting from collecting and processing personal data as part of the situational risk data may indiscriminately also affect persons who do not intend to travel and cross the border at all. As long as only non-personal data is collected, no such impact will be present; hence, the impact on non-travellers directly corresponds to the different types of privacy and data protection impact. It is, therefore, important to minimize such impact as far as possible so as to also reduce negative impact on non-travellers.

3.2.1.3. ELSA category C: Restrictions of societal freedoms and liberties

Accosting travellers

Situational risk-based border checks may very well have a positive impact with regard to accosting travellers in border checking procedures (i.e. making crossing the border subject to intense privacy and time costs for travellers). This is true as long as the situational risk classifications as "low risk" outweigh classifications as "high risk". Since added security measures have a tendency to be kept in place indeterminately, it should hence be ensured that situational risk-based border checks result at least in a shift or better yet in a reduction of the overall privacy and time impact on travellers due to border checks. This is especially

[12] If, incidentally, identifying information is collected as part of the dark web crawling, it is still unlikely that this would result in a disproportionate impact for certain groups of travellers, for as long as only situational risk-based border checks take place. Using such incidentally collected identifying information to target specific travellers would be a case of traveller differentiation based on risk profiles (cf. section 3.2.2).

true if intensified checks affect travellers that should only be subjected to minimal checks (e.g. Union citizens and persons permanent residency status).

Enabling technology like the dark web crawler should, therefore, also address situational data that relates to "low risk" classifications; on the system-level of implementation at a BCP, this should allow for designs that result in less overall privacy and time impact for travellers, especially if they should only be subjected to minimal checks in the first place.

Lack of accountability

Additional impact due to a lack of accountability may result from situational risk-based border checks if no reason is given to the travellers for being subjected to additional checks. In this case, travellers would be considerably hindered in seeking judicial redress and holding officials accountable in case of disproportionate impact.

This means that, in order to minimize such impact for the enabling technology like the dark web crawler, it should not be necessary to keep it secret that such technology is used and that a passenger was classified "high risk" due to matching certain situational (non-personal) circumstances – e.g. due to flight used or BCP chosen. The more information can be given to the traveller without compromising the security effect of the additional checks, the lower will be the lack of accountability as traveller will be able to effectively seek redress if they believe the situational risk classification affects them disproportionately.

Restriction of self-determination and misuse of data

With regard to the restriction of self-determination and misuse of data, the impact will directly correspond to the amount of personal data that is (inadvertently) collected, processed and stored as part of the situational analysis. The impact can, hence, be minimized by also minimizing types of privacy impact and keeping storage times as low as possible.

Lack of transparency

Enabling technology for situational risk-based screening may intensify impact due to lack of transparency if it is problematic to declare and discuss openly that such technology is used at BCPs and what positive and negative effects are to be expected due to its use. While full transparency in the technologies may be impossible to achieve without also undermining their security effects, it should be possible for the traveling public to understand what costs and benefits are to be expected, how it may affect them at the border and how the risk classifications work on a general level.

3.2.2. Enabling traveller differentiation based on risk profiles

In the second variant of risk-based traveller differentiation (cf. textbox on page 25), checks will vary in intensity based on a risk analysis of "background information" related to specific travellers. While the term "profiling" is often avoided in this context (due to being closely associated with racial profiling), it is meant here as a general description of using personal data to classify travellers according to certain risk criteria. Depending on the specific security goal to be reached, the risk categories may be aimed to identify travellers for example with a high and low risk of smuggling drugs or of being involved in terrorist activities. The main assumption behind this form of risk-based traveller differentiation is that travellers who are in fact involved in such activities are likely to be part of a specifiable group that can either be effectively defined by intelligence services or by statistical calculations (Adey 2004).

In order to enable traveller differentiation based on risk profiles, it is necessary to reliably establish the identity of all travellers so that they can be checked according to the assigned

risk classification. While it is mandatory for border crossings to establish the identity of all travellers, this process has to be especially robust for this form of risk-based border checks, as subsequent checks may be too lax if it is possible to use another identity so as to be classified as low risk. To support robust traveller identification, a registration process may be foreseen that may be supported by a corresponding app (Traveller Registration Application). TRESSPASS makes use of a range of existing, off the shelf technologies (e.g. fingerprint scanners) to support a robust identification and, as part of its development efforts, adapts one technology that is meant to prevent the spoofing of facial recognition systems (Thermal Counter Spoofing Sensor).

Apart from such identity focused technologies, TRESSPASS also adapts an open web analysis tool to enable traveller differentiation based on risk profiles as part of the Web Intelligence module. As opposed to the dark web crawler, this tool will analyse information on a given identity that is *publicly accessible for anyone on the internet* (clear web). Here, background information is provided for risk assessment as part of the screening process itself. Other forms of risk-based checks currently performed in the US make use of *external risk data*, i.e. the risk classification of travellers is performed externally to the checking procedures, for example by intelligence agencies, and only a matching of identities is performed as part of the checks itself (Weydner-Volkmann 2017).

In some of the national political debates in Europe around aviation security, differentiated checks based on risk profiles have been met with scepticism and outright rejection. In Germany, for example, the political discussion on a possible introduction of this kind of differentiated checks terminated quickly after prominent politicians suggested a parallel between differentiated screening based on risk profiles and the selection programs in Nazi Germany (Weydner-Volkmann 2017). Risk profiling of traveller should, therefore, be considered a highly sensitive political topic – and depending on the technologies used and processes in place, the ethical, legal and societal repercussions can be prohibitively severe (Weydner-Volkmann 2017). The following sub-sections focus on the technologies developed in TRESSPASS, but also aim at a broader understanding of these repercussions in relation to the categories and types of impact identified in deliverable D9.6.

3.2.2.1. ELSA category A: Privacy and data protection

Intrusion into spatial privacy

For an ethical assessment of spatial privacy with regard to passenger differentiation based on risk profiles, it is important to assess how virtual private spaces are affected. This, however, is highly dependent on the kind of technical solution used. While technical tools exist that collect, store and analyse data also from deep web sources (cf. section 3.2.1), TRESSPASS's open web analysis tool is meant to only access public information in the clear web. Information that is redacted (e.g. via a privacy setting in social media), shared only among a set of people (e.g. social media friends) or simply not indexed to prevent unfettered access from anyone is not used.

Since spatial privacy, ultimately, relates to the travellers' choices of what information should be known about them and visible to the public, another technical aspect is very relevant: Is it possible for the traveller to decide, at a later point in time, to redact information that is available on them? From a technical point of view, the question is whether (a) the information gathered from the clear web is stored for later analysis, and (b) whether stored or archived information is used that is not under the control of those travellers. Minimizing the impact on spatial privacy would mean that (a) only information that is currently still available in the clear web is analysed and that (b) only data from deliberately selected web services is analysed

which give travellers a meaningful way to redact such information. This would allow people to "curate" what can be part of the open web analysis and react accordingly if they wish to do so. This route is followed with the TRESSPASS open web analysis tool. Furthermore, it would minimize the impact of deliberate attacks from third parties that maliciously make information from virtual private spaces available to the public (so-called "doxing" attacks) or spread false information.[13]

Intrusion into bodily privacy

With regard to bodily privacy, it is important to assess how biometric technologies for robust identification affect travellers. As mentioned above, it is mandatory for border crossings to establish the identity of all travellers and it is hoped that biometric technologies will allow a more robust identification. However, collecting, processing and storing biometric information can create an impact on bodily privacy beyond the mere fact of being identified at the border. For example, storing biometric information can allow re-identification (i.e. tracking) for future border crossings if the data is stored or even for other contexts if the data is accessible for other purposes, as well.

In order to minimize such impact when using biometric technology, it is preferred to use token-based biometrics wherever possible. This means the identity of travellers should only be verified *against a token in the possession of the travellers* (typically a biometric passport) whenever possible; no biometric template should be stored after the traveller is cleared to cross the border. This is also true for technology that detects forms of biometric tempering, like masks or faked fingerprints: Minimizing the impact on bodily privacy means avoiding the collection, processing and especially the storing of biometric information, wherever possible. The Thermal Counter Spoofing Sensor developed/adapted in TRESSPASS meets this requirement as it does not collect or process biometric information.

Intrusion into private life

For assessing the impact on travellers' private life, technologies that deal with an analysis of "background information" are central. This means that an ethical assessment needs to deal with how profiling technologies are affecting the distinct personality or character of travellers. This may relate to, among others, information on their familial, political, professional, religious, or sexual relations, their communications or activities (cf. Deliverable D9.6). While collecting information that has deliberately been made public should not be seen as an intrusion into privacy, this might be a higher bar to meet than it appears at first glance. Firstly, it is often unclear whether information has actually deliberately been made public – or just inadvertently. The latter could happen, for example, because default settings on social media or may entice people to share much more information much more widely than they actually want. This would imply that privacy aspects of processing such "public information" are much more problematic and, depending on the specifics of the data collection and the social media platform, it can mean that the privacy settings cannot be relied on. This is also true for information on travellers that is being made available not by the travellers themselves but by, for example, social media friends. In order to minimize the impact of using technologies like

[13] Spreading false information is, strictly speaking, not an impact on travellers' private spaces and it is debatable whether it is a privacy problem at all. Similarly, information may inadvertently be made public because, e.g. social media friends post information that *also* relate to a specific traveller. Again, I would argue that this is not an issue of spatial privacy, as the data collection and processing does not affect virtual private spaces. This does not mean, however, that does not pose a another type privacy problem (in this case, a private life impact).

the “open web analysis” as part of border checks, measures should be taken to minimize collecting and processing public information that has not deliberately been made public.

Similarly, the systematic compilation of information from different sources, especially if they have been posted under different pseudonyms, can reveal more information on travellers’ personality than originally intended. Hence, the creation of profiles on travellers’ character and dispositions goes much further than a simple use of publicly available information, especially when this is used as part of the border checks procedures to determine a traveller’s risk category. This could also lead to chilling effects, i.e. situations where travellers refrain from legitimate actions in their daily lives out of fear that they may have negative consequences later on (e.g. at the border). On the other hand, the privacy impact seems much smaller if open web analysis is used not to analyse a travellers character or personality, but to check the veracity of specific claims made by a traveller (in answer to legitimate questions asked by border guards), e.g. regarding the reason for travelling or regarding their means of living during their stay.

Disclosure of private information

Regarding a privacy impact through disclosure of private information, it is clear that the more biometric or person related information is stored and shared, the higher the risk of disclosure. Hence, in order to minimize this type of impact, technologies should avoid storing and further communicating information wherever possible. For example, the open web analysis technology developed in TRESSPASS is not meant to store any personal information. Instead, it offers a “live” analysis of currently available information.

3.2.2.2. ELSA category B: Unfair distribution of impact across different social groups

Disproportionate impact due to infeasibility of standard checks

Especially with regard to biometric identification, it is important to check if there is a disproportionate impact due to infeasibility of standard checks: For all existing biometric methods, there also exists a certain percentage of travellers that cannot have their identity verified biometrically – a fact that is often not considered sufficiently when thinking about the deployment of biometrics (Petermann and Sauter 2002). For such cases, alternative methods for establishing the identity of travellers must be in place. From an ethical perspective, it is important to ensure that the BCP as a whole offers such alternatives that are comparable with regard to the privacy and time impact on the traveller. This means that such an impact is expected to depend mostly on system-level aspects of a risk-based BCP. From a value-sensitive technology development perspective, however, it is important to keep the rates of false biometric rejection and failure to enrol in mind. This is also true for technology that is meant to make biometrical identification more robust (e.g. the Thermal Counter Spoofing Sensor), as it may change those rates depending on their functioning.

In certain situations, biometrics and supporting technology to improve robustness may also *mitigate* impact due to the infeasibility of standard checks: While veils worn in public for religious reasons will work for facial biometrics so long as the face is not covered, burkas and other veils that cover large portions of the face prevent facial biometrics from working. However, robust, automated biometrics may make it feasible for border checks to be more forthcoming than traditional methods if the technology operates to a level of robustness where separated “private” booths become viable. From a value-sensitive design perspective, this use case should, therefore, be kept in mind.

Disproportionate impact due to accumulation of false alarms

Disproportionate impact due to accumulation of false alarms is relevant both for biometric technologies and for analysing background information on travellers. As mentioned above, from a value-sensitive technology development perspective, it is important to keep the rates of false biometric rejection and failure to enrol in mind with regard to the development of biometric technology. Apart from keeping those rates as low as possible and offering, wherever necessary, viable alternatives, it should also be checked whether "false alarms" are generated disproportionately for salient groups of travellers. For example, AI-driven methods to biometrically screen facial images for known fugitives have been shown to disproportionally affect structurally disadvantaged social groups (ACLU 2018). During technology development, such effects should be tested and minimized, e.g. by using more representative, purpose tailored training data for machine learning. Similarly, technologies to increase the robustness of biometric identification should be tested if and to what degree it disproportionately generates false alarms for salient social groups or recurrently for the same travellers.

Apart from privacy and data protection aspects, one of the most prominent concerns regarding the use of risk profiles for differentiated checks of travellers are disproportionate alarms for salient social groups. Especially in cases where a (semi-)automated risk analysis takes place, the use of certain indicators can disproportionately affect members of social groups that are already structurally disadvantaged. In cases where AI-driven technologies are used for risk categorization, this is especially endemic, as it often remains unclear what data patterns resulted in a specific categorization. Since fully automated decision making is not foreseen in the legal framework (cf. deliverable D1.4; EU 2016b), it is central for technology designers to make decisions taken by humans in the loop as meaningful as possible. In case AI is used at central points during the risk analysis, developers should make an effort to incorporate explainable AI technologies. In any case, however, it does not suffice to check whether risk indicators are based on "discriminating information" (e.g. the use of data on religious or political beliefs), but it is also necessary to check if (semi-)automated risk classification *results* in disproportionate false alarms for salient social groups as well as for the same travellers over and over.[14] Such impact is much less likely to occur for open web analysis technology if it is used not to analyse a travellers character or dispositions, but to check the veracity of specific claims made by a traveller (as has also been mentioned with regard to the impact of travellers' private life). This is assumed because in such use cases, the scope of data processing on persons would be more targeted.

Disproportionate impact due to false or incomplete external data

Disproportionate impact due to false or incomplete external data relates to the use of external data related to identifiable individuals. Regarding risk-based border checks, one example is the use of externally provided risk profiles on travellers. Such risk profiles may be provided by law enforcement organizations and, in some countries, also by intelligence services. As opposed to a situation where risk profiles are established as part of the border checks procedures, it is not possible to assess in the ethics framework how externally provided risk data was generated and what kinds of technologies were used in this process. As mentioned

[14] From a value-sensitive design perspective, this entails that during the training of AI-driven technology in component development, misrepresentations and other factors that lead to discriminating effects should be minimized as much as possible in the training data.

above, during the border check itself, only matching of travellers' identities with lists may take place.

This also limits the possibilities for mitigation and minimization of disproportionate impact during technology design, but two possible avenues would be to (1) ensure through facilitating technology that false positives (i.e. traveller's being falsely identified as a person listed in the high risk category, e.g. because of a name similarity) are reduced as much as possible during the list matching. Additionally, such technology should allow for systematic tests to check whether known salient groups are affected disproportionately of being falsely identified with persons categorized as high risk. Furthermore, (2) technical solutions should prevent the same travellers from being misidentified over and over again, ideally without them having to give up additional personal data.

Another way to mitigate at least long-term effects of disproportionate impact due to external risk profiles would be to ensure that the grounds for being categorized as high risk are made available to the travellers. This would enable them to seek legal redress in case they feel they were subject to discriminating effects. From a technology development perspective, such a functionality should be supported, ideally without disclosing additional information to the border guards.[15] Similarly, in case a kind of "trusted traveller" category exists, it should be possible to communicate the grounds for denying this status to applying travellers.

Impact on non-travellers

With regard to the impact on non-travellers, it depends on how specific "background information" on travellers is collected, processed and stored. For example, biometric information should only be collected and processed for persons who intend to cross to the border, not from persons accompanying other travellers (e.g. in the airport). With regard to risk profiles, data collected, for example, through open web analysis should relate specifically to the travellers. This makes collecting or processing information on contacts and social media friends highly problematic since this affects non-travellers as well. In order to partly mitigate such impact, specific queries are preferable, e.g. whether specific social media accounts (e.g. of known drug traffickers) are among the list of contacts.

3.2.2.3. ELSA category C: Restrictions of societal freedoms and liberties

Accosting travellers

Three main aspects are relevant when we consider a potential impact on the freedom of movement by accosting travellers: (1) The overall time loss and (2) the overall privacy cost travellers are subjected to due to customs and border checks as well as (3) potential monetary costs for travellers to mitigate such restrictions of their freedom of movement. As mentioned in deliverable D9.6, however, there must be a factual and legal possibility to cross the border. With regard to border checks, we therefore need to make a distinction between travellers with EU citizenship (or a similar status like permanent residency), where this factual and legal possibility is already established; and third country nationals, where the legality of the border crossing may need to be established first.

[15] On the system level, i.e. from the perspective of designing a BCP as a whole, one could also try to mitigate such disproportionate impact by laying the burden of proof on those who provide the lists, i.e. it should be a requirement that agencies compiling such lists can demonstrate that and how they have taken measure to protect travellers from discriminating effects.

With regard to risk profiling and the use enabling technology like the open web analysis tool, the ethical impact could, therefore, be positive or negative, depending on whether travellers currently eligible to be subject only to minimal checks "in order to establish their identities" including a "rapid and straightforward verification" (EU 2016a, Art. 8) will now systematically be subject to higher privacy costs and time loss and on whether a larger number of third country nationals may become eligible for minimal checks.

As mentioned above, the use of the open web analysis tool in and of itself will likely create a certain privacy impact (especially regarding travellers' private life). Unless this impact can be minimized thoroughly and unless, at the same time, it seems plausible that through its use a large number of travellers may become eligible for minimal checks, an indiscriminate use for all travellers should be avoided. Even if this is the case, it must be considered that this implies an ethical and political trade-off: Union citizens (and persons of similar status) will be subject to a higher privacy impact so as to lower the number of third country nationals to be subject to more intensive checks. Hence, the use case of supporting targeted checks for specific travellers (e.g. to check claims made by the travellers as part of second line checks) will likely constitute less impact than the use case of making open web analysis part of initial risk profiling for all travellers.

Similarly, if the use of Thermal Counter Spoofing Sensor enables more robust biometric verification, thereby allowing more travellers to be checked faster (e.g. through automated, electronic gates) time loss may be considerably reduced for them (if enough electronic gates exist to actually allow an increased traveller flow). However, this also depends on the false alarm rate of the technology, i.e. it must be ensured that only in seldom cases travellers are falsely subjected to further checks due to the use of Thermal Counter Spoofing Sensor.

Open web analysis can also become relevant in the use case of registered, trusted traveller programs, where participating travellers may benefit from shorter waiting times and/or reduced overall privacy impact (in case the privacy impact of open web analysis is offset by relaxed intensity of other checks). However, if membership in these programs comes at a monetary cost, an additional societal impact is created: While the benefit of security checks at the border can be enjoyed by all, the connected burdens in time loss and/or privacy loss have to be carried by those who are not willing or not able to pay extra.[16] Hence, the starker the contrast in time loss and privacy impact between checks of paying members and checks of other travellers, the higher the societal impact.

Lack of accountability

Especially with regard to risk profiling, it is important for travellers to know what to expect as part of the checks and what rights they have in order to avoid a lack of accountability. As a guiding principle, it should be possible for the border guards to give an explanation of why a traveller has been categorized in a higher risk category or refused some kind of a trusted traveller status. This is especially pertinent, since much of what could be used as "background information" may not change over longer periods of time so that specific travellers are recurrently subject to more intense screening. From an ethical point of view, the minimum requirement for accountability is (1) that travellers know if they were subject to risk profiling and (2) that it is feasible to seek legal redress – which in turn implies that it must be possible during a lawsuit to check the specific reasons that lead to being classified as higher risk than other travellers.

[16] For a similar argument made in the context of aviation security, cf. (Sandel 2012).

For enabling technology like the open web analysis tool, a functionality should be foreseen that notifies travellers when their public information was screened by border checks. In case travellers are classified as higher risk due to open web analysis, it should be possible for them to learn the justification for this. As a minimum level of protection, this information must be accessible when seeking legal redress. Since this may imply storing even more personal information about a traveller, any form of access to the corresponding documentation should require some form of cooperation from the traveller (e.g. access to a redress code handed out to the passenger).

Restriction of self-determination and misuse of data

With regard to the restriction of self-determination and misuse of data, it has been argued in deliverable D9.6 that according to data protection legislation, it is necessary for the persons and entities that process personal data to demonstrate that the legal requirements have been met. Following the principle of data minimization, this means that enabling technologies should store as little personal information as possible. In fact, it is foreseen for the open web analysis tool in TRESSPASS that no personal information is stored whatsoever after the analysis is complete.

As mentioned above, it cannot be excluded that some personal information is processed as part of the analysis. For example, some of the publicly available information on social media may have inadvertently been made publicly accessible. Therefore, the principle of purpose binding should apply, i.e. technologies like open web analysis should not be used to determine a traveller's "risk" in a very broad and indeterminate sense. Instead, it should only be used with regard to specific security purposes, e.g. in order to check answers by a traveller that were given to questions related to a specific security goal laid down in law.[17]

Lack of transparency

In order to avoid a lack of transparency that may make governmental security measures opaquer and more restrictive, it should be possible to publicly describe the use of enabling technologies for traveller differentiation based on risk profiles without undermining the intended security goal. This means that it should be possible to provide publicly available information on what is happening during border checks, including on how risk profiles are created and used in the process. As stated in deliverable D9.6, this does not mean that very detailed information on the capabilities and detection rates of enabling technologies or on the specifics of risk indicators has to be given out, but it should be possible to describe on a high level that risk profiling is happening, e.g. using open web analysis of social media data (including the intended security goals and information on what information is stored for how long).

It should also be noted that many countries within the EU have far-reaching freedom of information laws. Thus, it may be not only ethically desirable but also legally required to make stored risk profiles on travellers available to them upon request. Access to such information may be limited, especially in case of an ongoing investigation, but in case risk profiles are stored for specific travellers, it may be necessary to store them in a fashion that meets the requirements of freedom of information laws without undermining the security goal. Since

[17] In case external risk profiles are used, only the broad suggestion can be given from a value sensitive design perspective that the use of such profiles may fuel a highly problematic demand for far reaching processing of personal data (Weydner-Volkmann 2017), especially if the risk profiles are meant to be provided by intelligence agencies that are subject to legal restrictions oversight to a considerably lesser degree than police and border guards.

the TRESSPASS open web analysis tool does not store any information for further use, we deem this requirement to be met.

3.2.3. Enabling traveller differentiation based on behavioural analysis

The third variant of risk-based traveller differentiation (cf. textbox on page 25) is based on behavioural data collected as part of the border checks procedures immediately before or during the checking process (e.g. in the arrivals terminal). This variant is based on the following psychological hypothesis: Attackers will unwittingly show certain behavioural peculiarities that can hardly be controlled by them (Weydner-Volkmann 2017). Trained security personnel or automated analysis can then make use of these peculiarities for risk classification. The fundamental idea behind this is that it is possible to detect 'bad intent' and use that as a basis for risk classification (US DHS 2013; US GAO 2013; Weinberger 2010). From an operational point of view, two ways to apply behavioural analysis can be distinguished: (1) analysing travellers' behaviour while they move towards and through the BCP, and (2) analysing travellers' behaviour as part of a longer interview situation, e.g. as part of second line checks.

Analysing travellers on the move by trained security personnel had been done in aviation security for some time. In the US, it was implemented by the TSA as part of the 'Screening Passengers by Observation Techniques' (SPOT) program. Here, so-called behavioural detection officers engage passengers in brief interactions and look for a predefined set of 'behavioural cues'[18] that are said to indicate elevated levels of stress, fear or the intention to deceive (US GAO 2013). From the flow of passengers, some are then selected for additional screening measures. Similarly, passengers can also be categorized as low risk, when their behaviour is assessed accordingly.[19]

For the time being, fully automated analysis of said behavioural cues of travellers on the move is technically not feasible. However, it is often also assumed that *mala fide* travellers will show unusual movement patterns as they walk through the BCP area. While it may be infeasible to link a specific risk indicator to specific movement patterns (like stalling or loitering), tracking all travellers' movements may provide enough data to use machine learning algorithms that find *uncommon* movement patterns (anomaly detection based on location tracking).[20]

In TRESSPASS, several enabling technologies are developed for tracking the location of travellers over time within the immediate border crossing area and adjacent facilities in order to establish movement patterns. A first technology, the Video Tracking Component (VTC), makes use of CCTV video surveillance technology; the aim is to recognise trajectories for each person in the video and find better ways to re-identify persons as they move from one camera to the next (multi target multi camera tracking). In this way, several cameras can be used together to analyse the movement patterns of travellers. A second technology may rely on

[18] This predefined set of of cues is based on a theory by Paul Ekman, according to which inadvertent facial micro expressions can be used in interactions to detect emotions and deception.

[19] The scientific basis for these programs has been fundamentally put into question in the past, as the empirical evidence does not seem to support that such criteria are effective. Instead, some studies suggest that the probability of detecting deception is only little higher than pure chance (Ormerod and Dando 2015; Weinberger 2010; US GAO 2013).

[20] It is unclear at this point how unusual movement patterns can be linked to *mala fide* travellers as part of risk-based border checks. Especially in border crossing situations, the movement of travellers through the facilities and across the border may often be rather uniform, as there may be little else one can do or where one can go. Furthermore, it must be assumed that *male fide* travellers may learn that their location is being tracked for the purpose of risk analysis. Hence, behaviour that is easy to influence on a conscious level (like the movements through a border crossing area) may be adapted accordingly by male fide travellers.

location tracking on the basis of a smartphone app, which is installed by travellers on their smartphones. This requires the cooperation and consent of travellers, which means that it has to be made transparent exactly what kind of data the app collects and whether or not it is used for security related risk assessment (and how).[21]

Due to this, the app may be part of a registered traveller bonus program and offer some added value; for the security component, on the other hand, a continuous collection of the location data in the vicinity of the BCP is central. This technology can also be used in combination with a third technology developed in TRESSPASS: RFID based traveller and luggage tracking (Travellers and Luggage Tracking Sensor Platform). Based on this, the location of luggage can be tracked throughout the border crossing area, e.g. by attaching a sticker containing an RFID transponder.

With regard to the second form of behavioural analysis, i.e. longer interactions with travellers in interview situations, it has been shown that this can reveal whether someone is trying to deceive the interviewer with a fair amount of reliability. In aviation security, such structured interviews have been used in Israel for some years now, where such interactions take between some minutes and several hours (Wagner 2014). This form of behavioural analysis is based on the hypothesis that in order to maintain a lie, we have to spend more cognitive effort than when answering truthfully. As a result of this, the amount of detail in the interviewee's answers will vary significantly depending on whether he or she is trying to maintain a lie or telling the truth. A publicly available study supports the hypothesis that this technique may indeed help to detect persons who are trying to deceive the interviewer (Ormerod and Dando 2015). From a border management perspective, this form of behavioural analysis therefore offers clear advantages, but due to the long interaction times, it only seems plausible to apply such a method selectively as part of the second line checks, not as part of an initial risk classification for all travellers.

Such deception detection techniques are currently infeasible to automate and will, at least for the foreseeable future, need to be conducted by trained and experienced humans. However, a supporting technology is developed in TRESSPASS that is meant to give improved feedback to the interviewer regarding the level of stress and cognitive load induced as part of the interview (Multi Modal Communication Analysis Tool, MMCAT). Since the level of detail in travellers' answers will vary from person to person, this may help to better establish a baseline under cognitive load for each interviewee and may allow a better assessment of signs of deception by interviewers.

3.2.3.1. ELSA category A: Privacy and data protection

Intrusion into spatial privacy

Due to the nature of the two main variants of behavioural analysis, we currently do not foresee direct forms of impact with regard to spatial privacy. Instead, we assume, that behavioural analysis may lead to a change in the distribution of bag searches, depending on how it is implemented.

However, even though the TRESSPASS traveller mobile app is being designed in a way that does not affect virtual private spaces, it may be difficult to convince travellers of this. Hence,

[21] At this point in the project, the corresponding Traveller Companion App is not foreseen to collect data for traveller risk assessment.

traveller perception should be taken into account and it doesn't seem prudent to make it mandatory for them to install the app on their smartphone.

Intrusion into bodily privacy

With regard to bodily privacy, two of the aspects mentioned in deliverable D9.6 are relevant with regard to the tracking of travellers' movements. First, it seems plausible that, by analysing movement patterns, one can infer a range of medical conditions, either pertinent to *how* travellers move through the BCP area (e.g. passengers with reduced mobility) or *where* they go to and where they linger (e.g. frequency of using the restrooms). This issue affects any form of location tracking as an enabling technology, but it may become especially relevant when AI-based anomaly detection tends to flag *uncommon movement patterns that result from medical conditions*. For now, it is unclear in how far medical conditions will be flagged as suspicious and in how far this may reveal non-obvious medical information to border guards. In any case, it should be avoided to store or communicate location tracking information and resulting movement patterns that can be linked to specific travellers.

This also links to the second relevant aspect, the ability to re-identify a traveller on a future trip or in another place. With regard to the Video Tracking Component (VTC) developed in TRESSPASS, this raises the question of the robustness in the ability to re-identify travellers over time and space. In case markers used to re-identify travellers in the images are likely to change fairly quickly (e.g. type and colour of clothing, current location and distance to next camera), the impact will be smaller than if more robust techniques are used, such as gait detection or facial biometrics. Hence, the relevant impact depends on the person re-identification capabilities across different application contexts, i.e. whether travellers could be re-identified from other video feeds and when wearing different clothing, assuming the generated tracklets and other identifying information were stored. In this case, information collected should be treated very similarly to collecting and processing facial biometric information. On the other hand, technical limitations towards the re-identification capability in other temporal or locational contexts will help reduce this impact. In any case, a timely deletion of identifying information will greatly reduce potential bodily privacy impact, although not to zero.

Regarding interview situations, it may be assumed that close observation of bodily reactions may, in some instances, reveal medical aspects about the interviewee. It is unclear in how far the use of technology like the interview support tool is likely to increase this issue. In any case, due to the potentially sensitive nature of the data collection, recording and storing such additional information from interviews should be considered a sensitive data collection; appropriate measures should be in place to prevent that such technologically enhanced interview recordings become part of a file or dossier on a suspect or be communicated to other persons. Data collection and processing should be strictly bound to the purpose of supporting specific interview situations. Using such data to improve the respective algorithms should only be possible upon explicit informed consent.

Intrusion into private life

With regard to the intrusion into private life, it can be argued that systematic tracking and analysis of travellers' movements produce so-called chilling effects (cf. Solove 2009, 178), i.e. travellers may refrain from certain legitimate actions out of fear of being flagged as high risk travellers. This affects the VTC, RFID based tracking as well as the tracking through the Traveller Mobile App. Such impact is likely to be exacerbated if the respective data is stored once the border checks have concluded and the traveller has crossed the border. Hence, the privacy impact of tracking can be reduced if the tracking data is not linked to a traveller's real

identity as part of the automatic processing and if it is deleted once the border checks have concluded.

In order to further limit the impact, data collection can be restricted to certain areas within the border crossing facilities. For example, it may be refrained from tracking the traveller's movements in the vicinity of shops, restaurants, restrooms and nursery areas, as this may reveal additional information on their personality and life choices. Due to the sensitive nature of this, tracking travellers to and within praying areas or other areas offered for recluse should be considered highly problematic with regard to privacy protection. In any case, the privacy impact of tracking increases with the area covered and it must be clearly communicated to the travellers, which areas are under surveillance.

A further aspect that has a considerable influence on privacy impact is whether or not a correlation of travellers' movements with that of other travellers takes place. This may be done in order to learn about whether certain travellers know and communicate with each other or whether they actively try to stay close to each other but avoid any open contact. While such information may be a helpful risk indicator for some specific threats, a systematic collection and analysis of social relationships in a public area reaches deep into the travellers' privacy. While strong pseudonymisation and data deletion policies may help to decrease this impact, the privacy impact will remain considerable even under such circumstances.

With regard to the use of MMCAT, we do not consider the impact on travellers' private life to be substantially higher during interview situations than in a conventional setting. While the interviews must be seen as having a high impact on travellers' privacy due to far-reaching questions being asked about their private situation, reasons for travelling, economic situation, etc., we believe that the use of MMCAT will make it more difficult to deceive the interviewer, but it will not create additional impact.[22]

However, this is only true for as long as no additional data is stored for further use due to the use of MMCAT, i.e. as long it remains restricted to purely supportive tasks during the interview itself. While video and/or audio data may already be collected for evidence collection, collecting further data on bodily reactions for the purpose of revealing an intent to deceive will also create additional impact. Due to the nature of the interview situation, it must be assumed that such data obtained under stress and in an interrogatory setting may reveal information on emotional states or character traits not relevant to security purposes. Storing such information for further security purposes, e.g. as part of a dossier or file on travellers classified as high risk, creates a considerable impact on the affected persons' privacy.[23]

Disclosure of private information

Regarding negative impact due to disclosure of private information, it should be ensured that travellers have no opportunity to access tracking information on other travellers. Hence,

[22] Any interview situation is designed to facilitate the detection of deception and interviewers are trained in conducting such interviews in a way that induces stress and cognitive load. The use of MMCAT may increase the effectiveness of this, but the main privacy impact remains dependent on the questions being asked by the interviewer. The use of MMCAT would gain a central ethical relevance, if one would assume a *right to deceive* the interviewer that would, by using supportive technologies, be hindered. In functioning liberal democracies, we do not believe that such a right should be considered during the analysis, but this argument points towards the research ethical issue of dual use: MMCAT may be used, e.g. by autocratic governments in support of undermining the legitimate right to refuse cooperation. This issue, however, is not in scope of this report.

[23] While this impact may be partially reduced by storing such information in a pseudonymous fashion (e.g. to further improve the technology), the highly personal nature of the information makes it unlikely that identifying information can be effectively scrubbed and treated as non-personal data.

especially when accessible for border guards through mobile devices, it must be prevented that travellers can use lost or forgotten devices or see information displayed on these devices' screens when used in public areas or in booths. On the other hand, since second line checks are likely to take place in private rooms, we do not foresee privacy impact through indiscrete use of technologies with regard to MMCAT.

3.2.3.2. ELSA category B: Unfair distribution of impact across different social groups

Disproportionate impact due to infeasibility of standard checks

With regard to behavioural detection based on tracking travellers' movements, e.g. by using RFID based tracking, the traveller mobile app or the VTC technologies, disproportionate impact due to infeasibility of standard checks may become an issue if they cannot feasibly be applied to travellers with different medical conditions (e.g. when using a wheelchair, or for travellers with reduced mobility). An inability to apply such behavioural detection may make it necessary to use other kinds of border check procedures. Depending on the design of the border checks procedures, this may produce a disproportionately higher impact as part of alternative border checks procedures.

Similarly, if any traveller should choose to use an unsupported smartphone platform or model or if they should choose not to use a smartphone in the first place, for whatever reason, this may result in a disproportionately higher impact as alternative screening measures may need to be used. Again, this will be dependent on the specifics of the design of the border checks procedures.

With regard to interview based behavioural detection, it may be infeasible to use supportive technologies like MMCAT due to medical reasons (e.g. an incompatibility with facial tics). Hence, a disproportionate impact may result from alternative measures to be used, but this depends on the specific design of the border checks procedures.

Disproportionate impact due to accumulation of false alarms

Disproportionate impact due to accumulation of false alarms may be the result of using behavioural detection based on tracking traveller's movements. Especially regarding certain medical conditions, it may still be feasible to apply technology like VTC, tracking via RFID or tracking based on the traveller mobile app. However, due to medical conditions causing uncommon movement patterns, a disproportionately higher amount of false alarms may be the result.

In regard to second line checks that make use of traveller interviews, it still needs to be carefully analysed whether technological support via MMCAT will have a higher false alarm rate in comparison to traditional second line checks for salient social groups. Relevant aspects to be considered are certain medical conditions that affect facial expressions, but it should also gender and ethnic aspects should also be taken into account.

Disproportionate impact due to false or incomplete external data

Since behavioural detection does not make use of external data, we do not foresee a heightened potential for disproportionate impact due to false or incomplete external data with regard to interview support technologies or the tracking of travellers' movements.

Impact on non-travellers

Impact on non-travellers may occur especially when behavioural detection is used in areas that are accessible to both, travellers and non-travellers. Technologies that require the active

cooperation of travellers (e.g. by installing and activating a smartphone app or by carrying an RFID tag on themselves or on their luggage) have an advantage here, as their use may not affect non-travellers. Similarly, if behavioural detection based on video surveillance, e.g. using the VTC technology) requires an active enrolment procedure and if no data is retained to re-identify persons on a different occasion, non-travellers may not be affected even in areas that make use of such technology. Otherwise, the use of tracking technology in combination with behavioural detection should be limited to areas that are only accessible to travellers, not to people accompanying them or picking them up (e.g. at the airport), so as to minimize such impact.

3.2.3.3. ELSA category C: Restrictions of societal freedoms and liberties

Accosting travellers

As mentioned above, three main aspects are relevant when we consider a potential impact on the freedom of movement by accosting travellers: (1) The overall time loss and (2) the overall privacy cost travellers are subjected to due to customs and border checks as well as (3) potential monetary costs for travellers to mitigate such restrictions of their freedom of movement.

With regard to behavioural analysis based on tracking of movements, potential enrolment processes may have a negative impact on overall time loss. This can true for video tracking like the VTC technologies, but also for RFID and traveller mobile app-based tracking. In all cases, an interaction between travellers and personnel or machines is prone to cause queues and waiting times unless a lot of space and equipment/personnel is available. In this respect, smartphone-based enrolment may constitute a way to reduce such delays since many travellers may enrol via the app well in advance and since additional equipment may not be necessary. However, even when this can reduce queuing time, time spent for enrolling via the smartphone app still constitutes a negative impact; it can be reduced by keeping the enrolment process short and straightforward and by ensuring minimal delays for travellers even if they enrol “last minute” before the border checks.

In the other hand, second line checks are likely to take a considerable amount of time. Hence, false alarms should be reduced as much as possible overall so that as few travellers as possible should be delayed. Regarding the use of supporting technologies like MMCAT, we do not see likely potential for additional delays.

For all technologies that enable behavioural analysis, it should be assumed that additional privacy impact will manifest for all travellers affected by it (see above). Currently, we assume risk-based border management will entail that *all travellers* will undergo behavioural analysis, irrespective of their status. This constitutes a considerable challenge, however: While ways can and should be pursued to offset this additional privacy impact, e.g. by relaxing existing checks for the large majority of travellers, it should also be considered that, under the border checks regulation, a large part of the travellers are EU citizens and travellers of similar status, who undergo only minimal checks at external borders. Whether it is feasible to sufficiently relax the intensity of border checks for third country nationals to offset the additional privacy impact remains to be seen.

Lastly, as argued in section 3.2.2 above, a potential monetary impact to offset the time loss or privacy impact may exist in the form of a trusted traveller program. Enabling technology for behavioural analysis based on tracking travellers’ movements may be used for this, if it allows either faster checks or a reduced privacy impact. This may also give travellers the choice between faster checks at the expense of less privacy or vice versa. In this case, the group of

travellers taking part such a program may benefit from a lesser "accosting travellers" impact. However, if membership in these programs comes at a monetary cost, an additional societal impact is created: While the benefit of security checks at the border can be enjoyed by all, the connected burdens in time loss and/or privacy loss have to be carried by those who are not willing or not able to pay extra. Hence, the starker the contrast in time loss and privacy impact between checks of paying members and checks of other travellers, the higher the societal impact.

Lack of accountability

As mentioned in section 3.2.2 above, it is important for travellers to know what to expect as part of the checks and what rights they have in order to avoid a lack of accountability. As a guiding principle, it should be possible for the border guards to give an explanation of why a traveller has been categorized in a higher risk category or refused some kind of a trusted traveller status. From an ethical point of view, the minimum requirement for accountability is (1) that travellers know if they were subject to behavioural analysis and (2) that it is feasible to seek legal redress in case of unreasonable impact – which in turn implies that it must be possible during a lawsuit to check the specific behavioural indicators that lead to being classified as higher risk than other travellers.

With regard to automated behavioural analysis based on tracking travellers' movements, this can sometimes be challenging. Since fully automated risk categorization is highly problematic from an accountability point of view, a human with sufficient understanding of the rationale of the specifics of the behavioural alert should be kept in the decision loop. Even when a border guard takes the final decision, providing sufficient accountability means that it is possible for him or her to articulate what specifically constitutes the risk that may or may not be assigned to a traveller. Alerts about uncommon movement patterns should, hence, be explainable to the travellers.

This is especially relevant in case false alarms may accumulate for individual travellers in order to enable them to seek legal redress; in such cases, a high level of accountability would mean that it is possible to provide some form of documentation of the risk categorization justifications.

The problematic scientific basis (BGH 1998; 2010) for forms of deception detection during interviews and police interrogations lead some European countries to prohibit using results generated by such technologies in criminal proceedings. One way to reduce the impact to the detriment of the accused could be to allow using such technology only to their benefit, i.e. to negate allegations (OLG Dresden 2013). Due to the fact that supporting technologies like MMCAT have similar issues, a similar use case (i.e. help disprove suspicion with regard to certain risk indicators) may be an option. This would address accountability issues when using body information during veracity assessment, which cannot be disproved or explained by the traveller.

Restriction of self-determination and misuse of data

With regard to the restriction of self-determination and misuse of data, the principles of data minimization and storage limitation imply that enabling technologies for behavioural analysis should collect and store as little personal information as possible. Hence, for behavioural analysis based on tracking travellers' movements, it should be ensured that no tracking information that can be linked to the identities of travellers is retained after the border checks have concluded and the travellers have left the BCP area. For VTC, this also implies that no

additional images from CCTV are being stored and that no tracklet or other re-identifying information is retained.

This is also important with regard to the principle of purpose binding: If the collection of behavioural data is being justified by the purpose of risk classification during border checks, there is no need in retaining such data after the checks have concluded. This also reduces the risk of "mission creep", i.e. using the data for further purposes once they are available and stored. This includes other forms of security provision, but also further training of algorithms in commercial (security) products. Retaining no behavioural data also reduces the risk of unauthorized access to such data. This is especially true for data collected during high stake interviews as part of second line checks, as supporting technologies like MMCAT will likely to collect highly sensitive information. Such information should only be retained as part of a legal investigation subject to corresponding data protection measures.

Lack of transparency

To avoid a lack of transparency, travellers should be actively informed about the use of behavioural analysis and enabling technologies like VTC, RFID tracking or smartphone tracking technology and some level of detail about its main technical functionality should be openly accessible to them. Hence, the availability of such information should not undermine the intended gains in security.

Similarly, information should be openly available about the use of supporting technologies for interviews in second line checks, such as MMCAT. Here, it should also be ensured that scientific scrutiny regarding deception detection is sought in order to strengthen the public debate around the use of such technology.

3.3. Preliminary findings on ethical trade-offs in traveller risk assessment

3.3.1. Risk assessment, traveller risk categorisation and ethical responsibility

The introduction of RBBM involves the implementation of different methods for traveller risk assessment. Since the outcome of the risk assessment will change the way in which travellers are checked at the border, their application implies ethical trade-offs that need to be analysed (cf. deliverable D9.6). For the most part, current concepts for risk assessment are data-driven, i.e. traveller risk categorisation takes place based on the collection and processing of data related to certain threat scenarios. Based on this data, risk assessment is expected to be indicative of the probability that one traveller is either a *mala fide* or *bona fide* traveller with respect to this threat.

However, it is important to understand that the probability element of the risk is not inherent to the travellers themselves; the risk assessment *doesn't uncover risk properties* of specifiable travellers. Instead, it starts from the knowledge that *some of the travellers* are in fact *mala fide* travellers. Different risk assessment methods then make use of data collection and processing to group travellers in such a way that *mala fide* travellers are expected to be a more likely part of one group (higher risk), and less likely part of another (lower risk). Hence, the outcome of risk assessment is the *assignment of artificially created risk categorisations*. A higher risk traveller is, in fact, still either a *bona fide* or *mala fide* traveller and the assigned risk is not a property of the traveller. It denotes the probability that if a border guard were to choose one traveller at random from the artificially created higher risk group, that traveller will more likely turn out to be a *mala fide* traveller than one chosen from the lower risk group.

Different risk assessment methods will create different groups and classify travellers differently. They can be used in combination so as to better attain the goals of border checks,

which may be measured in performance metrics such as the flow rate of travellers or improved protection against threats scenarios. However, since each of these risk assessment methods also comes with ethical trade-offs, deciding for or against the introduction of certain variants of risk-based border checks carries a high responsibility: It effectively constitutes a decision of how to create and apply risk categories for grouping travellers – not about "finding out" about existing risk properties of travellers. In order to appropriately address this responsibility for the decision makers, TRESSPASS treats the level of ethical, legal and societal compliance as an additional performance measure.

At the current stage, there is no broad consensus on how to differentiate different variants of implementing risk-based border checks. For now, it seems plausible that different forms of differentiation between risk assessment methods help highlighting different characteristics. For an analysis of the ethical, legal and societal impact of enabling technologies, we have differentiated risk assessment methods according to how they make use of personal data to perform the risk classifications: either (1) they only use purely situational data, or the make use of traveller related data, which can be related to (2) "background information" on the travellers, or to (3) the travellers behaviour during the approach and at the BCP (cf. section 3.2 above).

Of course, these variants for implementing RBBM are not mutually exclusive. It is likely that depending on the specific threats to be addressed, they will be mixed and risk data generated through different methods may even be fused to generate overall risk classifications. As shown in section 3.2 above, TRESSPASS brings technologies for each of the three variants to the table. It remains to be seen which of the technologies will be used in which pilots and which risk assessments methods will be applied in each case. Furthermore, it will also remain to be seen how enabling technologies and risk assessment methods interact with existing state of the art (SOTA) sensors and procedures at currently at place at the BCPs.

Due to this, for now, findings on the ethical trade-offs for each of the three variants of RBBM are only preliminary. We expect that results from the pilots will help us better understand how the introduction of risk assessment and traveller risk categorization will play out with respect to ELSAs. Currently, we expect that each of the three TRESSPASS pilots will perform differently with respect to ethical compliance. From a validation point of view, this is quite important so that it can be shown that TRESSPASS's ethical framework can reflect qualitative differences with respect to ethical concepts in each of the pilots' BCP design. At the same time, in order to demonstrate that the TRESSPASS tools and methods can be used to design privacy respecting BCPs, one of the designs should perform close to or better than current checkpoint with respect to ethical compliance.[24] In the following sections, we will discuss some preliminary findings on the ethical trade-offs of introducing each of the three variants of risk assessment and risk classification.

3.3.2. Preliminary findings on ethical trade-offs in situational risk-based checks

3.3.2.1. Privacy and data protection impact of situational risk-based checks

As argued in section 3.2.1 above, situational risk-based checks do not make use of traveller related information, which limits the privacy and data protection impact of this variant of checks considerably. Instead, additional information regarding the threat situation is used to

[24] We currently expect this to be the Polish case as high ethical compliance has been explicitly mentioned in the stakeholder requirements in D1.1 and since it allows us to make use of the feedback from the first pilot.

differentiate groups of travellers, e.g. based on where they started their journey, where they cross the border, what kind of luggage they have with them etc.

Enabling technologies like the dark web crawler developed in TRESSPASS are meant to collect such situational information; this technology indexes information from the dark web that is usually already presented in an anonymized fashion, but publicly available to anyone. While some privacy issues exist and should be taken into account, virtual private spaces are unlikely to be heavily affected under such conditions; similarly, we don't expect that other forms of privacy impact will increase considerably. Privacy and data protection related impact can be further reduced by limiting the indexing activities of the dark web crawler technology to sites that are *known* to be relevant for specifiable threats that are to be countered at the border (whitelisting e.g. market places for illegal trading of weapons or of stolen passports). Furthermore, if indexed information is retained only for a relatively short time, it can be ensured that information is only retained and used as long as it is still publicly available.

We expect that it is quite possible to offset the additional privacy impact if it is possible to achieve even a minor reduction in overall privacy impact, e.g. by reducing the intensity of checks for some third country nationals. On the other hand, situational risk-based border checks will only have positive effects if there is an organisational and political will to reduce the intensity of checks in some situations. Otherwise, raised intensity for checks in other situations cannot be offset by these reductions and they will layer on top of each other over time. Hence, there is a need for clear situational criteria not only for higher but also for lower risk.

3.3.2.2. Unfair distribution of impact for situational risk-based checks

Since situational risk-based screening doesn't make use of traveller related data, we did not find plausible scenarios where the corresponding traveller categorisation would result in disproportionate impact for some societal groups due to measures like dark web data collection being infeasible for them (cf. section 3.2.1.2). Similarly, since no external traveller related data is used, disproportionate impact due to false or incomplete external data is unlikely.

However, we foresee that it is plausible that false alarms will accumulate for some larger groups, e.g. based on nationality: Since situational risk indicators tend to be rather broad and corresponding criteria for risk categorisation will be in place for some time, travellers on certain routes will be more likely to be affected repeatedly than travellers on other routes. For example, if flights from a certain point of departure in a third country are considered higher risk than other flights, it is plausible to assume that citizens from that country will be affected more than others. Still, all travellers on that flight will then be affected similarly unless other criteria are added, e.g. as part of a fused risk assessment.

In order to address this issue, it needs to be ensured that checks need to remain proportional and do not accumulate too much for a broad group of passengers, especially if this group corresponds with a vulnerable societal group.

Similarly, situational risk-based border checks will like not only impact actual travellers: data collection and analysis, e.g. as enabled via the dark web crawler, will affect all persons alike. Hence, it is important to minimise the privacy and data protection impact generated by the dark web crawler as much as possible.

3.3.2.3. Societal restrictiveness of situational risk-based checks

More stringent checks due to situational risks have a tendency to be kept in place indeterminately. However, if enabling technologies is also used to identify lower risk

situations, it is plausible to assume that situational risk-based checks can result in a *shift or reduction* of over privacy and time impact on travellers. This seems plausible for as long as reductions can realistically offset potential additional checks on Union citizens and travellers.

By reducing the privacy impact as much as possible, this also reduces potential restrictive impact on travellers' self-determination and the risk of data misuse: Since situational risk-based checks are not dependent on non-traveller related information, only inadvertently collected and stored personal data create concerns in this respect. Minimising the "by-catch" as outlined in the privacy section helps to minimise this impact as well.

It also seems plausible that it can be openly communicated that situational risk-based checks are in place and that threat information is gathered from the dark web without undermining the security effect. Hence, it seems plausible that travellers can also be informed about reasons for being checked as higher risk. This would effectively address potential issues regarding transparency and accountability issues.

3.3.3. Preliminary findings on ethical trade-offs in risk profiling

3.3.3.1. Privacy and data protection impact of risk profiling

Depending on what kind of information is used for risk profiling, impact on virtual private spaces can be higher or lower. On the lower end of impact, we see for example TRESSPASS's open web analysis tool, which focuses on public information that is not stored – effectively allowing travellers to curate what is publicly known about them and used as "background information" on them. While this can reduce the impact, it does not fully mitigate it. Such additional impact could be plausibly compensated if, for example, fewer searches could be conducted on traveller's luggage or vehicles. On the more severe end of such impact would be accessing data on traveller's electronic devices, asking to unlock virtual private space or demanding passwords.[25] Due to the invasive nature of such searches, however, it seems implausible that fewer searches of other private spaces could compensate for such practices.

Furthermore, due to the importance of linking existing risk profiles with traveller identity, supportive biometric technology will likely be in place to allow a robust identification of travellers. Again, this can be done in more or less privacy friendly ways (token-based biometric verification vs. creation and storage of biometric templates), but some additional impact on bodily privacy is likely to be created for all travellers who currently don't need to have their identity checked biometrically during border crossings. On the other hand, if risk profiling will lead to fewer secondary checks of the travellers' body (e.g. as part of customs checks), it is plausible to assume that impact bodily privacy can be reduced overall.

It is in the nature of risk profiling of travellers to reveal at least some aspects of travellers' distinct personalities and about how they lead their lives. Otherwise, no data could be gathered that can provide threat-related information on aspects like a traveller's intent or capability. However, a first distinction has to be made regarding the question of whether the actual risk profiling *is part of the actual border checks procedures* or if the border checks *make use of risk data that is provided by another party (e.g. lists of trusted travellers and watch*

[25] While access to virtual private spaces can be justified (potentially also in a border checks situation) upon reasonable suspicion and as part of an official investigation, due to the invasive nature of such data access, random searches and searches without specifiable cause for suspicion must be seen as disproportionate and highly problematic. Consistent with this, corresponding activities that had been conducted by border guards in the US under a border search exception to the fourth amendment were recently ruled unconstitutional, unless a reasonable suspicion could be articulated (Brodkin 2019).

lists). In the latter case, the immediate impact on travellers' private lives at the BCP can in principle be limited, since border guards "only" need to match the travellers' identities to the names listed on the relevant lists in order to enforce a predetermined risk categorization.[26] Hence, in such cases, the impact on travellers' private lives depends on whether and how much additional information is provided to the border guards.

In the former case, however, risk profiling and categorization are part of the actual border checks procedures. Hence, in case such collection and processing of traveller related data are not voluntary (e.g. as part of a trusted traveller programme), it is likely to create a considerable impact if the goal is to provide sufficient and reliable information on risk-related aspects like intent or capability. A less severe variant of such impact could be the exclusive use of publicly available background information on a traveller so as to check some of their claims made (e.g. reason of travel). While the exclusive use of information *that has deliberately been made public by the traveller* would be preferable, determining if some of the information processed may have inadvertently been made public (e.g. some social media post) is notoriously difficult and will be error-prone – especially since compiling pieces of information from different sources can reveal additional information that was not meant to be public and is also prone to lead to false conclusions about a person's character. Hence, people may feel inhibited to take part in legitimate online or offline activities out of fear that information on this becomes available on the internet and leads to negative consequences at the border. In case risk profiling in conducted involuntarily on all travellers, it seems implausible to assume that privacy benefits for some travellers could compensate for mass intrusions without suspicion.

Storing any traveller related information also increases the risk of disclosure to unauthorized access. On the other hand, by accessing only information that is currently openly available (as done by the TRESSPASS open web analysis tool), this risk can be mitigated and some level of control can be given back to the travellers, as they could effectively "curate" what is known about them.

3.3.3.2. Unfair distribution of impact for risk profiling

The heavy reliance on biometrics makes risk profiling prone to cause disproportional impact on some groups of travellers, since for all biometric technologies, a certain percentage of the population exists that cannot be enrolled or verified consistently. Hence, risk assessment methods should foresee viable alternatives to robustly verify a traveller's identity so that disproportionate impact through longer delays, higher privacy impact or repeated false alarms can be avoided at the system level of a BCP. Similarly, the use of biometrics or counter spoofing technology should remain largely compatible with different forms of religious clothing or foresee viable alternatives in BCP design.

In case risk profiling is conducted as part of the actual border checks for all travellers, some of them are likely to be consistently categorised as "higher risk" or denied entry to a "lower risk" category based on some background information – especially since the same methods of risk assessment may be used at many BCPs. Here, it is paramount that risk assessment methods allow for individual cases of repeated false alarms to be mitigated, e.g. as part of a redress

[26] While this means that the BCP's impact on travellers' private lives could be limited, it also means that this privacy impact happens elsewhere. We try to address this fact as part of the ELSA type "restriction of self-determination and data misuse" that is part of the societal restrictiveness category (ELSA category C): As a kind of "customer" of such risk profiling, border checks create a demand for a large scale collection and processing of personal data; this demand may be satisfied by governmental (or private) actors like intelligence agencies, which may be subject to fewer legal restrictions and little effective oversight. We therefore do not see it as a viable option to opt for "externally provided" risk profiling in order to reduce a BCP's privacy impact.

system. Still, especially for AI-driven and semi-automated methods of risk profiling, it is to be expected that more false alarms will be produced for some societal groups of travellers than for others. This is likely to be dependent on the specifics of the risk indicators as well as the technologies used to collect the corresponding data. It is unlikely, however, that this effect can be fully mitigated. Apart from reducing the situations in which background information is used as part of risk-based checks, a somewhat counter-intuitive method to reduce this impact may involve semi-randomized alarms to ensure a fairer distribution of impact at least amongst the members of each of the two groups of travellers, Union citizens (including persons of similar status) and third country nationals.

On a similar note, in case of risk profiling based on externally provided watch lists or trusted traveller lists, experiences from aviation security suggest that, even when no sensitive data like religious or political conviction is used as part of the risk profiling, some salient groups are likely to be affected disproportionately by being falsely categorized as "higher risk". Again, it is unlikely that this effect can be fully mitigated, especially because for higher risk travellers to be reliably identified, it is necessary to subject all travellers to list matching and, possibly, enhanced biometric identification. Furthermore, consistent false matching of travellers with names on the watch lists proved to be a consistent problem in aviation security (Weydner-Volkmann 2017). Hence, it is likely that a similar redress system must be foreseen as part of the risk assessment methods so as to address individual cases of consistent false categorization due to list mismatching.

Depending on when during the journey the risk profiling takes place, it is quite possible that non-travellers are affected by the profiling as well. This is especially true for externally provided risk categorizations, either because of a large-scale surveillance program, but possibly also due to misidentifications (e.g. when a risk profile is conducted on someone else because of a name similarity to a traveller). Such impact on non-travellers could be minimized by collecting and processing background information only at the BCP and only based on social media handles provided by the travellers. It is unclear, however, if such restrictions wouldn't undermine the intended security goals.

3.3.3.3. Societal restrictiveness of risk profiling

One of the central hopes in relation to the introduction of risk profiling at borders is faster and less intensive checks for the majority of the travellers. Reducing the average intensity and amount of time that travellers have to reserve for potential checks would also reduce the impact type "accosting travellers".

However, it needs to be taken into account that applying risk profiling for all travellers can also have the opposite effect since the very large majority of Union citizens and travellers of similar status currently only undergo minimal checks. Hence, applying risk profiling to all travellers means that for each Union citizen and traveller of similar status that is selected for more intensive checks, at least one third country national should be included in the group subject only to minimal checks. In general, the possibility to include volunteering travellers in a (at best non-monetary) trusted-traveller programme that makes them eligible for minimal checks seems like a plausible way this could be implemented. Technology that may be needed to make (biometric) forms of identification more robust, then, also need to have a sufficiently low false alarm rate to allow for faster checks.

On the other hand, mandatory, far-reaching risk profiling for all travellers will considerably increase the privacy cost and could also create false alarms at a rate that severely impacts the flow of travellers. From an ethics perspective, even less intrusive technologies like the open

web analysis should, therefore, only be applied in cases where there are actual grounds for suspicion, not as part of a generalized risk classification measure.

From a transparency and accountability perspective, it must be possible to provide a certain level of detail about the risk profiling measures used. For travellers categorized as higher risk or denied classification as lower risk, it must be possible to provide an explanation that allows avenues to seek legal redress. This means, it must be possible for travellers to be informed about the grounds for classification and to hold border guards or other involved organizations accountable when they feel they have been subject to disproportionate or discriminatory measures, e.g. recurrent false alarms or a disproportionate impact due to dietary requirements[27].

This is also true for making use of external profiling data (e.g. through watch lists). In addition to that, while such measures entail a comparably low privacy and data protection impact at the BCP itself (cf. section 3.3.3.1 above), it creates a demand for collecting and processing sensitive personal data elsewhere, i.e. it creates a higher demand for surveillance activities (Weydner-Volkmann 2017). This is because added security of watch lists is directly dependent on the validity and completeness of these lists – and therefore on the validity and completeness of upstream intelligence gathering. By making the effectiveness of border checks at least in part dependent on the effectiveness of intelligence gathering, it is very likely that this also implies a heightened demand for more surveillance activity at large. In the past, as the global surveillance disclosures by Edward Snowden have revealed, sweeping surveillance activities have in part relied on the misuse of personal data, e.g. from electronic communications or non-public data from social media. Due to the level of cooperation between intelligence agencies in the Western world, it seems plausible that using external risk profiling as part of risk-based border management will also create a higher demand for sensitive and potentially illegal collection of personal data on as many travellers as possible.

3.3.4. Preliminary findings on ethical trade-offs in behavioural analysis

3.3.4.1. Privacy and data protection impact of behavioural analysis

As opposed to risk profiling, certain forms of behavioural detection promise to not make use of personal data but rely on outwardly observable behaviour instead. Brief interactions with trained security personnel (e.g. in the form of the US TSA program SPOT) in order to detect deceptive behaviour could be an example of this. However, as mentioned in section 3.2.3, the scientific basis for this is contended.

While it is currently infeasible to automate behavioural analysis of cues for deception, the detection of unusual movement patters promises to detect *mala fide* travellers. For this, travellers' approach to and movements within the BCP are tracked and uncommon patterns are detected. However, this may make it necessary to collect personal data since for any automated forms of behavioural detection alarms need to be relatable to specific travellers in some way. This is especially true for CCTV based tracking that makes use of biometric information to re-identify travellers across different cameras. Certain types of biometric information, like gait analysis, may make it technically feasible to re-identify a traveller in later border crossings or in border unrelated contexts. Depending on how robust such re-identification is, the impact on bodily privacy can be considerable; less robust re-identification

[27] Since dietary requests are collected as part of PNR data in aviation and since some religious beliefs involve dietary requirements, it is sometimes feared that using this information in risk assessment could lead to discriminatory effects (FRA 2011).

that only works if recordings are temporarily coherent from cameras that are located close to each other reduce this impact. Further minimization of impact should consist of very short retention times and a strong pseudonymisation of tracked identities, which should ideally never be processed in direct relation with their actual identity. This is also true for RFID and smartphone app[28] based tracking.

The detection of uncommon movement patterns or the correlation of movements between travellers based on unrestricted tracking throughout the BCP and adjacent areas is likely to reveal a range of sensitive information about medical conditions, personal relations, religious convictions and other private life circumstances: Uncommon movements or frequent visits to certain places like restrooms may be due to a range of medication conditions that could be revealed by flagging corresponding behaviour as suspicious; many bigger airports offer areas of seclusion, such as nursery areas and prayer rooms, but also social service facilities for travellers in a personal crisis. To avoid a strong privacy impact, it must be ensured that common medical conditions cannot be correlated with suspicious behaviour alarms (e.g. frequent visits to the restrooms). Furthermore, tracking should be restricted to as few areas as feasible so that it can be avoided that tracking can allow conclusions visits to places that reveal sensitive information.

In order to avoid unwanted disclosure of tracking information, it must be ensured that such information is never displayed visibly to other travellers, that tracking information cannot be correlated to travellers in a data breach and that tracking data cannot be accessed through security personnel's mobile devices in case of loss or theft.

Close observation of behaviour in interview situations may reveal medical conditions not openly apparent to interviewers. In general, the impact on travellers' privacy is expected to be very high in general, depending on the questions being asked by border guards. In order to mitigate even further privacy impact, no additional data should be stored for further analysis as part of using supportive technologies such as MMCAT.

3.3.4.2. Unfair distribution of impact for behavioural risk analysis

Forms of behavioural analysis that (semi-)automatically detect uncommon patterns of movement are prone to disproportionate impact on those who will show uncommon patterns due to medical, cultural or religious reasons. Such a disproportionate impact can manifest in different ways: Commonly, if the application of a detection method is known to be infeasible alternative screening procedures will be applied (e.g. since it is infeasible to use a metal detector portal on a person in a wheelchair, other methods for screening are foreseen in aviation security). Here, it must be ensured on the level of system design, that alternative methods foreseen in case of medical issues do not imply a systematically higher impact with regard to travellers' privacy and time loss.

Alternatively, if the application of detecting uncommon movement patterns is not considered infeasible, a disproportionate impact can manifest by causing a disproportionate amount of false alarms that trigger further, more intensive checks. Since forms detecting uncommon patterns (anomaly detection) are by definition rather broad in scope (as opposed to the detection of specific threats, e.g. the presence of trace explosives), efforts must be taken to minimize discriminating effects due to the accumulation of false alarms for some vulnerable groups of passengers. Due to the spectrum of uncommon behaviour caused or motivated by

[28] While the TRESSPASS traveller mobile app does not employ any form of covert data collection, it remains to be seen if travellers trust to install a border checks related app that collects security related information on them. Since this is an acceptance related issue, this is out of scope for this report and will be addressed as part of T6.3.

medical, religious or cultural factors, however, it is unlikely that discriminating effects can be reduced to zero.[29]

With regard to the use of trained officers for the detection of behavioural cues for deception in brief interactions, the reliance on opaque decision criteria have exacerbated discriminating effects in the past.[30] Apart from this, some religious and cultural factors may cause a much higher impact for some groups during interactions if the use of such detection methods requires travellers to fully reveal their faces during the interactions.[31]

With regard to the application of supportive technology for traveller interviews, again, it must be ensured that the technologies are well tested for (in)feasibility regarding the use with travellers who have certain medical conditions (e.g. facial tics), so as to avoid disproportional impact for those travellers during interview situations.

In order to avoid unnecessary impact on non-travellers (e.g. persons who accompany travellers to an airport), behavioural detection should be either restricted to areas that are only accessible to travellers who intend to cross an external border or require some form of active cooperation on behalf of the travellers (e.g. installation and registration of an app, biometric enrolment, etc.) that minimizes the possibility of non-travellers being inadvertently subjected to risk assessment.

3.3.4.3. Societal restrictiveness of behavioural analysis

In the past, officers trained for the detection of behavioural cues in brief interactions were not only used by the US TSA to classify aviation passengers as higher risk, but also to include passengers in a lower risk category (managed inclusion). At least for some passengers, this may allow for less impact regarding the "accosting travellers", i.e. the average impact in time, privacy and money (in case of trusted traveller programs) on the freedom of movement due to border checks. In case this is applied in a way that allows more travellers to be subject only to minimal checks, it is plausible to assume that even if such brief interactions are applied to all passengers, this may still allow an overall reduction of time loss and privacy impact (while remaining neutral with regard to potential monetary impact).

In principle, the same holds true for behavioural analysis with regard to the movements of travellers. This depends on the specific technologies and overall system design, however, as many of these technologies require an enrolment process. For the TRESSPASS RFID tracking and VTC technologies, we currently find it more plausible to expect additional delays due to the added enrolment procedures and connected queues that are to be expected. Depending on the specifics of the solution, smartphone app-based tracking may be a better solution, here, but this may be a more fitting solution for trusted traveller programs since not all

[29] In fact, such characteristics should specifically be taken into account whenever defining the training data during the development of AI-driven algorithm for behavioural detection, so as to reduce the effect of corresponding technologies to produce more false alarms for certain groups of travellers.

[30] Experiences from the US show that even from within the ranks of the trained Behavioral Detection Officers, some believe that it aggravated systematic discrimination. For example, the New York Times writes: "More than 30 federal officers … say the operation has become a magnet for racial profiling … 'They just pull aside anyone who they don't like the way they look – if they are black and have expensive clothes or jewelry, or if they are Hispanic,' said one white officer, who along with four others spoke with The New York Times on the condition of anonymity." (Schmidt and Lichtblau 2012).

[31] Since travellers may have to reveal their faces as part of the identity checks, the additional impact due to the introduction of such forms of risk analysis *may* be rather small, depending on what requirements the travellers have to comply with currently.

travellers can be required to carry a compatible smartphone with them. Furthermore, as mentioned above, location tracking causes additional privacy impact. If applied to all travellers, it seems plausible to assume that potential gains in privacy and time savings for some won't offset the additional impact especially for Union citizens and travellers of similar status, unless as part of a voluntary trusted traveller program.

However, it seems unclear if forms of anomaly detection plausibly provide sufficient risk data for low risk classifications (as it only allows detecting deviance from the norm). This would mean that "normal" or "average" movement patterns imply a lower classification.[32]

Additionally, all forms of behavioural analysis are likely prone to a lack of accountability, since it is very hard to assess whether the criteria for analysing behaviour have a sound scientific basis, are proportionate and have been applied correctly: This is not only true for the detection of behavioural cues by trained officers (Weydner-Volkmann 2017), but also for any application of automated, AI-driven pattern analysis.[33] Hence, it seems almost impossible for travellers to challenge such opaque decisions later on, even if they are repeatedly falsely categorized as higher risk. A redress-system may mitigate some of the impact, but not fully. It is also unclear how a redress system could feasibly be implemented for such cases. In case of support technologies for traveller interviews, some of the scientific concerns regarding automated support for deception detection could be addressed by using such data only in the interest of the traveller, i.e. help towards a lower risk classification or disprove suspicion.

One of the advantages of behavioural detection is that it does not require previously collected personal data and may allow the deletion of all collected and processed data after the border checks have concluded. Of course, if data is retained for re-use or to collect data for training AI technologies, additional impact will manifest due to the higher threat of misuse of data (e.g. due to mission creep in different security contexts or due to unauthorized access to the data).

In any case, forms of behavioural analysis should be communicated openly (up to a certain level of detail). This is only possible if abstract knowledge about the measures and technologies doesn't undermine their effectiveness. In any case, it is highly desirable to allow enough transparency regarding the methods and technologies to allow for scientific scrutiny.

3.3.5. Preliminary recommendations regarding risk assessment methods

Based on the preliminary findings discussed in the previous sections, it becomes clear that the data driven aspects of RBBM can raise profound ethical issues that need to be addressed as part of defining and choosing risk assessment methods and technologies that enable their implementation. It is important to understand that the impact of introducing RBBM will, in the end, depend on the specifics of the implementation, i.e. the details of the design of the BCP procedures. Nevertheless, based on the preliminary findings, we find it plausible to argue that situational risk based border checks will likely imply less severe negative ethical impact than checks based on risk profiling or behavioural analysis.

While any form of border checks will involve negative ethical impact, situational risk based border checks may offer benefits for a lot travellers that outweigh added negative impact for some. Since traveller differentiation is not reliant on personal or identifying data and since

[32] It would also require to make the assumption that "normal" movement patterns cannot be faked deliberately, since otherwise *male fide* travellers could simply deliberately "follow the crowd" through the BCP so as to be checked less intensely.

[33] Efforts in the development of "explainable AI" may change this in the future.

risk categorizations affect larger cross-sections of society equally (e.g. all travellers on a flight), privacy and data protection impact, as well as discrimination and restrictiveness issues can be limited if implemented accordingly.

The nature of risk profiling, on the contrary, involves the processing of personal and identifying information which raises considerable privacy and data protection issues. This is exacerbated by potential discrimination and restrictiveness issues, such as the feasibility of robust biometric identification, false or incomplete external data, transparency and accountability issues as well as the risk of data misuse of collected data. Since the use of such measures tend to yield a high negative impact, enabling technologies should be designed to limit the impact wherever possible. Even then, if implemented, risk profiling should be limited to special cases like second line checks, e.g. to support current forms of interviews that already imply a high privacy impact. Otherwise, it is likely that ethical benefits cannot plausibly outweigh the negative ethical impact on travellers.

Similarly, behavioural analysis raises considerable issues with regard to potential discrimination or other unfair distribution of impact across certain traveller groups. Hence, special attention must be paid to such issues when deciding about implementation in border checks – especially with regard to the feasibility for certain groups of travellers or regarding the fair distribution of false alarms. If such impact can be limited effectively, ethical benefits can still only plausibly outweigh negative impact with regard to restrictiveness, if potential enrolment processes do not effectively increase the temporal costs for travellers and if the transparency and accountability impact can be minimized effectively.

4. MODES OF ETHICAL, LEGAL AND SOCIETAL IMPACT

In this chapter, we will further differentiate the twelve types of ethical, legal and societal impact presented in Table 2 on page 18. Based on the analysis of enabling technologies developed in TRESSPASS, for each type of impact, we will identify different *modes of impact*, i.e. different ways in which those types of impact become manifest as part of the (risk-based) border checks procedures.

As mentioned in chapter 2 above, we will build on a similar framework for the ethical evaluation of airport security screening (Volkmann 2017; Weydner-Volkmann 2018), which was developed as part of the FP7 project XP-DITE. Each of the modes of impact, however, will need to be carefully adapted to reflect the contextual peculiarities of RBBM.

Furthermore, we need to ensure that these modes of impact also adequately reflect the changes implied by moving to a risk-based border management concept. Since much of the technology behind differentiating border checks according to risk indicators is data driven, the (semi-)automated processing of personal data will play a central role in this adaptation of the framework. Hence, we will need to ensure that the modes of impact allow an adequate reflection of qualitative differences with respect to the corresponding types of impact.

In this preliminary and incomplete version of the framework, we will identify the modes of impact for only two of the twelve impact types. Deliverable D9.8 will contain the complete version of the framework, i.e. it will complete the development of the modes of impact, it will provide the full list of observable indicators used for assessment and it will add the coding and aggregation rules for the assessment of qualitative indicators.

4.1. Modes of impact and indicators for ELSA category A: Privacy and Data protection

4.1.1. Modes of impact and indicators for spatial privacy

As mentioned in deliverable D9.6, in times where either much of people's conduct leaves digital traces or where personal interactions and other events in people's lives happen entirely or at least in part digitally, it is important to consider new forms of shielding mechanisms. Such "virtual private spaces" include privacy settings in apps and social media platforms to allow a certain measure of control about how widely certain information about person's conduct is accessible publicly – i.e. even though private information may exist somewhere on the web, it may be part of the deep web by limiting the access to it to a certain number of people (friends or persons who know a shared secret like a password or a specific link). Such virtual private space can also be created through encryption-technologies, even if the underlying data is publicly accessible. Similarly, some people choose to store certain information only "offline" on devices they physically control to further limit the chances of a breach of the shielding effect.

As part of the revelatory function of border checks, such "digital shielding" may be removed. This must be adequately reflected when we adapt the modes of impact identified in the XP-DITE project for spatial privacy (Volkmann 2017) for the purpose of this document.

We propose that there are four modes of impact, i.e. ways in which the border guards (BGs) may intrude into the travellers' private spaces: They can reveal the insides of such spaces directly by (1) either physically accessing items that travellers have placed in private spaces (such as bags or vehicles) or (2) by exposing the insides of private spaces to the gaze of the BGs; furthermore, BGs can (3) process data from the insides of virtual private spaces and they can (4) store information on the insides of physical or virtual private spaces for later use. These

four modes of impact form the first set of concept leaves. Table 3Table 1**Error! Reference source not found.** presents the observable indicators that have been selected for evaluating the impact of intrusion into private spaces for border checks.

TABLE 3: MODES OF IMPACT AND INDICATORS FOR SPATIAL PRIVACY

Impact type:	**Intrusion into (virtual) private spaces**			
Modes of impact	**BGs access belongings**	**Exposure of insides to BGs' gaze**	**BGs process data from virtual insides**	**BGs store collected data**
Observable Indicators	• Are vehicles or luggage items searched? ➢ *Is it a full search?* ➢ *Is it a mobile home?* • Are declared items touched?	• Do declared items have to be displayed? • Are images taken from insides of luggage or vehicles? ➢ *Can border guards see more than 2D shapes and densities?* ➢ *Is there no technology that removes known benign objects from images?*[34] ➢ *Is anomaly detection2*[35] *used on luggage or vehicles?*	• Is information collected from dark web? ➢ *Is it limited to sites pertinent to specific risks?* ➢ *Is doxing data actively removed from collection?* • Is information collected from deep web? ➢ *Are passwords to be provided?* • Are electronic devices accessed? ➢ *Do devices have to be unlocked?* • Is there a penalty for not installing a mobile app on a private device? ➢ *Is the app fully open source?*	• Is collected data stored ➢ *Is it not stored in a traveller-controlled (e.g. token-based) way?* ➢ *Is it stored after the traveller is cleared in a non-traveller-controlled way or communicated to others?* • Is darknet data linked to traveller identities? ➢ *Is it retained longer than 6 months?* • Is deep web data retained? • Is data from electronic devices kept? • Do apps access virtual private spaces?

4.1.2. Modes of impact and indicators for bodily privacy

In deliverable D9.6, we argued that allowing the border guards (BGs) to touch or gaze at (parts of) travellers' bodies that they chose to conceal from the public intrudes into the travellers' privacy. The same is true for collecting and processing information about travellers' body, which is an important aspect of many forms of data driven forms of traveller differentiation in RBBM. This must be adequately reflected when we adapt the modes of impact identified in the XP-DITE project for spatial privacy (Volkmann 2017) for the purpose of this document.

[34] The software component "benign object removal" can be used to remove common objects from x-rays that are known to be harmless. This helps BGs not be distracted but may also reduce the privacy impact.

[35] Anomaly detection provides more information about the insides of divested items as it detects whether the contents differ from the norm in a certain way. On the other hand, trace detection and malicious object recognition are used to find specific things directly, and not by checking against an expected norm.

We propose that there are four modes of impact, i.e. ways in which the border guards (BGs) may intrude into the travellers' bodily privacy: They can reveal more of the body than is plainly visible by (1) removing covering clothes, either physically or by using image producing technologies or (2) by taking physical access to travellers' bodies; Furthermore, they can (3) gain information about (parts of) the body and they can (4) store biometric information for identification or verification. These four modes of impact form the second set of concept leaves. Table 4 presents the observable indicators that have been selected for evaluating the impact of intrusion into bodily privacy.

TABLE 4: MODES OF IMPACT AND INDICATORS FOR BODILY PRIVACY

Impact type:	**Intrusion into bodily privacy**			
Modes of impact	**BGs take access to travellers' bodies**	**Exposure of body to BGs' gaze**	**BGs gain info about body**	**BGs store biometric data**
Observable Indicators	• Do BGs touch the traveller's body? ➢ *If BGs touch the traveller's body, do they touch beneath covering clothes or is it an enhanced frisk search?* ➢ *If BGs touch the traveller's body, is it a full search or are intimate zones included in a directed search?* ➢ *If BGs touch the traveller's body, does the BG have a different sex than the traveller?* • Do BGs handle prostheses or other body extensions? • Do BGs interfere with travellers' bodily functions?	• Is traveller required to divest covering clothes? • Is an image[36] of the body visible to BGs? ➢ *If a body image is visible, are intimate zones included?* • Are images collected from illegal (darknet) trading sites? ➢ *If images are collected, are the trading sites pertinent to specific security goals?* ➢ *If images are collected, are images victims[37] excluded?* • Are traveller's bodily reactions recorded in interviews? ➢ *If reactions are recorded, is (non-evidence)[38] data stored after checks have concluded?*	• Is trace detection used on the body that where known medicinal products cause false alarms? • Is anomaly detection used on body images? ➢ *If anomaly detection is used on body images, will medical implants, piercings, amputations, etc. cause false alarms?* • Are movement patters tracked and analysed? ➢ *If movement is tracked, does reduced mobility cause false alarms?* ➢ *If movement is tracked, are sensitive areas[39] included?*	• Are high resolution images of travellers stored after they are cleared? ➢ *If high resolution images remain stored, are they stored longer than three months?* • Is biometric data stored after the traveller is cleared? ➢ *If biometric data is stored, is it not stored in a traveller-controlled way?* ➢ *If biometric data is stored, can it be used to re-identify travellers later or in another context?* ➢ *If biometric data is stored, is it communicated outside of the BCP?*

[36] Here, "image" means any type of visual representation of the traveler's *specific* body. Completely *generic* images like avatars or stick figures (often used to visualize localized alarms) are not considered body images.

[37] "Victims" of doxing, sexual assault, harassment and other crimes should be protected from additional impact due to unspecific data collection from illegal trading sites.

[38] Recordings that serve the official purpose of evidence as part of a formal investigation and are subject to corresponding data protection regulations are not considered here.

[39] Here, "sensitive areas" refers to restrooms, medical aid facilities, nursery areas and similar body-related places.

5. TRESSPASS-PERSONA COOPERATION ON ACCEPTANCE DATA

There is an important distinction to be made between what an assessment of the ELSA related impact can deliver and what can be gained by collecting empirical acceptance data. The ethical impact assessment comparatively evaluates specific BCP designs against a set of relevant normative concepts. The meaning of those concepts must be well specified within the cultural and historical context in order to be able to assess whether different BCP design perform better or worse in these respects (cf. chapters 2-4).

Contrary to that, measuring acceptance amongst the traveling population is an empirical task that involves collecting perception related data on a highly complex concept. Since measuring acceptance would require a separate approach from TRESSPASS's ethical framework, TRESSPASS will (in accordance with the original call and the DoW) cooperate with the EU H2020 project PERSONA, that deals with assessing acceptance of no-gate border crossing solutions – i.e. TRESSPASS will not conduct an acceptance study on its own, but rather collaborate with the PERSONA project on adding an acceptance dimension to the overall CONOPS framework. This collaboration is organized in T6.4 in collaboration with T9.3.

At a first workshop in June 2019, it became clear that PERSONA aims to deliver two separate forms assessments: (1) an empirical assessment of acceptance amongst the travelling public regarding no-gate border crossing solutions in a simulated environment; and (2) a normative impact assessment along the lines of a data protection impact assessment.

Feedback on PERSONA's initial presentation of the intended method was generated from T9.3 (together with T6.4) to start the collaboration. Two further workshops are planned for 2020 and we currently explore the possibility of having PERSONA assess one of the TRESSPASS pilots. As of yet, details on the PERSONA assessment method, on the PERSONA simulation environment and on the plans for the TRESSPASS pilots are not available. Hence, the details of the collaboration remain to be specified further in the coming months.

Nevertheless, it must be stressed that assessing public acceptance in a simulated environment is a complex task that may not be readily transferred from one environment to another to the next. Furthermore, it is important to see that, while no-gate border crossing solutions and risk-based border management have many aspects in common, the two concepts are not the same. Hence, it will remain to be clarified as part of T6.4 in what way TRESSPASS can genuinely profit from integrating PERSONA's method as part of the overall CONOPS approach.

From the perspective of the ethical framework, we foresee that the empirical assessment of acceptance will not interact with the ethical framework on a methodological level. Instead, it may be used separately as part of the CONOPS framework to generate an additional dimension of planning data for BCP design (and European border policy drafting). Depending on the method of data collection and interpretation, it may be possible to complement certain parts of TRESSPASS's typology of relevant ELSAs with corresponding perception related data from PERSONA. For example, the overall score for ELSA category A (privacy and data protection) for two separate BCP designs could each be complemented by score that gives some indication how travellers would perceive being subject to border checks in such BCP designs.

In case there are contradictory results – e.g. when the TRESSPASS score would indicate that a BCP 1 has more negative impact on travellers' privacy than a BCP 2 while the acceptance score would indicate a higher *perceived* protection of privacy for BCP 1 – this would not be a methodological problem, as the two scores reflect different aspects of BCP designs. Instead,

the two scores could give complementary information to decision makers on what aspects of a BCP design to focus on and what trade-offs are to be expected. On the contrary, if the two scores are congruent, this would also mean that the perception related score cannot be used for validating TRESSPASS's score.

On the other hand, exchanges with the PERSONA project indicate that the project's work on normative impact assessment aims at a similar goal as TRESSPASS framework for ethical impact assessment. For this, PERSONA plans indicate that they are developing a method that is closely modelled after a data protection impact assessment. While it is hard to predict what PERSONA's method will look like in the future, this suggests that the project will develop an independent method for assessing differences in the quality of privacy and data protection impact, which is the focus of ELSA category A.

Depending on what normative concepts their methodology will assess in the end and how, it may be possible to use this outcome of the PERSONA project to validate at least parts of the TRESSPASS framework for ethical impact assessment. Ideally, in case of conflicting results between the two frameworks, it should be tried to explain this discrepancy and improve one of the frameworks. In reality, however, due to the different foci of the two projects, conflicting results may be explained in ways that illustrate that the two seemingly conflicting results are, indeed, compatible – but conflicting results should still be explained. On the other hand, this also means that while congruent results from the two independent methods will support each other, this should only be seen as *part of a more comprehensive validation* unless it can be plausibly shown that the differences in focusing on RBBM vs. focusing on no-gate border crossing solutions is irrelevant.

In addition to the WP6 reports that focus on the CONOPS integration, Deliverable D9.8 will describe how input from the TRESSPASS-PERSONA collaboration was used to complement, improve or validate the TRESSPASS framework for assessing ELSA related impact.

6. FRAMEWORK INTEGRATION IN TRESSPASS

This chapter outlines how we foresee the integration of the TRESSPASS framework for ELSA related impact assessment with WP6's overall CONOPS framework (section 6.1) and WP7's simulation activities (section 6.2).

6.1. Integration with the overall CONOPS framework

6.1.1. Ethical impact assessment along three Key Performance Indicators

We currently foresee that there will be three discrete overall ethical performance scores. They will assess the ELSA related impact of risk based BCP designs on the level of the three categories of ELSAs as ethical Key Performance Indicators (KPIs) within the overall CONOPS framework. The idea is to allow decision makers to identify not only trade-offs between ethical impact and, for example, throughput, but also between the different categories of ethical impact, e.g. between privacy intrusions (ELSA category A) and potential for discrimination (ELSA category B).

While it will ultimately depend on how the rump scales from deliverable D9.6 will be further developed into four point ordinal scales and scoring defined as part of the aggregation rules in deliverable D9.8, we currently foresee that it should be possible to define "minimum levels of protection" for these categories. These can act as ethical restrictions during BCP design and evaluation. These minimum levels of protection could then be complemented by desired performance goals that can be defined during BCP design.[40]

6.1.2. Brief description of the CONOPS framework

As outlined in deliverable D6.1 ("Observational Studies Methodology and Research Framework"), according to the IEEE (1998), a Concept of Operations (CONOPS) refers to a communication document which is used to support the process of system development or system change. It provides a consensus document that communicates the vision for change from the current system to the prospective system to all system actors and stakeholders. Such a document aims to establish the core concept behind this vision for change, in this case the TRESSPASS system for the development of risk-based border management.

To develop the TRESSPASS CONOPS, we are drawing on Engeström's (1987) cultural-historical activity theory (CHAT) (cf. deliverable D6.1 for our analysis of the ethical implications of the Current CONOPS and the Future CONOPS). To summarize, CHAT refers to a psychological framework for analysing how individuals, groups, organizations and communities interact with their material, historical and sociocultural worlds. Engeström's activity system (as illustrated in **Error! Reference source not found.**) depicts the interconnectedness of persons (subject) with various nodes (i.e. object, instruments, community, division of labour, rules and outcome). Subjects act towards a goal (object) to achieve a particular outcome. The achievement of the outcome is dependent on a range of dependencies, including the tools or artefacts (i.e. instruments) available to them; the socio-cultural contexts (i.e. community) in which the subject operates; the interconnectedness and interdependencies of the subject with others whose activities influence each other (division of labour) and hence, facilitate or impede the subject's ability to achieve their goal and lastly, the ethical, legal, procedural and socio-cultural constraints (i.e. rules) that limit activity. Collectively, all of these nodes and their

[40] A similar approach was taken in XP-DITE (Volkmann 2017).

relation to each other constitute the activity framework. None of these nodes is meaningful by themselves but must be considered in relation to all the others. Based on this framework, we have formulated a set of questions to investigate the operational processes the three BCPs of the TRESSPASS pilots (Schiphol Airport, Piraeus Sea Port and Terespol):

- **What** are the border guards doing?
- **Why** are they doing it?
- With **whom** are they interacting?
- **What** tools or resources do they use or require?
- What are their **constraints**?
- What is the overall **motivation** for their collective activity?

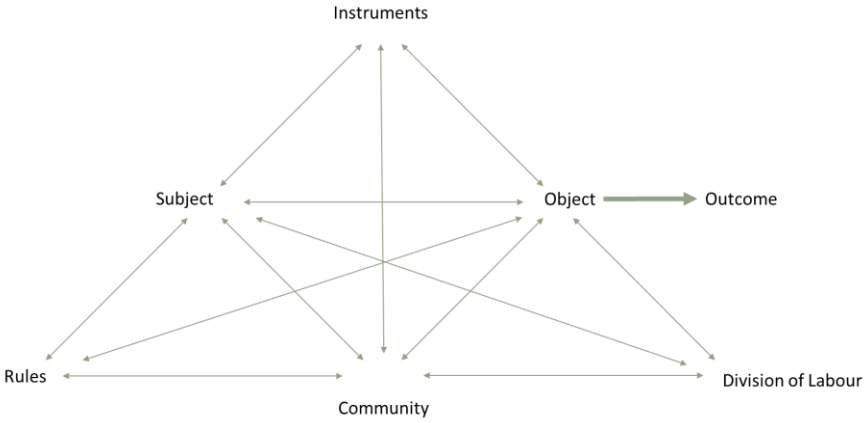

FIGURE 8: ENGESTRÖM'S (1987) ACTIVITY SYSTEM

Observations of the operational processes at the BCPs combined with conversations with border security staff resulted in a CONOPS of the current "as is" situation (cf. deliverable D6.2, first version). A graphic representation of a generic CONOPS is provided in **Error! Reference source not found.** below. The model depicts a generic airport departure process mapping the paths travellers take from start to finish. Separate paths ("swimlanes") illustrate the operational processes that take place simultaneously depicting the activities of border security "actors" including border guards, airport and security staff, airline staff as well as "external actors" (i.e. databases). The swimlanes also illustrate at which point the operational processes of other "actors" intersect with the traveller as they proceed from one stage to the next. As the CONOPS develops, it will also incorporate the system architecture (tools) and risk indicators that will lead to a more evolved illustration of the future TRESSPASS system.

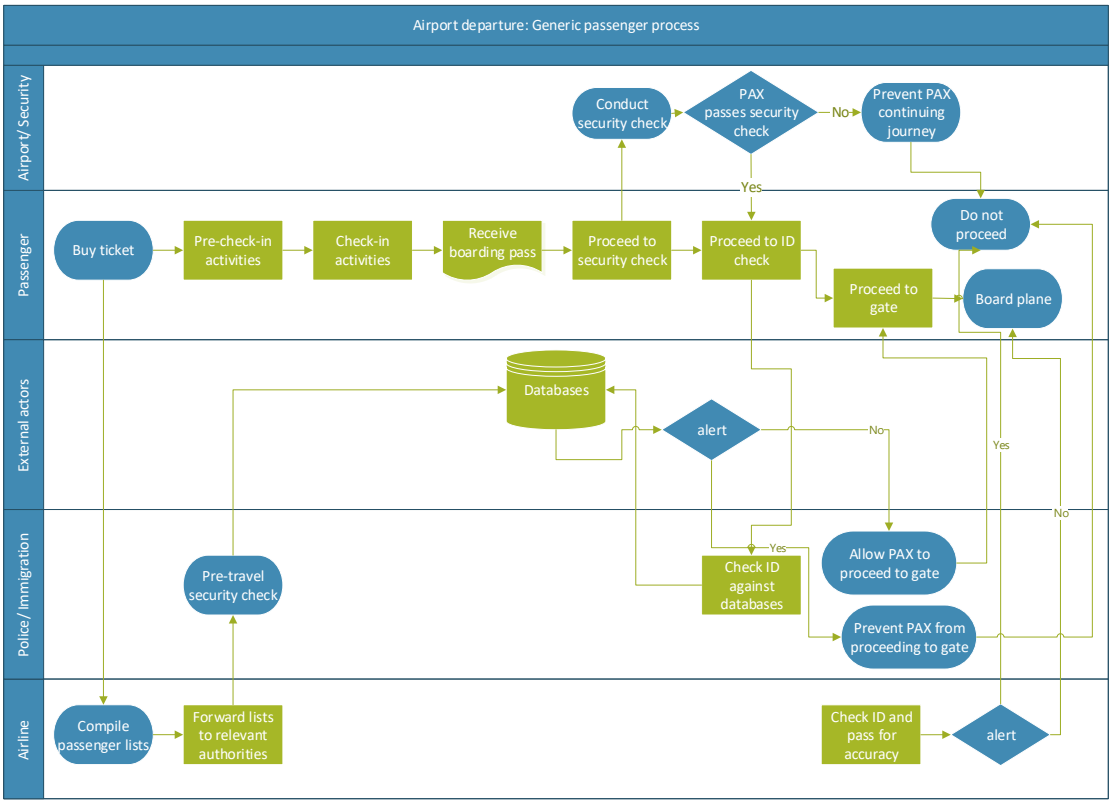

FIGURE 9: GENERIC CONOPS DEPICTING THE DEPARTURE PROCESSES AT AN AIRPORT

6.1.3. Integration of the overall CONOPS framework with the ethical framework

Both the "as is" and "to be" CONOPS depictions can be meaningfully translated back and forth between the ethical framework's formalized depiction of paths through the BCP as described in Section 2.2.1 (cf. Figure 3 on page 17). Using both the CONOPS and Engeström's (1987) CHAT system (**Error! Reference source not found.**), we can review and interrogate each activity at each stage of the travel process in terms of its ethical, legal and societal implications using the ethical framework and its scoring on qualitative scales (sections 2.2.2 to 2.2.4).

For example, during the security check the traveller interacts with security staff. Both actors share a similar goal: the traveller seeks to pass through the security checkpoint as quickly and securely as possible; the security officer seeks to ensure that travellers pass through the security checkpoint as quickly and securely as possible. Although the emphases on security and speed may differ (e.g. security guards may prioritize security over speed), both security

guard and traveller share the same goal, or object. This shared goal is called 'boundary object' (Star and Griesemer 1989) and is illustrated in Figure 10.

.

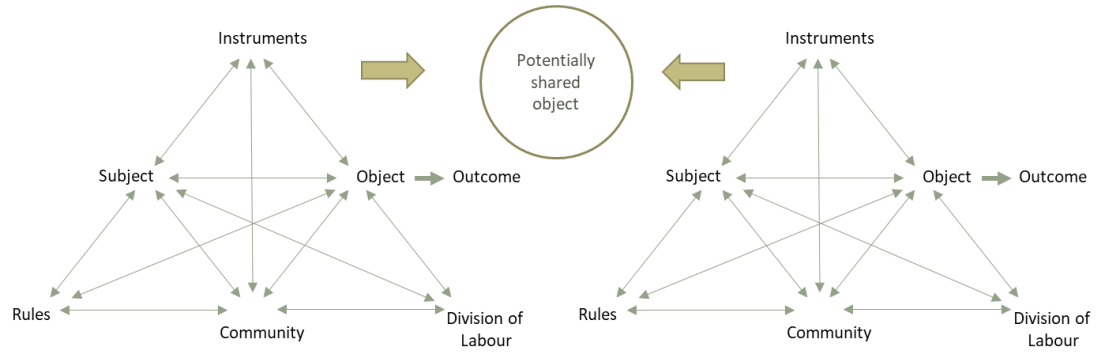

FIGURE 10: BOUNDARY OBJECT – OVERLAPPING ACTIVITIES

However, despite shared goals, the interaction between travellers and security staff may have ethical, legal and societal implications. Pursuing the example of the security checkpoint, we may interrogate the operational processes in relation to the tools available to the security personnel. We may evaluate the activities based on the assessment categories of the ethical framework (i.e. intrusiveness, restrictiveness and fairness) and label these activities based on the scoring on the scales concept (e.g. close to a best-case scenario 'A' or close to a worst-case scenario 'D'). The following fictional scenario serves as an example for a procedural design of border checks close to the worst case with regard to the fairness category:

> **Fictional scenario**:
>
> A person with a prosthetic leg approaches an airport BCP with CCTV based tracking and behavioral detection based on movement patterns is in place. Since the system is known not to be working for travelers with reduced mobility, she cannot enroll in the system and has to be checked in an alternative procedure. Because some cases of drug trafficking are known that involved travelers with reduced mobility hiding drugs in their luggage and in some cases even in prosthetics in order to avoid behavioral detection measures, travelers with reduced mobility are subject to a full search of their luggage and, if applicable, an x-ray of prosthetics that is done on an x-ray luggage scanner that can be seen from the outside.

The bad score in the fairness category and its reasoning based on paths through the BCP of a given ("as is") BCP CONOPS may help to establish a knowledge base of the ELSA implications of current operational processes. Ideally, this guidance document will inform the CONOPS outlining a future design based on the TRESSPASS methods and tools. For example, the activity described in Scenario A is likely to score badly in terms of the fairness category, and the ethical framework allows highlighting the reason for this: because the impact of customs checks is clearly much higher for travellers with reduced mobility than for travellers who can enrol in the movement tracking system: not only are they much more often subject to a bag search (spatial privacy impact), the impact for travellers with prosthetics are particularly high since prosthetics are handled by border guards in the open.

In a future "to be" CONOPS, this score in the fairness category could be improved in terms alternative border checks designs for travellers with reduced mobility. Regarding the human

factors, the interaction between the respective subjects (i.e. border guard and traveller) could be improved by checking prosthetics in another fashion without requiring removal or x-raying in public; in regard to the system-level design of the BCP, it must be ensured that potential privacy benefits due to risk-based screening are shared fairly and not to the detriment of travellers with reduced mobility. Hence, the BCP design could be enhanced by conducting a comparative evaluation of the "as is" and plausible versions "to be" CONOPS. A comparative evaluation of the ethical implications of the "as is" and potential "to be" CONOPS will enhance informed decision-making processes.

6.1.4. Acceptance as a complimentary dimension in the evolving CONOPS document

As the findings from the PERSONA project become available (cf. chapter 5), we can integrate them into the evolving CONOPS document and thus, the TRESSPASS approach to BCP design, which is informed by the CONOPS. The integration of traveller acceptance of risk-based or, in keeping with the terminology used by PERSONA, no-gate border crossing solutions may play an important role in the development of the future CONOPS. Even for system compliant with EU and national legislations and regulations, and a comparatively good score in the ethical impact assessment, user acceptability of a BCP design may perform badly in terms of traveller acceptance. Hence, user acceptability may facilitate or limit the usability and thus, implementation of a risk based BCP based on the TRESSPASS methods and tools. As mentioned in chapter 5, acceptability data may, thus, provide a complementary perspective.

The integration of the ethics framework and its overall scoring of BCP designs in three ELSA categories with the CONOPS (as described in this section) will allow us to identify potential obstacles or challenges which can be addressed as part of the evolving CONOPS. The ELSA scores and the scoring rules will allow us to identify border checks procedures that travellers may find problematic or unacceptable and if possible, to adjust them. Although there may be activities or processes that cannot be changed due to a range of factors (e.g. technological limitations, staffing issues, rules and regulations, economic constraints), the future CONOPS will nevertheless act as a document based on which the TRESSPASS system can evolve alongside future technological innovations, socio-economic and legislative changes.

6.2. Integration with the TRESSPASS simulation and evaluation activities

6.2.1. Evaluation platform for RBBM and the ethical framework and

As part of Task T7.5, it is foreseen in TRESSPASS to include a performance evaluation (e.g. legal, regulatory and ethical compliance, traveller and operational acceptance, throughput, seamlessness) from a regulator, operator and traveller perspective as part of the evaluation platform for risk-based border control systems (BCSEP). Part of this, namely the assessment of ELSA related impact on the traveling public, will be based on the framework presented in a preliminary, incomplete version in the present report (the final version will be submitted as deliverable D9.8).

As mentioned in section 2.2.3 above, the framework's methodology has been developed in such a way so as to facilitate a largely computational evaluation of a given BCP design, so long as all relevant information is included in the formalized BCP design description used by the BCSEP for performance evaluation. Although the framework can be used for a manual evaluation, this implies that the majority of the yes/no questions that serve as binary observable indicators as part of the framework should be answered based on the formal design descriptions. This way, only a manageable number of indicator values should need to be established by hand.

Furthermore, it is necessary for the framework to know how many travellers will be checked on each path through the BCP, i.e. it is necessary that the BCSEP can establish for what rate of travellers an alert is raised for each border checks procedure. For example, with regard to the simplified border checks procedure illustrated in Figure 3 on page 17, it would be necessary for the BCSEP to establish the rate of travellers which are referred to second line checks due to a hit in one of the databases, as well as the rate of those travellers that can then subsequently be cleared to enter the EU.

6.2.2. Introduction to the performance evaluation of BCPs:

T7.5's performance evaluation at BCPs follows the system mechanism of the checkpoint process. According to the concept of risk-based border management outlined in deliverable D1.2, the BCP design will be evaluated in four performance areas, namely effectiveness (success-rate of stopping unauthorized travellers), flow-rate (speed of flow), efficiency (number of resources required), and level of ethical compliance. The performance area "level of ethical compliance" can be expressed in three ethical performance categories as foreseen in section 2.2.2) and along these, alternative designs can be compared to each other to find ways of ethical impact mitigation.

In the design modelling process, the simulation aims to enhance and optimize the first three indicators of the system performance and highlight negative ethical impact on the travellers. For this, the simulation utilizes a stochastic approach to generate random anonymous travellers based on crowd behaviour distribution. The process relies on queuing behaviour of crowded travellers from arriving at BCPs to waiting to be checked before exiting the system.

Figure 11, for instance, shows different sections of the screening and checking process at an airport. When a traveller arrives at the BCP, they approach the screening area where a set of pre-defined border guards receive some risk indicators on the traveller based on externally provided data and behavioural analysis of movement patterns. Due to a risk assessment based on the risk indicators, the traveller is selected to specific checking procedures.

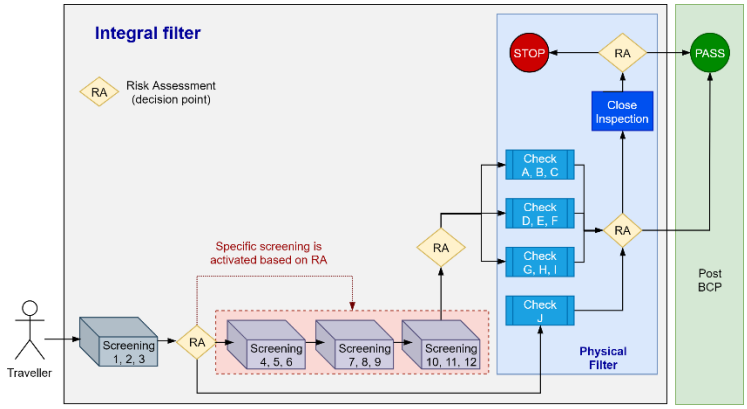

FIGURE 11: HIGH LEVEL BCP PROCESS DESIGN FOR SIMULATION

At this point, we assume that most travellers will be assigned to a lower risk category for basic checking and occupy the corresponding lanes for being checked. Some travellers, however, will be assigned a higher risk category and may be sent to second line checks. Hence, they will move through the BCP on other paths. The algorithm that simulates the path decision-making for a BCP design should reflect the output values generated by sensors and data analytics regarding risk indicators of a real system (e.g. the pilots). The simulator will try to imitate the real system in order to assess the performance of the BCP as realistically as possible.

Therefore, in the simulation, the majority of travellers are assumed to pass basic checking. At the basic level checking process, a number of travellers queueing at individual lanes are accounted according to their individual service time. The results of the simulation then allow a characterization of the flow of travellers throughout the system, which can help to optimize the flowrate and efficiency of BCPs. The corresponding data on service time can also be used as part of the ethical framework to assess relevant aspects like time loss. Details of the simulation approach and performance evaluation are addressed in deliverable D7.3.

6.3. Paths through the BCP as a compatible approach

During the CONOPS and ethics by design workshop in Freiburg (30 September – 01 October 2019), an open discussion and exchange of perspectives took place around the integration of the overall CONOPS framework, the ethics framework, the design activities and the simulation and evaluation activities in TRESSPASS.

One of the main outcomes of the workshop for this report was that partners agreed to use the schematic overview that was proposed by NUIM to depict the current “as-is” CONOPS (“swimmlanes”) as a common way for the consortium to exchange formal information on the design and procedures of border checks. In order to move towards realistic use cases for new technologies, it was decided to start filling in those technologies available to the TRESSPASS project (and especially those developed as part of the project) that enable a risk-based design. For this it was decided to undertake a “mapping” of the TRESSPASS risk indicators to the available technologies and then proceed with including those technologies in a meaningful way as part of a future “to-be” CONOPS.

As has become clear in sections 6.1 and 6.2 above, the ethical frameworks assessment based on paths through a BCP is compatible with the “swimmlanes” depiction as part of WP6’s CONOPS framework, but also works well with the formalization of BCP designs for simulation and evaluation as part of WP7.

7. CONCLUSIONS

Based on the typology of ethical, legal and societal issues of risk-based border management presented in TRESSPASS's deliverable D9.6, this report has presented a preliminary and incomplete description of a framework for impact assessment. As shown in chapter 2, the framework will allow a *comparative assessment* of different procedural designs for border checks. It aims at allowing a better understanding of the trade-offs involved in introducing risk-based border checks as part of a future border management regime. This is done by comparing the effects of the procedural designs of border crossing points along the twelve types of relevant impact specified in the typology. We hope that this will allow ethically informed and well-balanced decisions about the kind of border checks desired for Europe's external borders (Ethics and Data Protection by design).

Chapter 3 of this report has analysed how each of those types of ELSA related impact relate to the TRESSPASS developed enabling technologies for risk-based border management. This has ensured that the framework can adequately reflect the impact of *risk-based* checks. As part of this analysis, options for value sensitive design during technology design were generated an included in deliverable D3.1 and D4.3. The preliminary findings on the ethical trade-offs implied by different forms of traveller risk assessment were included in chapter 3. They also formed input to deliverable D2.3.

For two of the twelve types of impact, modes of impact and observable indicators for impact assessment were already defined in chapter 4 of this report. In deliverable D9.8, this will also be done for the rest of types of impact defined in the typology along with the definition of the coding and aggregation rules for qualitative evaluation of the impact along four-point ordinal scales. This complete framework will form the basis for deliverable D9.9's guidelines for decision makers.

In chapter 5, a preliminary outline was presented on how WP6's collaboration between the TRESSPASS and PERSONA projects on traveller acceptance data could be used in relation to the ethical framework. Lastly, in chapter 6, the ethical frameworks integration with WP6's overall CONOPS framework and WP7'ssimulation and evaluation platform has been described in more detail. Given the compatibility of the ethical framework's differentiation of paths through the BCP with both, the CONOPS depiction of "swimmlanes" and the formal designs for simulation and evaluation, we foresee no great obstacles for the integration and compatibility.

While the framework for impact assessment is focused on assessing the impact of *risk based* border checks, it is designed in such a way so as to also allow an assessment of current, rule based checks. Hence, the ambition for D9.8 is, as part of WP6's overall CONOPS framework, to allow a comparison of rule based and risk based checkpoint designs. While the impact assessment will always depend on the specifics of the BCP designs, abstracting from some specific examples of BCP designs would then enable a more informed decision making about the ethical implications of moving from rule based to risk based border management.

Based on the results of applying the framework to a range of risk based and rule based BCP designs, it should then be possible to develop an ethical argument for confirmation or falsification of hypotheses such as the following ones:

- "When designed and applied in the proper way, risk based BCPs can allow a reduced privacy impact for the majority of travellers compared to traditional rule based BCPs."

- “Risk-based BCPs tend to offer less impact with regard to some types of ethical impact, but do so at the expense of a higher impact with regard to other types.”
- “There is no clear ethical argumentation for or against risk based border management. It all depends on the specifics of the BCP designs.”
- “Due to the added requirements in the collection and processing of personal data, certain forms of risk based BCPs necessarily have a more severe privacy and data protection impact than rule based BCPs.”

8. REFERENCES

ACLU. 2018. “Amazon’s Face Recognition Falsely Matched 28 Members of Congress With Mugshots.” American Civil Liberties Union. July 26, 2018. https://www.aclu.org/blog/privacy-technology/surveillance-technologies/amazons-face-recognition-falsely-matched-28.

Adey, Peter. 2004. “Secured and Sorted Mobilities: Examples from the Airport.” *Surveillance & Society* 4 (1): 500–519.

BGH. 1998. Polygraphentest als Beweismittel JurPC Web-Dok. 13/1999, Abs. 1 - 75. BGH.

———. 2010. 2011_6_504.pdf, 6/2011 Zeitschrift für das Juristische Studium 557. BGH.

Brodkin, Jon. 2019. “US Violated Constitution by Searching Phones for No Good Reason, Judge Rules.” Ars Technica. November 13, 2019. https://arstechnica.com/tech-policy/2019/11/us-cant-search-phones-at-borders-without-reasonable-suspicion-judge-rules/?comments=1.

Bühlmann, Marc, Wolfgang Merkel, Lisa Müller, and Bernhard Weßels. 2012. “The Democracy Barometer: A New Instrument to Measure the Quality of Democracy and Its Potential for Comparative Research.” *European Political Science* 11 (4): 519–36. https://doi.org/10.1057/eps.2011.46.

———. 2017. “Country Profile: Germany.” Democracy Barometer. 2017. http://www.democracybarometer.org/country_profiles/EN/Germany/.

Dewey, John. 1988. *The Later Works Volume 13: 1938-1939.* Edited by Jo Ann Boydston. Carbondale, Ill. [u.a.]: Southern Illinois University Press. https://katalog.ub.uni-freiburg.de/link?id=008002460.

Dewey, John, and James Hayden Tufts. 1985. *The Later Works Volume 7: 1932.* Edited by Jo Ann Boydston. Carbondale (Ill.): Southern Illinois University Press.

Engeström, Yrjö. 1987. *Learning by Expanding: An Activity-Theoretical Approach to Developmental Research*. Helsinki: Orienta-Konsultit.

EU. 2016a. *Regulation (EU) 2016 of the European Parliament and of the Council of 9 March 2016 on a Union Code on the Rules Governing the Movement of Persons across Borders (Schengen Borders Code)*. *CELEX*. Vol. 32016R0399. https://publications.europa.eu/en/publication-detail/-/publication/42fba6c3-f0c5-11e5-8529-01aa75ed71a1.

———. 2016b. *Directive (EU) 2016/ 681 of the European Parliament and of the Council of 27 April 2016 on the Use of Passenger Name Record (PNR) Data for the Prevention, Detection, Investigation and Prosecution of Terrorist Offences and Serious Crime*. *OJ L 119*. http://data.europa.eu/eli/dir/2016/681/oj.

Fesmire, Steven. 2014. *Dewey*. The Routledge Philosophers. New York: Routledge.

FRA. 2011. “Opinion of the European Union Agency for Fundamental Rights on the Proposal for a Directive on the Use of Passenger Name Record (PNR) Data for the Prevention, Detection, Investigation and Prosecution of Terrorist Offences and Serious Crime (COM(2011) 32 Final).” FRA Opinion 1/2011. Vienna: European Union Agency for Fundamental Rights.

Greenberg, Andy. 2014. “The Dark Web Gets Darker With Rise of the ‘Evolution’ Drug Market.” *Wired*, September 18, 2014. https://www.wired.com/2014/09/dark-web-evolution/.

———. 2015. “The Dark Web’s Top Drug Market, Evolution, Just Vanished.” *Wired*, March 18,

2015. https://www.wired.com/2015/03/evolution-disappeared-Bitcoin-scam-dark-web/.

IEEE. 1998. “IEEE Guide for Information Technology - System Definition - Concept of Operations (ConOps) Document.” *IEEE Std 1362-1998*, December, 1–24. https://doi.org/10.1109/IEEESTD.1998.89424.

Jäckle, Sebastian, Uwe Wagschal, and Rafael Bauschke. 2012. “Das Demokratiebarometer: „basically theory driven“?” *Zeitschrift für Vergleichende Politikwissenschaft* 6 (1): 99–125. https://doi.org/10.1007/s12286-012-0133-6.

Lauth, Hans J. 2007. *Demokratie und Demokratiemessung: Eine konzeptionelle Grundlegung für den interkulturellen Vergleich*. 2., durchges. u. aktualis. Aufl. Wiesbaden: VS Verlag für Sozialwissenschaften.

Müller, Thomas, and Susanne Pickel. 2007. “Wie lässt sich Demokratie am besten messen? Zur Konzeptqualität von Demokratie-Indizes.” *Politische Vierteljahresschrift* 48 (3): 511–39. https://doi.org/10.1007/s11615-007-0089-3.

Munck, Gerardo L., and Jay Verkuilen. 2002. “Conceptualizing and Measuring Democracy Evaluating Alternative Indices.” *Comparative Political Studies* 35 (1): 5–34. https://doi.org/10.1177/001041400203500101.

OLG Dresden. 2013, 2013 BeckRS 16540. OLG Dresden.

Ormerod, Thomas C., and Coral J. Dando. 2015. “Finding a Needle in a Haystack: Toward a Psychologically Informed Method for Aviation Security Screening.” *Journal of Experimental Psychology: General* 144 (1): 76–84. https://doi.org/10.1037/xge0000030.

Petermann, Thomas, and Arnold Sauter. 2002. “Biometrische Identifikationssysteme. Sachstandsbericht, Büro Für Technikfolgenabschätzung Beim Deutschen Bundestag (TAB).” http://www.itas.kit.edu/pub/v/2002/pesa02a.pdf.

Poscher, Ralf. 2017. “The Right to Data Protection. A No-Right Thesis.” In *Privacy and Power: A Transatlantic Dialogue in the Shadow of the NSA-Affair*, edited by Russell A. Miller, 129–41. Cambridge University Press.

Sandel, Michael J. 2012. *Was man für Geld nicht kaufen kann: die moralischen Grenzen des Marktes*. Translated by Helmut Reuter. 7. Aufl. Berlin: Ullstein.

Schmidt, Michael S., and Eric Lichtblau. 2012. “Racial Profiling at Boston Airport, Officials Say.” *The New York Times*, August 11, 2012. http://www.nytimes.com/2012/08/12/us/racial-profiling-at-boston-airport-officials-say.html.

Solove, Daniel J. 2009. *Understanding Privacy*. Cambridge, Mass.: Harvard Univ. Press.

Star, Susan Leigh, and James R. Griesemer. 1989. “Institutional Ecology, `Translations’ and Boundary Objects: Amateurs and Professionals in Berkeley’s Museum of Vertebrate Zoology, 1907-39:” *Social Studies of Science*. https://doi.org/10.1177/030631289019003001.

Stevens, Stanley Smith. 1946. “On the Theory of Scales of Measurement.” *Science* 103 (2684): 677–80. https://doi.org/10.1126/science.103.2684.677.

US DHS. 2013. “Transportation Security Administration’s Screening of Passengers by Observation Techniques (Redacted).” Office of Inspector General OIG-13-91. Department of Homeland Security. http://www.oig.dhs.gov/assets/Mgmt/2013/OIG_13-91_May13.pdf.

US GAO. 2013. “TSA Should Limit Future Funding for Behavior Detection Activities.” Aviation Security. United States Government Accountability Office. http://www.gao.gov/assets/660/658923.pdf.

Volkmann, Sebastian. 2013. “Methods for Assessment and Quantification of Compliance with given Ethical Requirements.” D7.3. XP-DITE Deliverable. Albert-Ludwigs-Universität Freiburg. https://doi.org/10.6094/UNIFR/13816.

———. 2017. “Updated Methods for Assessment and Quantification of Compliance with given Ethical Requirements.” D7.10. XP-DITE Deliverable. Albert-Ludwigs-Universität Freiburg. https://doi.org/10.6094/UNIFR/14020.

Wagner, Katrin. 2014. “Vom Werkzeug zum Täter – ein Paradigmenwechsel im zivilen Luftverkehr?” In *Risikobasiert versus One Size Fits All. Neue Konzepte der Passagierüberprüfung im Flugverkehr*, edited by Katrin Wagner and Wolfgang Bonss, 21–33. SIRA Conference Series 3. München: Universitätsverlag Neubiberg.

Weinberger, Sharon. 2010. “Airport Security: Intent to Deceive?” *Nature News* 465 (7297): 412–15. https://doi.org/10.1038/465412a.

Wetering, Elbert van de. 2014. “A Risk-Based Passenger Screening Security Architecture Optimized against Adaptive Threats.” MA-Thesis Erasmus Universiteid Rotterdam. http://thesis.eur.nl/pub/16046/Scriptie-Econometrie-2014-Elbert-van-de-Wetering-publieke-versie2.pdf.

Weydner-Volkmann, Sebastian. 2017. “Risk Based Passenger Screening in Aviation Security: Implications and Variants of a New Paradigm.” In *Rethinking Surveillance and Control. Beyond the “Security versus Privacy” Debate*, edited by Elisa Orrù, Maria Grazia Porcedda, and Sebastian Weydner-Volkmann, 49–83. Sicherheit Und Gesellschaft. Freiburger Studien Des Centre for Security and Society 12. Baden-Baden: Nomos.

———. 2018. *Moralische Landkarten der Sicherheit: Ein Framework zur hermeneutisch-ethischen Bewertung von Fluggastkontrollen im Anschluss an John Dewey*. Ergon Verlag. https://doi.org/10.5771/9783956503788.

9. LIST OF FIGURES

10. LIST OF TABLES