

## H2020 – BES – 5 – 2015

### Research Innovation Action



Intelligent Portable Control System



*This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 700626*

## D4.1 Deliverable Title

<b>Report Identifier:</b>	D4.1 First version of the iBorderCtrl software platform		
<b>Work-package, Task:</b>	WP4	<b>Status – Version:</b>	V1.00
<b>Distribution Security:</b>	CO	<b>Deliverable Type:</b>	R
<b>Editor:</b>	STR		
<b>Contributors:</b>	ED, EVE, ICCS, JAS		
<b>Reviewers:</b>	ED, EVE		
<b>Quality Reviewer:</b>	ED		
<b>Keywords:</b>	Software, Development, Prototype		
Project website: <a href="http://www.iborderctrl.eu">www.iborderctrl.eu</a>			

### **Copyright notice**

© Copyright 2016-2019 by the iBorderCtrl Consortium

This document contains information that is protected by copyright. All Rights Reserved. No part of this work covered by copyright hereon may be reproduced or used in any form or by any means without the permission of the copyright holders.

## Table of Contents

<b>ABBREVIATIONS</b> .....	<b>8</b>
<b>EXECUTIVE SUMMARY</b> .....	<b>9</b>
<b>1 INTRODUCTION</b> .....	<b>10</b>
<b>2 IT INFRASTRUCTURE</b> .....	<b>11</b>
2.1    iBORDERCTRL SYSTEM DESCRIPTION .....	11
2.1.1    System technical description.....	11
2.1.1.1    Traveller User Application (TUA) .....	11
2.1.1.2    Border Guard User Application (BGUA) .....	12
2.1.1.3    Border Manager User Application (BMUA).....	13
2.1.2    System interfaces description/ integration within the iBorderCtrl system .....	14
2.1.3    System data flow description .....	18
2.1.3.1    Traveller User Application .....	18
2.1.3.2    Border Guard User Application .....	19
2.1.3.3    Border Manager User Application .....	20
2.1.4    System data structure.....	21
2.1.4.1    Database tables .....	21
2.1.4.2    Dictionary of iBorderCtrl Database .....	23
2.1.4.3    Database technologies.....	30
2.1.4.4    QR code.....	30
2.1.5    Technical requirements coverage .....	31
2.2    CLOUD INFRASTRUCTURE.....	42
2.2.1    Software stack implementation.....	42
2.2.1.1    Cloud computing services.....	42
2.2.1.2    Cloud Deployment Model .....	43
2.2.1.3    Open Technologies .....	44
2.2.1.4    Migration planning and application .....	45
2.2.2    Security of cloud infrastructure .....	47
2.2.3    Cloud Providers selection .....	52
2.3    SECURITY.....	53
<b>3 RISK BASED ASSESSMENT TOOL (RBAT)</b> .....	<b>54</b>
3.1    RBAT OVERVIEW.....	54
3.2    MULTI CRITERIA DECISION ANALYSIS (MCDA) - STEPS FOR THE "WEIGHT" DETERMINATION OF EACH iBORDERCTRL MODULE.....	56
3.3    SYSTEM TECHNICAL DESCRIPTION .....	63
3.3.1    Weight based algorithm for the final risk score calculation.....	63
3.3.2    Rule Authoring environment.....	64

3.3.3	Risk Database.....	66
3.3.4	Interfaces to iBorderCtrl system.....	66
3.4	RBAT TECHNICAL REQUIREMENTS COVERAGE .....	67
<b>4</b>	<b>EXTERNAL LEGACY AND SOCIAL INTERFACES (ELSI).....</b>	<b>70</b>
4.1	PUBLICALLY AVAILABLE INFORMATION FROM SOCIAL MEDIA PLATFORMS .....	70
4.2	BORDER CONTROL CONSULTS SYSTEM (VIS, SIS).....	72
4.2.1	SIS, VIS, Entry Exist Databases.....	73
4.2.1.1	SIS Schema and Description .....	73
4.2.1.2	VIS Schema and Description.....	74
4.2.1.3	EES Schema and Description.....	76
4.2.1.4	Universal Messaging System (UMF).....	79
4.3	ELSI TECHNICAL REQUIREMENTS COVERAGE.....	81
<b>5</b>	<b>BORDER CONTROL ANALYTICS TOOL (BCAT).....</b>	<b>83</b>
5.1	USER INTERFACE AND PRE-REQUISITE KNOWLEDGE REQUIRED .....	83
5.2	THE UNDERLYING SCIENTIFIC ANALYSES ALGORITHMS.....	83
5.3	EVALUATION OF INDIVIDUAL BORDER CONTROL SYSTEMS.....	84
5.3.1	Association testing.....	84
5.3.1.1	.....	84
5.3.1.2	<u>Showcase</u> Falsified Travel documents (VISA).....	84
5.4	KNOWLEDGE DISCOVERY AND PATTERN IDENTIFICATION.....	84
5.4.1	Advanced exploratory analyses of the collected data .....	84
5.4.2	Correlation Coefficients .....	84
5.4.2.1	Correlation Coefficients Showcase.....	85
5.4.2.2	Principal Component Analyses.....	86
5.4.2.3	Machine Learning Showcase .....	86
5.4.3	Clustering and Classification .....	87
5.4.3.1	Showcase Clustering and Classification for Pattern Discovery.....	89
5.5	BCAT TECHNICAL REQUIREMENTS COVERAGE.....	89
<b>6</b>	<b>USER INTERFACE .....</b>	<b>91</b>
6.1	TUA IMPLEMENTATION.....	91
6.1.1	TUA interface implementation .....	91
6.1.2	TUA screenshots .....	91
6.2	BGUA IMPLEMENTATION.....	102
6.2.1	BGUA interface implementation .....	102
6.2.2	BGUA Screenshots .....	103

6.3	BMUA IMPLEMENTATION.....	109
6.3.1	BMUA interface implementation .....	109
6.3.2	BMUA screenshots .....	109
<b>7</b>	<b>DATA PROTECTION IMPACT ASSESSMENT.....</b>	<b>111</b>
7.1	INTRODUCTION.....	111
<b>8</b>	<b>CONCLUSIONS.....</b>	<b>153</b>

## List of Tables

TABLE 1	TRAVELLERS' USER APPLICATION TECHNICAL REQUIREMENTS .....	31
TABLE 2	BORDER GUARD USER APPLICATION TECHNICAL REQUIREMENTS .....	36
TABLE 3	IBORDERCTRL DATABASE TECHNICAL REQUIREMENTS .....	38
TABLE 4	DIFFERENT APPLICATION MIGRATION OPTIONS SUPPORTED BY IBORDERCTRL PLATFORM. ....	43
TABLE 5	PROS AND CONS OF PRIVATE, PUBLIC AND HYBRID DEPLOYMENT CLOUD MODELS.....	44
TABLE 6	CLOUD SECURITY PRINCIPLES (SOURCE <a href="http://goo.gl/mUf5c2">HTTP://GOO.GL/MUF5C2</a> ) .....	48
TABLE 7	PENETRATION TEST AUDIT TOOLS .....	52
TABLE 8	PERFORMANCE MATRIX.....	59
TABLE 9	CONSEQUENCE TABLE FOR "TECHNOLOGY MATURITY" OBJECTIVE .....	60
TABLE 10	CONSEQUENCE TABLE FOR "ACCURACY AND RELIABILITY" OBJECTIVE .....	60
TABLE 11	CONSEQUENCE TABLE FOR "PERFORMANCE" OBJECTIVE .....	61
TABLE 12	CONSEQUENCE TABLE FOR "UNIVERSALITY" OBJECTIVE.....	61
TABLE 13	CONSEQUENCE TABLE FOR "PHASE APPLIED" OBJECTIVE .....	62
TABLE 14	RISK DATABASE FIELDS EXPLANATION .....	66
TABLE 15	RBAT TECHNICAL REQUIREMENTS .....	67
TABLE 16	ELSI TECHNICAL DESCRIPTION.....	81
TABLE 17	BCAT TECHNICAL DESCRIPTION.....	90

## List of Figures

FIGURE 1	INTEGRATION OF THE LOCAL BG SYSTEM WITH THE CENTRAL IBORDERCTRL SYSTEM .....	13
FIGURE 2	EXPECTED RISK DASHBOARD GRAPH FOR NEXT DAYS FOR BORDER MANAGERS .....	14
FIGURE 3	HIGH-LEVEL DATA FLOW, EXCHANGE AND INTERFACES FOR THE TRAVELLER USER APPLICATION (PRE-REGISTRATION PHASE) .....	15
FIGURE 4	HIGH-LEVEL DATA FLOW, EXCHANGE AND INTERFACES FOR THE BORDER GUARD USER APPLICATION (BORDER-CROSSING PHASE).....	16
FIGURE 5	HIGH-LEVEL DATA FLOW, EXCHANGE AND INTERFACES FOR THE BORDER MANAGER APPLICATION .....	17

FIGURE 6 TRAVELLER USER APPLICATION DATA FLOW .....	19
FIGURE 7 BGUA APPLICATION DATA FLOW.....	20
FIGURE 8 BMUA USER APPLICATION DATA FLOW.....	21
FIGURE 9 IBORDERCTRL DATABASE SCHEMA (“TRAVELLER TABLES”) .....	22
FIGURE 10 IBORDERCTRL DATABASE SCHEMA (“BORDER GUARD TABLES”).....	23
FIGURE 11 QR CODE DIAGRAM .....	31
FIGURE 12 CONSIDERATIONS FOR CLOUD SERVICE CHOICE.....	42
FIGURE 13 A TYPICAL MIGRATION PROJECT LIFE CYCLE.....	45
FIGURE 14 IBORDERCTRL MIGRATION PROCESS.....	46
FIGURE 15 SECURITY ASSESSMENT METHODOLOGY .....	49
FIGURE 16 PHASES IN PENETRATION TEST ANALYSIS.....	50
FIGURE 17 – RBAT – PRE-REGISTRATION PHASE.....	55
FIGURE 18 RBAT – BORDER CROSSING PHASE.....	56
FIGURE 19 VALUE TREE.....	58
<b>FIGURE 20 – RISKS OBJECTS GENERATION BASED ON THE TRAVELLER TABLE OF THE IBORDERCTRL DATABASE.....</b>	<b>64</b>
FIGURE 21 EXAMPLE OF RULE AUTHORIZING .....	65
FIGURE 22 RISK DATABASE SCHEMA .....	66
FIGURE 23 – RBAT INTERFACES TO THE IBORDERCTRL SYSTEM.....	67
FIGURE 24 TWITTER REST API FOR ELSI .....	70
FIGURE 25 TWITTER API VARIABLES.....	71
FIGURE 26 DIRECT FOLLOWERS RISK ASSESSMENT ELSI .....	71
FIGURE 27 MULTILEVEL FOLLOWERS RISK ASSESSMENT .....	72
FIGURE 28 UMF AS A LAYER BETWEEN SYSTEMS (SOURCE <sup>12</sup> ).....	80
FIGURE 29 THE POLICE INFORMATION MODEL (SOURCE ).....	80
FIGURE 30 UMF MAPPING (SOURCE ) .....	81
FIGURE 32 MARGIN DEFINITION .....	89
FIGURE 31 TUA INITIAL SCREEN .....	91
FIGURE 32 TUA REGISTER SCREEN.....	92
FIGURE 33 TUA LOGIN SCREEN .....	93
FIGURE 34 TUA HOME SCREEN.....	93
FIGURE 32 TUA TRAVEL INFORMATION SCREEN (STEP 1/5).....	94
FIGURE 33 TUA TRAVEL INFORMATION SCREEN (STEP 2/5).....	95
FIGURE 34 TUA DOCUMENT INFORMATION SCREEN (STEP 2/5) .....	96

FIGURE 35 – TUA VEHICLE INFORMATION SCREEN (STEP 2/5).....	97
FIGURE 36 – TUA DOCUMENT/VEHICLE UPLOAD SCREEN (STEP 3/5).....	98
FIGURE 37 TUA QR CODE SCREEN (STEP 5/5).....	99
FIGURE 38 TUA MY PROFILE SCREEN .....	100
FIGURE 39 – TUA MY DOCUMENTS SCREEN .....	101
FIGURE 40 – TUA MY TRIPS SCREEN .....	101
FIGURE 41 MAIN SCREEN OF THE PU APPLICATION.....	103
FIGURE 42 DATA ACQUISITION SCREEN FOR THE CAMERA.....	103
FIGURE 43 INITIAL LOGIN SCREEN OF THE PU BORDER GUARD AGENT USER INTERFACE.....	104
FIGURE 44 QR CODE CHECK SCREEN OF THE PU BORDER GUARD AGENT USER INTERFACE.....	104
FIGURE 45 DOCUMENT AUTHENTICITY CHECK SCREEN OF THE PU BORDER GUARD AGENT USER INTERFACE .....	105
FIGURE 46 VEHICLE INFORMATION CHECK SCREEN OF THE PU BORDER GUARD AGENT USER INTERFACE.....	105
FIGURE 47 MATCH FINGERPRINT CHECK SCREEN OF THE PU BORDER GUARD AGENT USER INTERFACE.....	106
FIGURE 48 FACE RECOGNITION INITIAL SCREEN OF THE PU BORDER GUARD AGENT USER INTERFACE.....	106
FIGURE 49 FACE RECOGNITION CHECK SCREEN OF THE PU BORDER GUARD AGENT USER INTERFACE .....	107
FIGURE 50 PALM VEIN CHECK SCREEN OF THE PU BORDER GUARD AGENT USER INTERFACE.....	107
FIGURE 51 HIDDEN HUMAN DETECTION CHECK SCREEN OF THE PU BORDER GUARD AGENT USER INTERFACE.....	108
FIGURE 52 OVERALL BCP SCREEN OF THE PU BORDER GUARD AGENT USER INTERFACE .....	108
FIGURE 53 LOST CONNECTION BETWEEN BODY MOUNTED CAMERA AND TABLET ERROR SCREEN OF THE PU BORDER GUARD AGENT USER INTERFACE .....	109
FIGURE 54 EXPECTED RISK DASHBOARD GRAPH FOR NEXT DAYS FOR BORDER MANAGERS .....	109
FIGURE 55 EXPECTED NUMBER OF TRAVELERS DASHBOARD GRAPH FOR BORDER MANAGERS.....	110

## Abbreviations

EEA	European Economic Area
DPA	Data Protection Authority
DPIA	Data Protection Impact Assessment
PIA	Privacy Impact Assessment
BGUA	Border Guard User Application
BMUA	Border Manager User Application
RIA	Research and Innovation Action



## Executive Summary

In this deliverable the on-going progress of the Development of the iBorderCtrl software platform and related interfaces (Work Package 4) is presented. As all tasks in WP4 are still ongoing until month 24, this report delivered at month 18 represents the first prototype version of developed software solutions that make up the iBorderCtrl platform. The next and final deliverable of this WP D4.2 to be delivered in m24 will represent the completed versions at the end of WP 4.

In this deliverable we begin with the iBorderCtrl System described from a technical point of view, together with the three user interfaces, that offer the user experience in the Traveller, Border Guard, and Border Manager Applications. An interface/ communication description and data flows are provided demonstrating the integration of the applications under the iBorderCtrl platform. The system data structure central to the iBorderCtrl is presented with explanation of the QR code key to the quick identification of travellers into the system, the underlying database and relevant technologies. The cloud infrastructure designed and set up for the platform in a way to enable scalability and robustness while ensuring security is presented next with a focus on services, the deployment model, open technologies relied on and the migration plan.

The design of the platform followed a security by design approach, this is presented along with the solution on how information is secured in transit, at the database as well as how Security constraints are documented and explored through the traceability matrix.

The major software systems developed in iBorderCtrl are presented; these are the RBAT, ELSI, BCAT modules and the three user interfaces. Each system is presented in terms of its technical design, functionality, and requirement coverage analyses.

The functionality of RBAT is presented demonstrating how the border managers (Risk authoring tool) would interact with it to set the rules and what information will be provided to the risk database. The method to derive the multi-criteria decision analysis steps -to determine the initial weights for each module- is presented. This will help iBorderCtrl function from day one, while it's deployed through the pilots and these weights are fine-tuned through the use of meta analyses on their actual performance.

The External Legacy and Social interfaces (ELSI) tool is presented, with a demonstration of how twitter information from will be incorporated from consenting travellers. A detailed explanation of how legacy databases will be incorporated utilizing the Universal messaging format (UMF) for SIS, VIS and EES.

BCAT is presented as the scientific workflow part of the project that will perform the meta analyses of collected data. The user interface and algorithms to be incorporated are presented. To help enhance the understanding of the potential outcome of this tool on the project and specifically to the border control authorities, scenarios are presented for each set of algorithms that are designed to demonstrate how BCAT will perform analyses and why this is expected to be impactful to border control and to the iBorderCtrl platform by enabling them to be adaptable.

Finally, the three user interfaces are presented this time with a focus on the graphical implementation with screenshots of the current versions and implementation details.

# 1 Introduction

This deliverable report is an interim report on the first version of the iBorderCtrl software platform and related interfaces developed as part of work package 4. The aim of the work package is to develop the project's software tools and relevant interfaces and more specifically:

- The data storage and communication infrastructure
- The risk-based assessment tool (RBAT) and the integrated automated border control analytics tool
- The External Legacy and Social Interfaces (ELSI)
- The central data repository to collect the data into a single environment
- The BCAT tool for meta analyses of data to enable a computationally intelligent adaptive border control.
- The user interfaces for the border control agent, the border control manager, and the traveller.

The data storage and communication infrastructure is based on a privacy by design enabled integration environment. Technical requirements were defined in WP2 of the project. The infrastructure was constructed with appropriate technical specifications, to be capable of supporting the large-scale content storage and processing requirements envisioned by a future Europe wide deployment of iBorderCtrl.

The RBAT tool although based on a pre-existing tool provided by partner ED was customised for the border control application and tailored to the requirements of iBorderCtrl. The authoring interface is presented and how RBAT is integrated with the project through the iBorderCtrl databases is presented through data flows.

ELSI's social component is implemented through the consideration of twitter as the social platform of choice. Twitter's API and user agreements met the requirements of the project. A tool was developed to communicate with twitter's API to recover data from consenting travellers for analyses. The legacy databases were focused on existing systems deployed at the border SIS and VIS as well as the Entry/Exit System (EES) that although not widely deployed is finalized and expected to be adopted -in the near future- by more countries. To achieve the communication, we relied on the Universal Messaging Format (UMF) communication procedures as per the description of Europol's technical documents.

BCAT enables the combinatorial analyses of all data collected in iBorderCtrl utilizing statistical, machine learning and data mining approaches to discover new patterns and knowledge that can be used through the RBAT tool to enhance the performance of the system. Furthermore, the tool will be used to evaluate all modules in iBorderCtrl through the analyses of data collected as part of the pilot phase. Finally, BCAT powers the real time generation of widgets that provide key information to border control managers (future risk and traffic prediction) as well as travellers (future waiting time).

The User interface for the Border Manager, Border Guard Portable Unit, as well as the traveller's pre-registration phase are presented with screen shots of the current development versions and all technical, design and requirement adherence information relevant to the project.

The data protection impact assessment is presented to demonstrate the on-going strict adherence to data protection requirements across the iBorderCtrl system.

## 2 IT infrastructure

### 2.1 iBorderCtrl System description

#### 2.1.1 System technical description

The iBorderCtrl system unifies the different interdisciplinary architectural components presented in D2.2 (Reference Architecture and Components Specifications) and D3.2 (First version of all technological tools and subsystems [redacted] and Avatar based dialogue, DAAT, BIO - fingerprints and palm vein - tool, FMT, HHD tool, Portable unit) and converges into an integrated system that provides to the traveller, the border guard and the border manager a useful tool to better and more effectively run their daily operations and working routines.

As explicitly described in D2.2, as far as the IT part is concerned, the iBorderCtrl system offers an integration environment to host mainly the iBorderCtrl Database (which consists of multiple tables as will be described in detail below) and specific software sub-systems concerning analytics and risk assessment, as well as the three applications to be handed over to all actors involved in the project; these are the Traveller User Application, the Border Guard User Application and the Border Manager User Application. The specific software subsystems of the integration environment, consisting of the rule-based risk analysis tool (RBAT), the border control analytics tool (BCAT) and the interfaces to the external and legacy systems (ELSI), will be presented as separate Chapters, due to their specific nature and development aspects.

In the following, in the framework of the present Chapter, the description of the Three User Applications and the main iBorderCtrl Database will be presented in detail, since they constitute the main interaction paths between all kind of systems users and the respective data storage and processing. The analysis that follows presupposes the connectivity and interfacing of the above main system components with the biometrics and sensor related modules [redacted]/avatar, DAAT, BIO, FMT and HHD) that are described in detail within D3.2.

##### 2.1.1.1 Traveller User Application (TUA)

The Traveller User Application will be responsible to manage the pre-registration procedure.

The application backend will store the information provided by the user to the iBorderCtrl database and will provide simplified access to the stored data. The traveller user application will include also the traveller user interface, which is the presentation layer (visualisation) of the application in a user friendly, self-explained and simple format for the traveller. Through this application, the travellers, following their consent, will be able to enter and update their personal information, upload travel related documents (such as VISA, passport) and travel information (tickets, hotel reservation, vehicles data etc.) and undertake the avatar interview. The application will send to the traveller a QR code upon completion of all required preregistration steps; when the QR code is handed over to the Border Guard at the BCP all information collected during the preregistration phase will be retrieved by the system in order to speed up the process.

From a technical point of view, the Traveller User Application - server-side application is implemented in a service-oriented approach using OSGi technology. The Open Services Gateway Initiative (OSGi) defines an architecture for developing and deploying modular applications and libraries. From a developer's perspective, OSGi offers the following advantages:

- Different modules of the application can be installed, uninstalled, start, and stop dynamically without restarting the container (see explanations below).
- The application can have more than one version of a particular module running at the same time.
- OSGi provides very good infrastructure for developing service-oriented applications, as well as embedded, mobile, and rich internet applications.

The Traveller User Application bundles are deployed using the Karaf OSGi container. Apache Karaf is a modern and polymorphic container powered by OSGi. Karaf can be used standalone as a container, supporting a wide range of applications and technologies. It also supports the "run anywhere" (on any machine with Java, cloud, docker images) using the embedded mode. Karaf OSGi container was chosen because it is a lightweight, flexible, powerful and is the perfect solution for systems integration.

The network communication between components is implemented using CXF REST services. CXF enables the development of RESTful services via annotations using the HTTP Binding. One advantage of using CXF is more flexibility in terms of deployment. It can be deployed in a web container or an embedded web container and run as a standalone application. CXF also provides integration with other frameworks like spring and also provides tools to work with schema / WSDL etc. The Database layer is build using the Hibernate ORM, and queries are executed in JPQL. As an Object/Relational Mapping (ORM) framework, Hibernate is concerned with data persistence as it applies to relational databases.

The Client-Side of the TUA is implemented using Ionic 3 and Angular 4. Ionic has the advantage to recognize the platform specific advanced CSS proportional to the native look and feel on different mobile operating systems. It reduces the requirement for code changing as it gives the codes of mobile-optimized HTML, JS, and CSS components. Aside from this, Ionic incorporates into AngularJS, which turns into a robust structure that makes code more manageable. With regards to creating mobile and web applications, AngularJS is a broadly preferred framework. The extensions to HTML's sentence structure offered by AngularJS are extremely useful for the design of mobile applications. The Ionic structure utilizes AngularJS to offer a bunch of centre functionalities to the designer with the goal that they can incorporate alluring components into the application. AngularJS is an ideal approach to make program-based applications, while with the assistance of the Ionic structure, versatile designers can make hybrid applications and web applications. As a result of the above, the client-side of TUA can be built for Web, Android, iPhone or Windows Phone.

### 2.1.1.2 Border Guard User Application (BGUA)

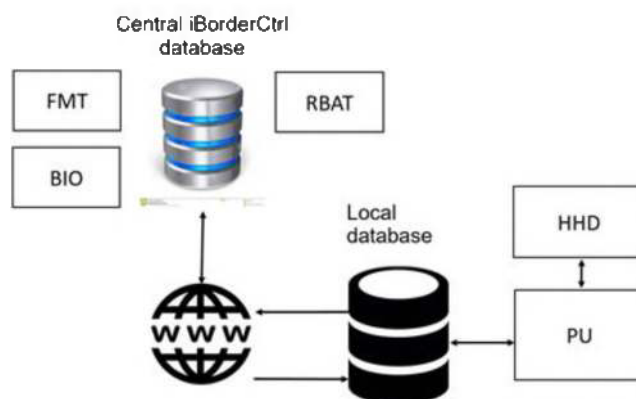
The Border Guard User Application will be responsible to manage the border-crossing procedure.

The Border Guard will interact with the iBorderCtrl platform through a dedicated web-based Border Guard User Application, which will be accessible and handled only by the Border Guards either through the portable unit or through a local terminal (i.e. PC) at the guards' booth at the BCPs. Through this application the Border Guards will be able to perform the Border Checks for each traveller crossing the BCPs: enter the travellers' and vehicles' documentation data during Border Checks, upload their travel related documents (such as VISA, passport, etc.), perform and upload scanners checks (fingerprints, palm vein, face matching etc.) and retrieve from the iBorderCtrl database the risk score per traveller. The BGUA interface workflow will be based on the "On-border" crossing point check general scenario that was described in the *WP2 Deliverable 2.2*.

The current advancement in the implementation of the Portable Unit (PU) involves the development of the application controlling devices and design of the harness playing the role of the frame for the hardware, worn by the Border Guard (BG). The system has three hardware levels, which cooperate in order to send the data to the central iBorderCtrl database. These are:

- **Local server**, responsible for the collection of data from the PU application from all Portable Units deployed per BCP. The data acquired from the devices will be stored **temporarily** and passed to each of the respective modules. After the border control procedure is concluded, these data can be deleted (as will be stored in the core of the system). In the server the **permanent** data will also be stored, referring to the details about the procedure: the date and time of starting and concluding the border checking procedure, identifier of the traveller / travel through the QR code and identifier of the BG checking the travellers, the GPS position of the BG, as well as the outcome of the checking procedure (passing or turning down the traveller). The local server will also be used to transmit the risk assessment between the central database to the PU.
- **Central Computing Unit (CCU)**, being the portable tablet / computer with the BG User Application. Here the control over the data acquisition devices is executed and the information about the risk assessment for the particular border checking procedure is presented and visualised. The procedure is performed using a set of screens, related with the particular acquisition stages. The respective GUI is also able to cooperate with the HHD tool, which directly provides its own individual risk score that is passed by the BGUA to the central iBorderCtrl database and the RBAT tool, contributing to the final risk assessment. The final risk score will be collected by the BGUA application and will assist the BG's final decision.

- **Data acquisition devices**, which must cooperate with the CCU. These data acquisition devices consist of all scanning devices that are used during the border checking procedure. The Integration of them with the BGUA application was the main advancement during this first period of the project.



*Figure 1 Integration of the local BG system with the central iBorderCtrl system*

### 2.1.1.3 Border Manager User Application (BMUA)

The Border Manager User Application will be responsible to manage all the information collected and stored, facilitating the BCP everyday work and will operate in the background.

The Border Manager User Application provides information concerning historical data as well as predictions of future days risk and traffic, providing statistics to assist the Border Managers on organizing and handling their everyday work. Consider for example decisions on the number of shifts and additional lanes to be used in case of heavy traffic or effective utilization of resources i.e. assignments of more experienced officers in days of high risk travellers expected. The BCAT tool is integrated within the BMUA to also allow the seamless expansion of the data analytics into novel scientific workflows making them directly available to border managers. The design of the Border Managers Application provides a transparent and seamless experience with all information of interest to them through a secure intranet portal. It combines detailed information in the form of a dashboard to the travel manager to help them plan ahead and more efficiently manage their resources in terms of expected traffic and risk, as well as it provides access to BCAT tools for meta-analyses and to the RBAT back-end for providing new rules to be instantly deployed in the field.

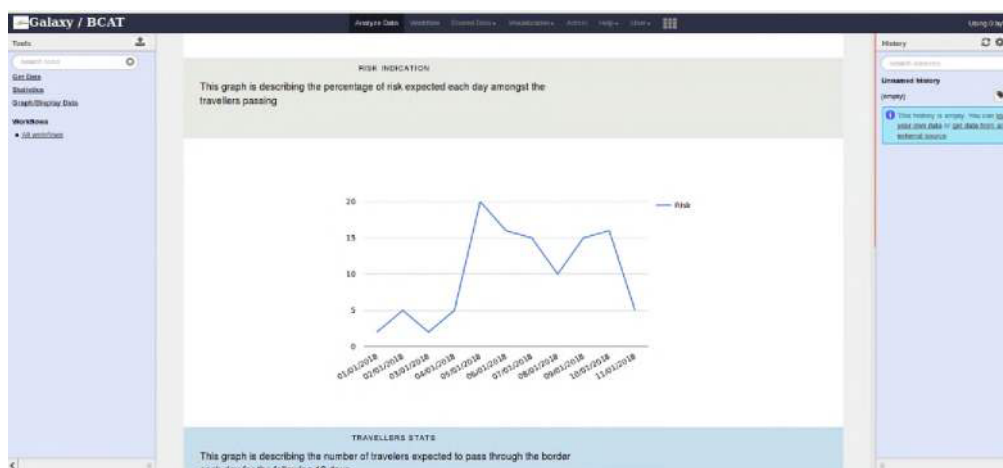
The BMUA needs to ensure scalability with potential deployment across all EU borders and a tremendous volume of data related both to upcoming travels (from pre-registration derived data of not yet border crossings) and those collected within the actual border crossings, as well as historical data collected with the scope of analysing them; to produce knowledge that would enhance automation, improve prediction accuracy and fine tune risk assessment based on empirically derived data. Furthermore, some of the analyses will need to run frequently to provide predictions about upcoming changes in overall border risks; therefore, ensuring scalability access to large high-performance computing resources is required.

To meet all those requirements, the BMUA application needs to rely on a strong scientific workflow system that would allow border managers to get access to information through the real-time analyses of data using state of the art analytical approaches. There is a clear need for a user-friendly interface that does not require any advanced analytical, programming or scripting expertise on the part of the border manager and that relies on a familiar web-based user interface capable of deploying computationally intensive tasks on secure cloud based resources, thus utilizing secure cloud based databases. To meet this task, we chose to repurpose the Galaxy Project<sup>1</sup>, an open, web-based platform for accessible, reproducible and transparent computational research; although initially designed for biomedical analyses, the Galaxy project has been also re-purposed

<sup>1</sup> <https://galaxyproject.org/>

across other fields, primarily due to its ability to provide a user-friendly approach to scientific workflows, therefore, enabling respective adjustments for the purposes of the iBorderCtrl project.

Dashboard type graphics (Figure 2) with real time analytics are implemented to present real time information to border managers that will empower them to manage their borders better, increasing their efficiency and reducing the cost. To achieve this, one of the novel features is the ability to utilize pre-registration data as well as historical data from past years to calculate relatively accurately both the number of people expected to cross the border at a specific time, as well as, the relative risk of each traveller.



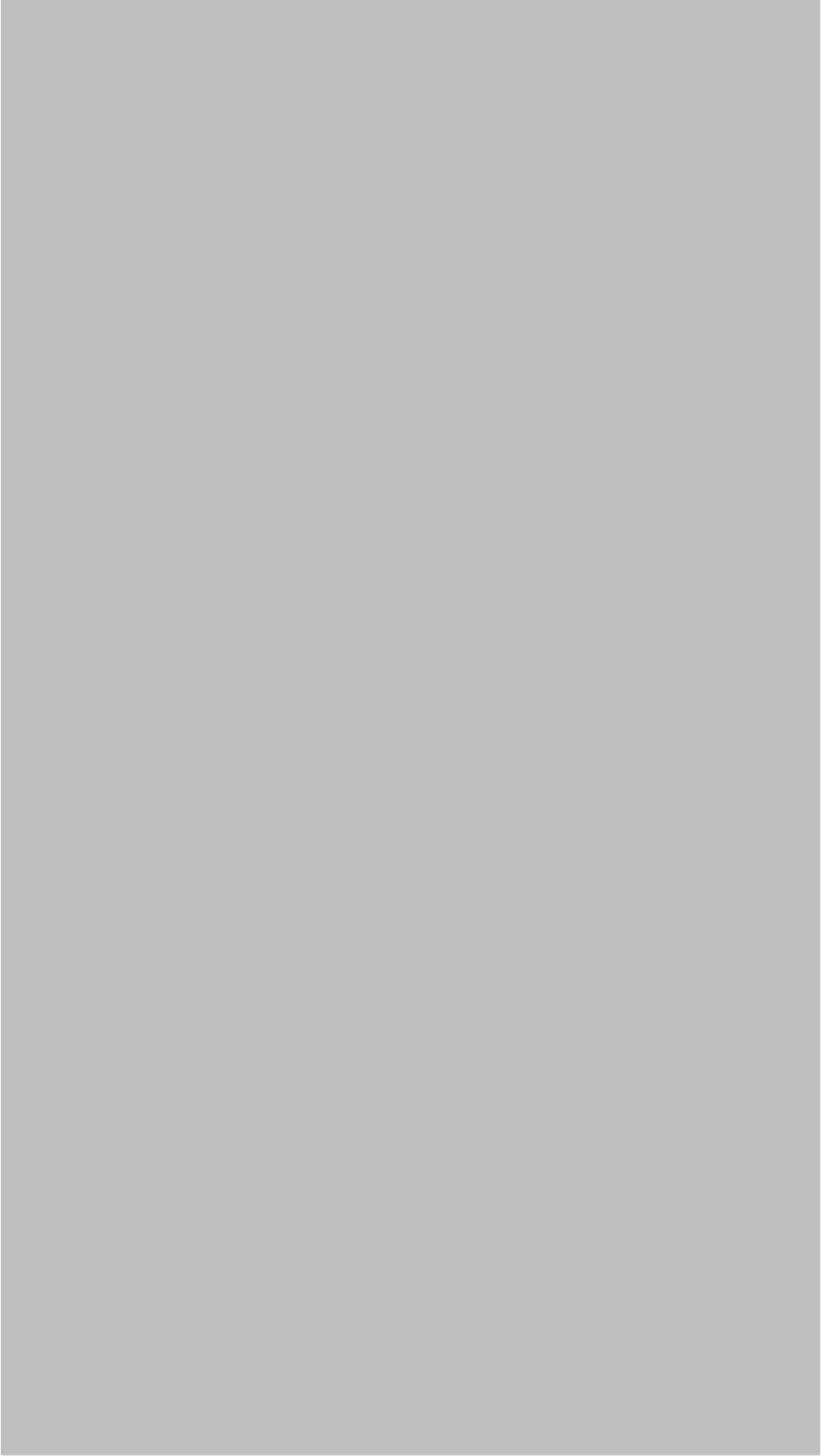
**Figure 2 Expected risk dashboard graph for next days for border managers**

### 2.1.2 System interfaces description - integration within the iBorderCtrl system

It is evident that the iBorderCtrl holistic platform with all its software and hardware components and dimensions is overall a complicated multi-disciplinary system; consisting of a blending of various types of data (text, video, images, dashboards etc.), various hardware and software technologies, scanning devices and advanced algorithms implemented along with various types of users involved with different needs each other. The overall architecture comprises the handling of various topologies and distributed processing along with local services at the BCPs (which are geographically distributed within the EU); all to be interconnected together through a cloud-based approach in the end.

In order to effectively tackle the inherent complexity of the relevant interactions and interfaces between all the above a step by step approach was initiated, starting from the high-level interactions among the various components in terms of their data exchange; as the development and software implementation further progresses, these interactions are being detailed in depth and will end-up in a format directly following purely technical software-development terms. At this particular project's stage, the first steps of this approach resulted in a detailed mapping (in the form of extended "technical responsibilities excel sheets") of the data exchanges and interfaces among the various components of the whole system per User Application; **in terms of "what is being consumed from and what is being provided to which"**, so that all interconnections between the various subsystems to be adequately described in terms of their interfaces, data description and data exchange to enable their appropriate integration, processing and handling.

For the sake of proper visualization and representation of all this extended information in a format comprehensive and understandable by the reader, the above detailed mapping has been "interpreted" in the following three diagrams, one per User Application. These diagrams represent in a visualised way, the **"high-level data flow, exchange and interfaces" per User Application** between the users, the various subsystems and the iBorderCtrl platform, mainly the iBorderCtrl database which is the core of the system. A **colour code**, explained in the upper left of each diagram, indicates the various types of data exchanged, while the main functions of the overall system are highlighted. Then, the detailed dataflow for each User Application separately will be described in the sections to follow these three general diagrams.



*Figure 3 High-Level Data flow, exchange and interfaces for the Traveller User Application (pre-registration phase)*



#### D4.1 First version of the iBorderCtrl software platform







***Figure 5 High-Level Data flow, exchange and interfaces for the Border Manager Application***

## 2.1.3 System data flow description

### 2.1.3.1 Traveller User Application

In order to perform the pre-registration of the iBorderCtrl system, the traveller with the use of a camera equipped mobile phone / tablet / laptop / PC enters the traveller user interface of the traveller user application (TUA). The initial screen welcomes the user to the iBorderCtrl preregistration procedure and briefly presents the project objectives. A text provides information to the traveller regarding the data which will be required during the procedure and asks for the traveller's informed consent (checkbox). The TUA will require the informed consent of the traveller in order to continue (ticking the checkbox).

Subsequently, the traveller is given two options: to register to the system or to log in to the system. If it is the first time using the iBorderCtrl system, the traveller should first register to the system, providing some initial required information (i.e. name, surname, gender, user name, password etc.). The TUA will verify that the user has entered all the obligatory information, whether the user name entered by the traveller is available and whether the password is strong enough. Once the user successfully completes the above step, TUA will store all provided information to the iBorderCtrl database. Moreover, the TUA will send an email to the traveller with a link for activation/verification of the requested account and a link for cancelling account request. If the traveller decides not to activate the account, TUA will revoke the whole procedure and will delete all information stored in the database. If the traveller decides to activate the account, TUA will activate the relevant user account and will display the login screen where the traveller will have to enter the correct credentials.

TUA will set access permissions to the traveller as long as the provided credentials match the ones in the database. The traveller once logged into the TUA, can decide to register a new trip, to view his/her profile or old trips, to see statistics or traffic information regarding planned trips or to provide feedback for the whole procedure. In the latter case, the TUA will store the user feedback to the iBorderCtrl database. If the traveller decides to register to a new trip, first of all he/she will have to add all the countries that he/she will cross until the final destination is reached. Following this procedure, the traveller is requested:

- to fill in travel-related information for every country he/she will visit (i.e. length of stay, purpose of trip, expected date of arrival at the borders etc.)
- to declare the travel document/s (passport or ID, visa, residence permit) he is going to use in each country and to enter respective information about these travel documents
- to follow the instructions in order to take a photo of the above registered travel documents using the camera of his/her mobile phone / tablet / laptop / PC
- to declare if he/she is going to use a private vehicle during the trip and enter relative information (i.e. license plate, ownership, driver license number etc.)
- to follow the instructions in order to take a photo of the driver's licence using the camera of his/her mobile phone / tablet / laptop / PC.

TUA will verify that the traveller has entered all the obligatory information and will store all information entered by the traveller (mentioned above) to the iBorderCtrl database.

TUA will generate the QR code which includes the traveller user and travel ID. The QR code associated with the RBAT risk result for the pre-registration procedure will be stored in the iBorderCtrl database. The generated QR code will be also sent via email to the traveller and also displayed on the screen of the traveller's mobile phone / tablet / laptop / PC. The traveller will have to download the QR code in order to present it at the border guards when crossing the border.

It should be noted that throughout the pre-registration procedure, the traveller is going to be able to revoke the whole procedure and withdraw his/her consent. In this case, TUA will immediately delete all information stored in the iBorderCtrl database.

*Figure 6 Traveller User Application Data flow*

### **2.1.3.2 Border Guard User Application**

In order to perform Border Check-in Procedure using equipped tablet, the Border Guard needs to pass login and password to the BGUA, the initial screen will welcome the user and inform about connection and battery level status of Portable Unit equipment: QR code reader, camera, document reader, fingerprints and palm vein reader. Not registered users will have to create an account and provide the system with required information such as: Name, Surname, border guard id number, e-mail address etc. The BGUA will verify if all required information was provided, if entered login is available or password is strong enough. The BGUA will send the verification email message on the provided e-mail address, in order to authorize the registration process.

The push button START will begin the check procedure. The BGUA will instruct the officer what exactly device is needed to perform each procedure and warn if any extraordinary situations appear. After each procedure of data acquisition, the BGUA will display gathered data -related to each check- as well as any requested additional traveller information and estimated risk score. The HHD tool is foreseen to be connected in the end, providing directly its own risk score. During check procedure the BGUA backend will be updating the local database, in order to provide rest subsystems with the required information, passing it to the main iBorderCtrl database and returning back the final risk score after the RBAT calculations.

After each task is completed the BGUA will display the final risk score of the traveller at border crossing and the Border Guard will have to pass to the system whether his/her final decision is positive or negative.



*Figure 7 BGUA application data flow*

### 2.1.3.3 Border Manager User Application

In order to have access to Border Manager User Application, an authorized person needs to provide a verified username and password. A successful verification leads the user to enter into the system from which -based on the access control of the account- a certain access of use is provided. The interface page will display information in the form of a dashboard (graphics and tables). It will be possible to engage the functionalities of BCAT directly from BMUA to perform targeted analytics using scientific workflows based on the user's access rights. The outcomes of these requests are provided either with graph representations or with a simple tabular response as required.

The dashboards that are the first information presented following successful log in by a border manager are calculated automatically providing weekly risk scores and traffic information based on the iBorderCtrl database pre-registrations. These are designed to help the border manager identify potential changes in traffic or risk that she/he should address by adjusting the schedule and planning of her/his border crossing point. The dashboard will be in place to give numbers as to the travellers expected daily and in more advance requests number of high risk travellers estimated to arrive etc. It provides essentially an application where border guard can visualize what to expect and properly prepare for it.

Beside that helpful information and the friendly environment created, the Border Manager User Application also includes BCAT tools where the main purpose is the use of statistical packages and data mining/machine learning algorithms to perform advanced analytics of the data in the iBorderCtrl database allowing for: the discovery of key knowledge that could act as new rules, or updated weights and thresholds for existing rules that the border managers after review may include in their own instantiation of RBAT to increase their efficacy and be able to introduce risk patterns, make use of the plethora of raw datasets, and transform them into valuable information.



*Figure 8 BMUA User Application Data flow*

## 2.1.4 System data structure

### 2.1.4.1 Database tables

The iBorderCtrl database stores all the information provided by the travellers during the pre-registration phase and the Risk scores (or additional information about the check outcome of the processed data) provided by each iBorderCtrl module during the pre-registration and the border crossing. The respective tables and schema of the iBorderCtrl database are presented in the following figure.

***Figure 9 iBorderCtrl database schema (“traveller tables”)***



**Figure 10 iBorderCtrl database schema (“border guard tables”)**

### 2.1.4.2 Dictionary of iBorderCtrl Database

Each field presented in the iBorderCtrl database schema presented in Figure 6 is explained in the tables below:

#### UserAccount Table

Field	Explanation	Remarks
<b>Type</b>	The type of user is going to be stored [traveller, Border Guard (Officer), Border Manager]	
<b>login_name</b>		The traveller enters his own user name and password, the System Administrator for Border Manager and the Border Manager for Border Guards.
<b>password</b>		Same as above
<b>status</b>	Active, disabled, authorized/unauthorised etc.	
<b>created_on</b>		
<b>modified_on</b>		

**Traveller Table**

Field	Explanation	Remarks
<b>Name</b>	TUA stores the name as entered by the traveller.	
<b>Surname</b>	TUA stores the surname as entered by the traveller.	
<b>Gender</b>	TUA stores the gender as entered by the traveller.	
<b>Date_birth</b>	TUA stores the date of birth as entered by the traveller.	
<b>Nationality</b>	TUA stores the nationality as entered by the traveller.	
<b>Country_residence</b>	TUA stores the country of residence as entered by the traveller.	
<b>address</b>	TUA stores the address of the traveller as entered.	
<b>Mobile_telephone</b>	TUA stores the mobile telephone as entered by the traveller.	
<b>email</b>	TUA stores the email of the traveller as entered.	
<b>Twitter_Account</b>	TUA stores the Twitter account of the traveller as entered.	
<b>Citizenship</b>	TUA stores if the traveller has single or multiple citizenship. (Options: Single, Multiple)	

**InitialTravellInfo Table**

Field	Explanation	Remarks
<b>Origin</b>	TUA stores the origin country of the trip as entered by the traveller.	
<b>Destination</b>	TUA stores all countries that the traveller goes through during his travel as entered by the traveller.	
<b>length_of_stay</b>	TUA stores the number of days the trip will last as entered by the traveller.	
<b>contact_information</b>	TUA stores the contact information during the trip (relatives, friends etc.) if provided by the traveller.	
<b>hotel_reservation</b>	TUA stores the hotel reservation if any and if provided by the traveller.	
<b>expected_date_arrival</b>	TUA stores expected date of arrival as entered by the traveller.	
<b>expected_time_arrival</b>	TUA stores expected time of arrival as entered by the traveller.	
<b>expected_date_departure</b>	TUA stores the expected day of departure as entered by the traveller	



<b>rating_of_procedure</b>	TUA stores the rate of procedure as entered by the traveller	1-5 stars evaluation at the end of the TUA procedure
<b>review_of_procedure</b>	TUA stores review of the whole procedure as entered by the traveller in text form.	Free-text evaluation at the end of the TUA procedure
<b>Custom Sponsor</b>	TUA stores who is paying for the travel expenses as selected by the traveller (Options: Employer, My Self, My Family, Other)	
<b>Custom Transport</b>	TUA stores the means of transportation to be used during the travel as selected by the traveller (Options: Car, Train, Ferry, Plane, On Foot, Lorry, Coach, Other)	
<b>Custom Purpose</b>	TUA stores the purpose of the traveller's trip as selected (Options: Business, Education, Vacation, Visit Family, Find Employment, Attend Conference, Other)	
<b>Refused_entry</b>	TUA stores the Y/N answer of the traveller whether he/she has ever refused entry to the Schengen area.	
<b>Removed</b>	TUA stores the Y/N answer of the traveller whether he/she has ever removed from the Schengen area	
<b>Health_insurance</b>	TUA stores this Y/N field if the traveller has health insurance.	
<b>Health_Country</b>	TUA stores the country where the traveller has health insurance to (if above field is Y).	

**Travel table**

Field	Explanation	Remarks
<b>qrcode</b>	TUA stores the generated QR code upon completion of the pre-registration procedure	
<b>rating_of_procedure</b>	TUA stores the rate of procedure as entered by the traveller	1-5 stars evaluation at the end of the TUA procedure
<b>review_of_procedure</b>	TUA stores review of the whole procedure as entered by the traveller in text form.	Free-text evaluation at the end of the TUA procedure

**Traveller Doc Table**

Field	Explanation	Remarks
<b>doc_number</b>	TUA stores the number of the traveller document as entered by the traveller	

<b>doc_issuing_office</b>	TUA stores the issuing office of the travel document as entered by the traveller	
<b>doc_issue_date</b>	TUA stores the issuing date of the travel document as entered by the traveller	
<b>doc_exp_date</b>	TUA stores the expiry date of the travel document as entered by the traveller	
<b>doc_country</b>	TUA stores the country who issued the travel document as entered by the traveller.	
<b>doc_status</b>	TUA will store the status of the document: (expired or valid) as entered by the traveller	
<b>doc_photo</b>	DAAT is going to store the photo of the whole passport page	The same is going to apply also for all travel documents (visa, id etc.)

#### Document Table

Field	Explanation	Remarks
<b>document_type</b>	TUA stores the type of travel document (passport, visa, id, residence permit)	
<b>document_status</b>	TUA will store the status of the document: (expired or valid)	

#### VehicleDoc Table

Field	Explanation	Remarks
<b>license_plate</b>	TUA stores the license plate of the vehicle to be used during the trip as entered by the traveller	
<b>Insurance_policy</b>	TUA stores the insurance policy number of the vehicle to be used during the trip as entered by the traveller	
<b>ownership</b>	TUA stores the name of the owner of the vehicle to be used during the trip as entered by the traveller	
<b>driver_license</b>	TUA stores the number of the driver license as entered by the traveller	
<b>Issuing_country</b>	TUA stores the issuing country of the driver license as entered by the traveller	
<b>Licence_photo</b>	TUA stores the drivers licence photo as uploaded by the traveller	

#### BORDER GUARD Tables

Field	Explanation	Remarks
<b>officer_id</b>	Officer's record identifier	
<b>login</b>	Officer's login to the BGUA	The officers enters his own user name and password

<b>password</b>	Officer's password to the BGUA	Same as above
<b>name</b>	Officer's name	
<b>surname</b>	Officer's surname	
<b>mobile_phone</b>	Officer's mobile phone number	
<b>email</b>	Officer's email	
<b>created_at</b>	Date of the account creation	
<b>last_login</b>	Date of the last login to the BGUA	
<b>officers_additional_info</b>	Filed for special characters/features of the officer	
<b>border_crossing_country</b>	Operation country for border guard	id of the country, used also by COUNTRY table
<b>border_crossing_designation</b>	Operation designation for border guard	id of the designation, used also by DESIGNATION

**COUNTRY Table**

Field	Explanation	Remarks
<b>country_id</b>	Country identifier	
<b>country</b>	Country	e.g. Poland-Ukraine

**DESIGNATION Table**

Field	Explanation	Remarks
<b>designation_id</b>	Designation identifier	
<b>designation</b>	Designation	e.g. Medyka-Shehyni

**BORDER GUARD SESSION Table**

Field	Explanation	Remarks
<b>session_id</b>	Session identifier	used also by BORDER CONTROL PROCEDURE table
<b>officer_id</b>	Officer id that performs the control session	
<b>start_shift_time</b>	BGUA stores the time of session begin	
<b>end_shift_time</b>	BGUA stores the time of session finish and time	

<b>start_shift_date</b>	BGUA stores the date and time of session begin	
<b>end_shift_date</b>	BGUA stores the date of session finish and time	
<b>number_of_controls</b>	BGUA stores the amount of completed controls during session	
<b>start_battery_level</b>	BGUA stores the start battery level	
<b>end_battery_level</b>	BGUA stores the final battery level	

**BORDER CONTROL PROCEDURE Table**

Field	Explanation	Remarks
<b>control_id</b>	Control identifier	used also by PICTURES table
<b>session_id</b>	Session identifier	
<b>start_time</b>	BGUA stores the time that control begin	
<b>end_time</b>	BGUA stores the time that control finish	
<b>start_date</b>	BGUA stores the date that control begin	
<b>end_date</b>	BGUA stores the date that control finish	
<b>travel_id</b>	BGUA stores the travel id based on QR code provide by traveller during control procedure	
<b>user_id</b>	BGUA stores the user id based on QR code provide by traveller during control procedure	
<b>qr_code_risk_score</b>	BGUA stores the pre-registration risk based on QR code provide by traveller during border control procedure	
<b>daat_risk_score</b>	BGUA stores the risk provided by DAAT	
<b>bio_risk_score</b>	BGUA stores the risk provided by BIO (palm vein)	
<b>fmt_risk_score</b>	BGUA stores the risk provided by FMT	
<b>fingerprint_risk_score</b>	BGUA stores the fingerprint match score	
<b>hhd_risk_score</b>	BGUA stores the hidden humans detection score	
<b>final_risk_score</b>	BGUA stores the final risk score provided by RBAT	

<b>pass_nopass</b>	BGUA stores the officer decision	
<b>palm_vein_timestamp</b>	BGUA stores the palm vein timestamp	
<b>bcp_localization</b>	BGUA stores GPS localization of border crossing control	

**BCP TEMPORARY DATA Table**

Field	Explanation	Remarks
<b>control_id</b>	Control identifier	used also by PICTURES table
<b>qr_code</b>	BGUA stores the QR Code pattern provided by the traveller during border crossing procedure	
<b>doc_type</b>	BGUA stores the type of the travel document provided by the traveller during border crossing procedure	
<b>doc_number</b>	BGUA stores the number of the travel document provided by the traveller during border crossing procedure	
<b>doc_issuing_office</b>	BGUA stores the issuing office of the travel document provided by the traveller during border crossing procedure	
<b>doc_issue_date</b>	BGUA stores the issuing date of the travel document provided by the traveller during border crossing procedure	
<b>doc_exp_date</b>	BGUA stores the expiry date of the travel document provided by the traveller during border crossing procedure	
<b>doc_country</b>	BGUA stores the country who issued the travel document provided by the traveller during border crossing procedure	
<b>doc_status</b>	BGUA stores the status of the document: (expired or valid) provided by the traveller during border crossing procedure	
<b>doc_photo</b>	BGUA stores the photo of the whole passport page provided by the traveller during border crossing procedure	
<b>rfid_fingerprints_pattern</b>	BGUA stores the fingerprints pattern fetched from the RFID biometric document provided by the traveller during border crossing procedure	

<b>sensor_fingerprints_pattern</b>	BGUA stores the fingerprints pattern of the traveller fetched from the biometric sensor during border crossing procedure	
<b>rfid_face_picture</b>	BGUA stores the face picture of the traveller fetched from the portable digital camera during border crossing procedure	

#### PICTURES Table

Field	Explanation	Remarks
<b>picture_id</b>	Picture identifier	few pictures per traveller
<b>control_id</b>	Control id to which picture belongs	
<b>face_picture_photo</b>	BGUA stores the face photo of the traveller	
<b>best_quality</b>	Is it the photo with best quality per traveller?	

#### 2.1.4.3 Database technologies

For the purposes of the iBorderCtrl project and since:

- the kind of data that will be stored is relation data,
- big-data analytics will not be performed,
- spatial data will not be stored,
- engine should offer support forums,
- marketplace trends (databases popularity) should be balanced,
- very large tables will be avoided

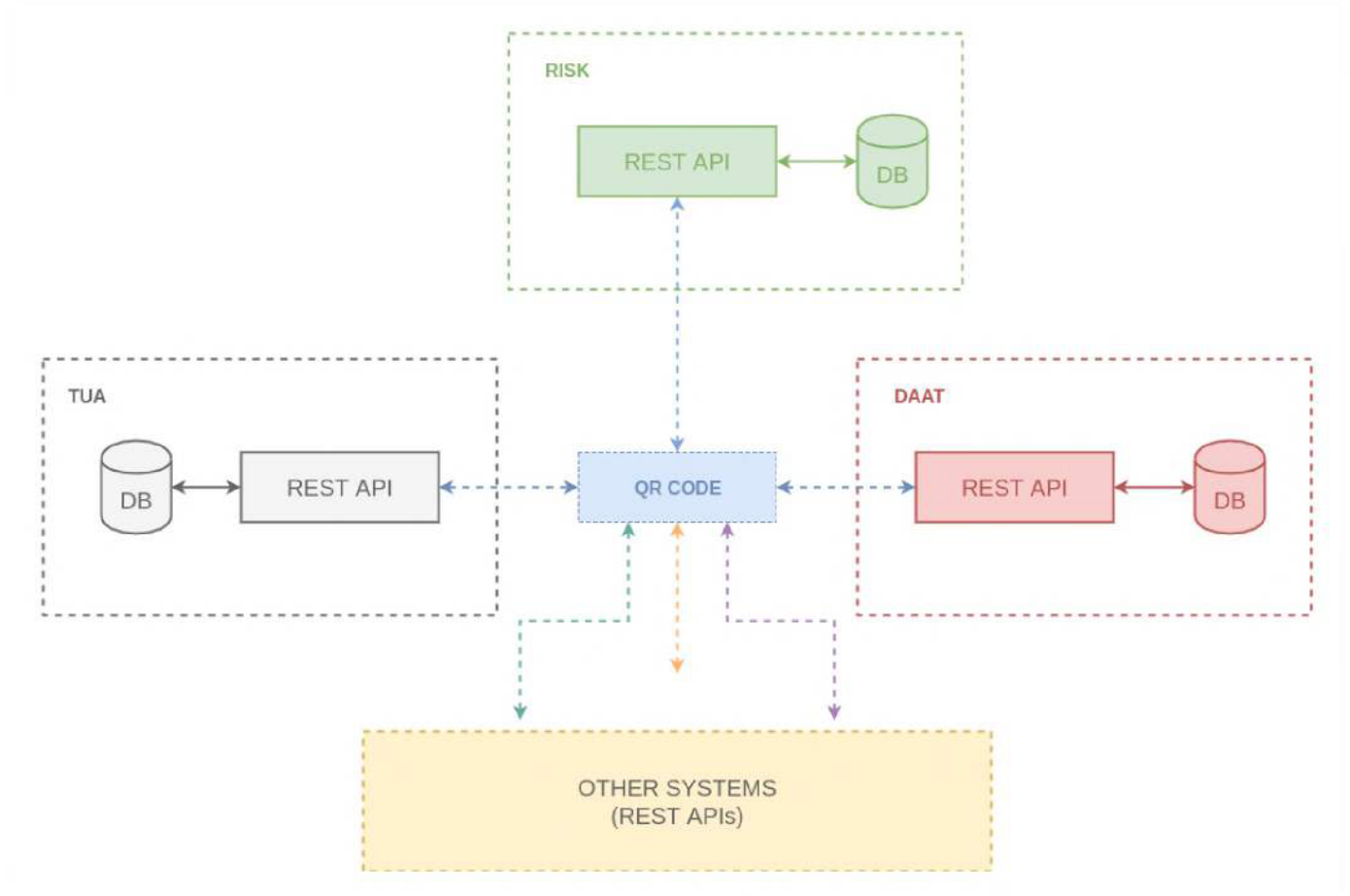
MySQL 5.7.X and its InnoDB engine will be used. The MySQL Database is a full-featured RDBMS and powers the most demanding Web, E-commerce and Online Transaction Processing (OLTP) applications. It is a fully integrated transaction-safe, ACID compliant database with full commit, rollback, crash recovery and row level locking capabilities. MySQL delivers the ease of use, scalability, and performance that has made MySQL the world's most popular open source database. Some of the world's most trafficked websites like Facebook, Google, Twitter, Uber, and Booking.com rely on MySQL for their business-critical applications.

#### 2.1.4.4 QR code

For the operation of the entire iBorderCtrl system and its different application, the connecting point for all information transfer (that also enhances the security of the system and introduces the pseudonymization privacy-enhancing technique) is the QR code. **The traveller's QR-Code will contain only the traveller\_id and the travel\_id as well.** None of the traveller's Personally Identifiable Information (PII) will be included in the QR code and as a result the traveller identity cannot be revealed through the QR code. This makes the communication between the iBorderCtrl modules easier and secure. Each of the subsystems will be able to get the information which they want using the QR-Code provided to them by the Traveller User Application.

The QR-Code is generated using the com.google.zxing library. ZXing library boasts the best maintenance record and is also the most popular one. It is a Google project from where is receiving regular updates. It consists of a core Java library that handles the decoding for all platforms, and a number of platform-specific

apps (including Android, Glass, Java SE, and the web). The Android application in the repository- is the application known simply as Barcode Scanner.

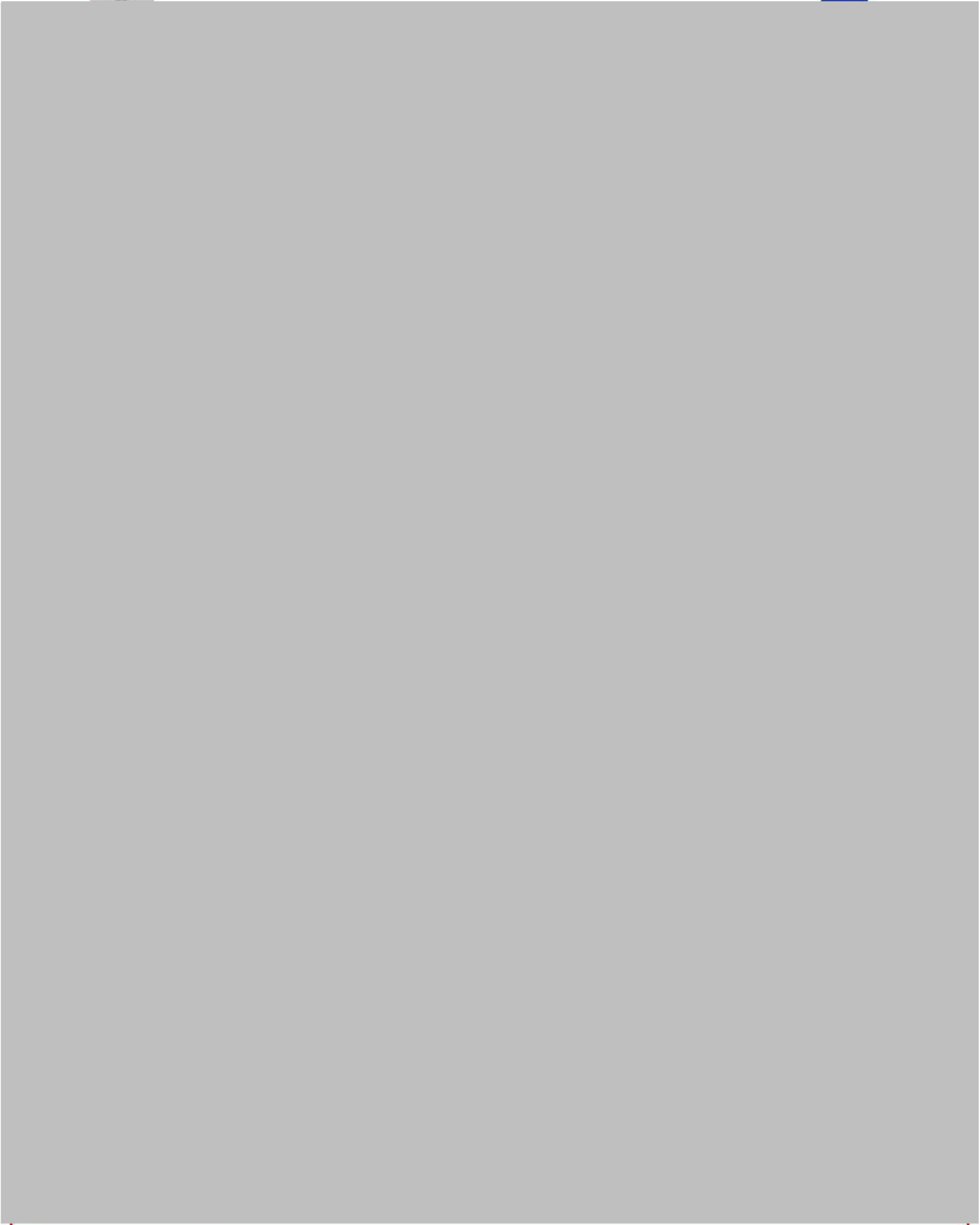


**Figure 11 QRcode diagram**

### 2.1.5 Technical requirements coverage

*Note: The BCAT and BMUA associated requirements have been inserted in the following Tables depending on the phase they address to.*

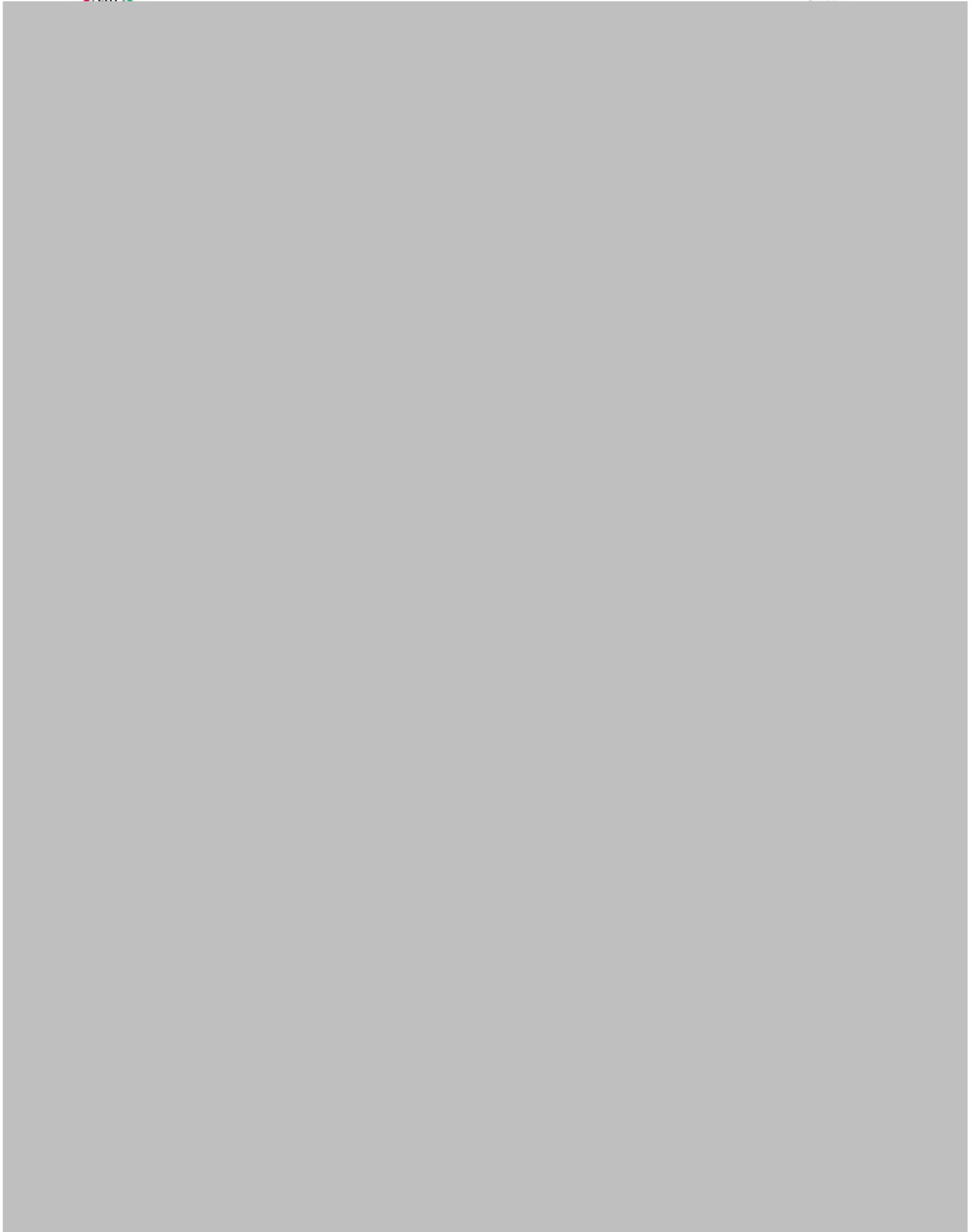
**Table 1 Travellers' User Application Technical requirements**







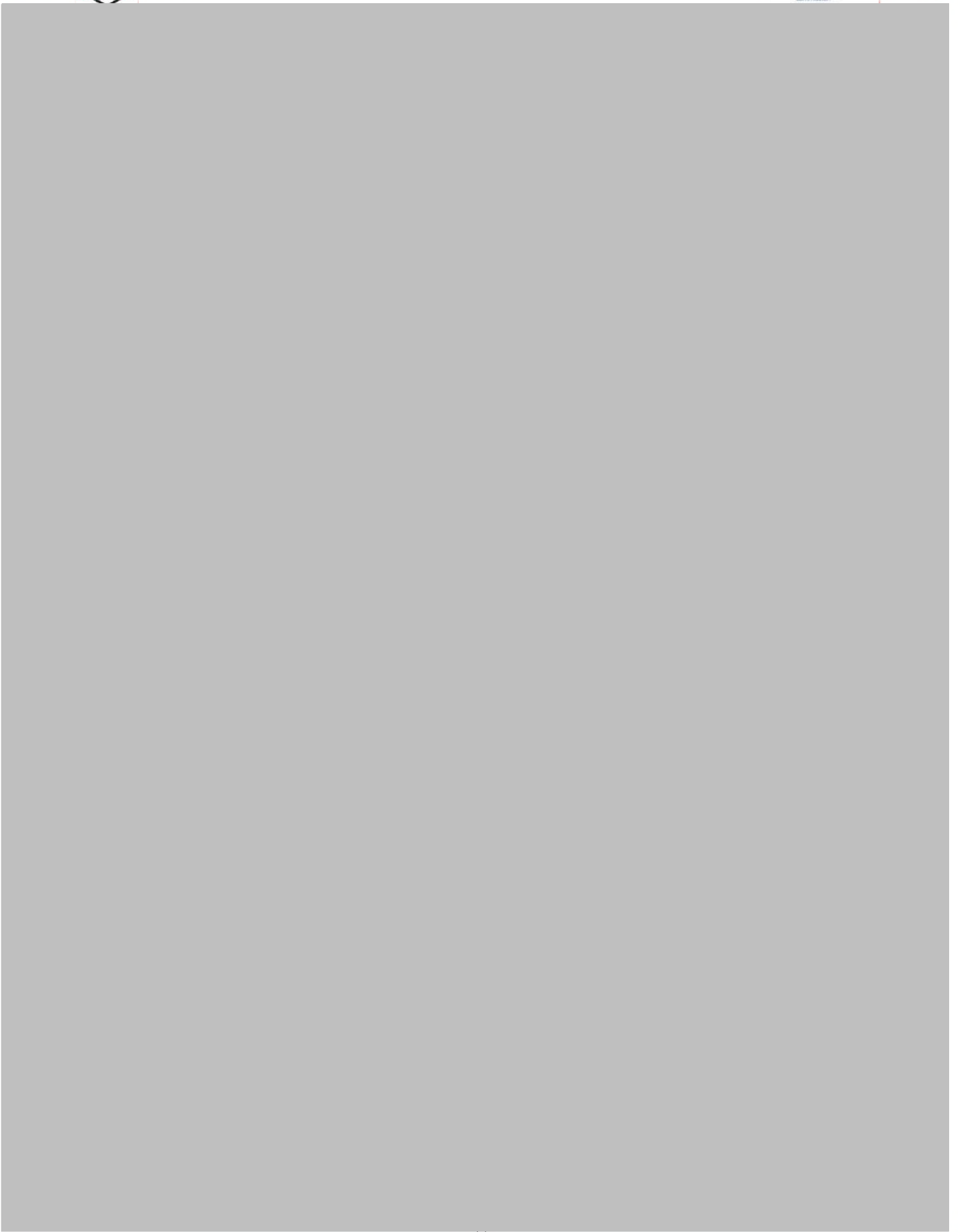






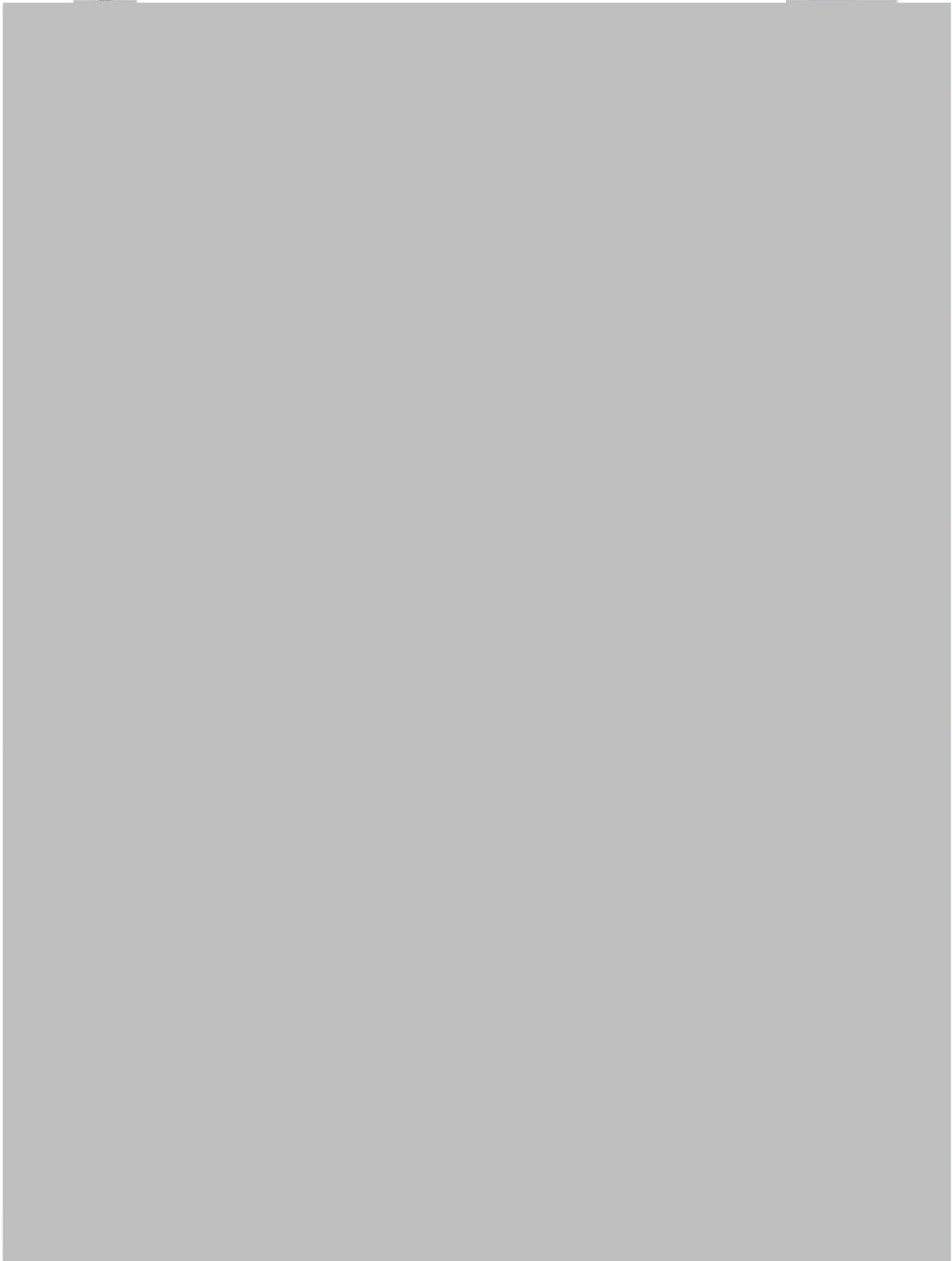


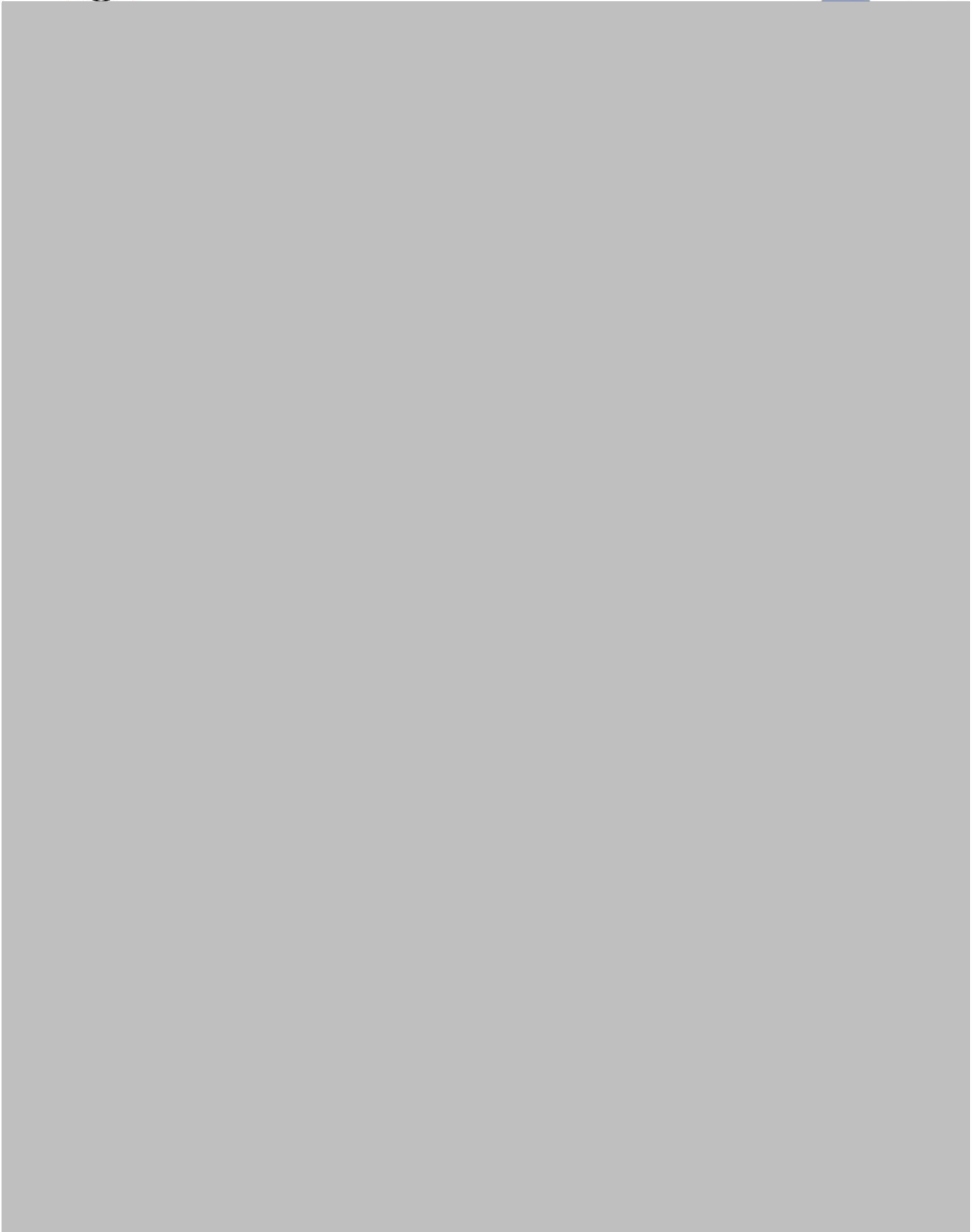
## D4.1 First version of the iBorderCtrl software platform





## D4.1 First version of the iBorderCtrl software platform











## 2.2 Cloud Infrastructure

### 2.2.1 Software stack implementation

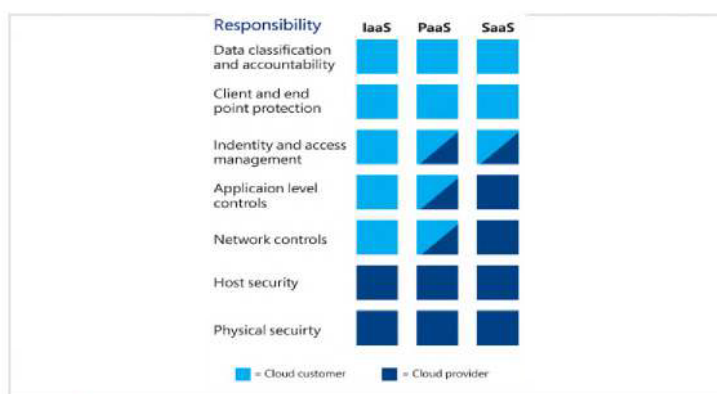
“The iBorderCtrl Platform will present a distributed computing architecture whose resources should be hosted on a cloud-based infrastructure. The goal is, when the platform reaches its full development status, to be directly implemented on a cloud-based infrastructure whereas the cloud provider to be chosen will provide the required virtual machines for the different components to perform, as well as a secure connection to the virtual machine (using SSH) for installing the aforementioned software modules and applications. Towards this goal, the iBorderCtrl project is already in the process of assessing options for cloud-based resources provision in order to result in a scalable and robust solution that will adequately meet the relevant security, storage and data processing along with the data privacy requirements defined in Chapter 4. Based on this roadmap set herein, the iBorderCtrl project will further refine the relevant cloud requirements in the framework of WP4 (task 4.1) in order to fully proceed in the implementation of the iBorderCtrl platform on a cloud-based infrastructure.”<sup>2</sup>

#### 2.2.1.1 Cloud computing services

Following the detailed description of the software stack as presented in Section 6.4 of D2.2, an assessment has been conducted between the different service categories of cloud computing<sup>3</sup>:

- **Software-as-a-Service (SaaS):** The capability provided to the consumer is to use the provider’s applications running on a cloud infrastructure. The applications are accessible from various client devices through either a thin client interface, such as a web browser (e.g., web-based email), or a programming interface. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, storage, or even individual application capabilities, with the possible exception of limited user-specific application configuration settings.
- **Platform-as-a-Service (PaaS):** The capability provided to the consumer is to use the provider’s development platform (programming languages, libraries, services, and tools) in order to create, test and host new applications. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, or storage, but has control over the deployed applications and possibly configuration settings for the application-hosting environment.
- **Infrastructure-as-a-Service (IaaS):** The capability provided to the consumer is to provision processing, storage, networks, and other fundamental computing resources in order to build a customized computing environment. The consumer does not manage or control the underlying cloud infrastructure but has control over operating systems, storage, and deployed applications; and possibly limited control of selected networking components (e.g., host firewalls).

The figure on the right depicts the various control responsibilities that cloud customers and providers have in IaaS, PaaS, and SaaS environments<sup>4</sup>.



**Figure 12 Considerations for cloud service choice**

<sup>2</sup> “D2.2 Reference Architecture and Components Specifications”, iBorderCtrl project Deliverable

<sup>3</sup> The Top Open Source Cloud Projects of 2014, viewed November 9, 2015 <<https://goo.gl/HQOutA2>>

<sup>4</sup> Transforming Government – Cloud policy framework for innovation, security, and resilience, 2015, Microsoft, viewed November 9, 2015 <<http://goo.gl/Tjz1jh>>

**The iBorderCtrl approach:** The iBorderCtrl platform enhances the IaaS solution with two modules that provide the high-availability and scalability features in a way that is transparent to the application owners. By accompanying the IaaS layer with the Data Service layer and Access layer, the data and the HTTP traffic management are delegated to the platform while the application’s business logic is still contained on the VM(s). This approach offers great flexibility as it does not require architectural changes to the applications but also keeps the deployment complexity low because the application owner “leverages” the high-availability and scalability features of the platform. The only drawback of this solution comparing with the SaaS is that the application owners are not entirely independent from platform administrators as the latter should configure the high-availability and scalability features per application.

The following table summarises the different application migration options:

Option	Description	Pros	Cons
Full IaaS	All the application components are deployed on VM(s) explicitly managed by the application owner	+ No architectural change of the application + Full control on the resources used for the deployment	- High deployment complexity because the application owner must take care of installing and configuring all the components for high availability and scalability
IaaS + Data Service Layer + Access Layer	Data and HTTP traffic management are handled by the platform, while the application business logic is still deployed on VM(s)	+ No architectural change of the application + Less deployment complexity because the application owner ‘leverages’ the high-available and scalable features of the platform layers	- Because of the centralized administration of the shared functions (e.g. data service layer), application owners cannot deploy their applications in full autonomy
PaaS + Data Service Layer	Applications are hosted on the PaaS Layer and use the Data Service Layer for storing data	+ No infrastructure management required by the user (the platform does it for her)	- Applications may require significant changes to comply with PaaS principles

**Table 4 Different application migration options supported by iBorderCtrl Platform.**

### 2.2.1.2 Cloud Deployment Model

Public Authorities should consider, when they plan their Cloud strategy, the different Deployment Models of Cloud Computing:

- **Private Cloud:** The cloud infrastructure is used exclusively for internal applications within an organisation comprising multiple consumers (e.g., business units). It may be owned, managed, and operated by the organization, a third party, or some combination of them, and it may exist on or off premises.
- **Community Cloud:** The cloud infrastructure is used exclusively by multiple organizations that have shared concerns (e.g., mission, security requirements, policy and compliance considerations). It may be owned, managed, and operated by one or more of the organizations in the community, a third party, or some combination of them, and it may exist on or off premises.
- **Public Cloud:** The cloud infrastructure is provisioned for open use by the general public. It may be owned, managed, and operated by a business, academic, or government organization, or some combination of them. It exists on the premises of the cloud provider.

- **Hybrid Cloud:** The cloud infrastructure is a composition of two or more distinct cloud deployment models (private, community, or public) that remain unique entities, but are bound together by standardized or proprietary technology that enables data and application portability (e.g. cloud bursting for load balancing between clouds).

The following table summarises the pros and cons of the different deployment models<sup>5</sup>.

Option	Pros	Cons
Private Cloud	<ul style="list-style-type: none"> <li>+ More control and reliability: IT can control the security of data, set compliance requirements, and optimize networks more effectively with cloud.</li> <li>+ Customizable: IT can customize storage and networking components so that the cloud is a perfect fit for the specific organization and its needs.</li> </ul>	<ul style="list-style-type: none"> <li>- Requires IT expertise: A high-level of IT expertise is required to ensure maximum effectiveness and optimal configuration of the deployment.</li> <li>- Costlier: The long-term costs may be higher due to increased management responsibilities and smaller economies of scale.</li> </ul>
Public Cloud	<ul style="list-style-type: none"> <li>+ Ease of management: Organisations IT departments do not manage their public cloud; they rely on Cloud provider to administer the cloud.</li> <li>+ Ease of deployment: With the public cloud, there is low barrier to entry, so you can quickly configure and stand up a cloud.</li> <li>+ Flexible: Users can add or drop capacity easily. Moreover, the environment is typically accessible from any Internet-connected device, so users don't need to jump through many hurdles to access.</li> </ul>	<ul style="list-style-type: none"> <li>- Can be unreliable: Public cloud outages are quite common, leading to headaches for users.</li> <li>- Less secure: The public cloud often has a lower level of security and may be more susceptible to hacks. In some cases, cloud providers may not be able to meet the strict constraints mandated by government institutions.</li> </ul>
Hybrid Cloud	<ul style="list-style-type: none"> <li>+ Flexible and scalable: Organisations are able to combine and match for the ideal balance of cost and security.</li> <li>+ Cost effective: Organisations can take advantage of the cost-effectiveness of public cloud computing, while also enjoying the security of a private cloud.</li> </ul>	<ul style="list-style-type: none"> <li>- Complexity of management: Moving parts between public and private clouds can be a challenge.</li> <li>- Requires IT expertise: A high-level technical staff is required to guarantee security vulnerability on all aspects is decreased.</li> </ul>

**Table 5 Pros and cons of private, public and hybrid deployment Cloud models**

In short, when choosing a specific cloud deployment model, it comes down to a series of trade-offs related to cost, management and security. While public clouds may be the best option for small organisation from a cost perspective, organizations that require more control and/or security may opt for a private or hybrid cloud — providing they have the manpower and budget to manage those deployments effectively.

**The iBorderCtrl approach:** The iBorderCtrl Platform offers great flexibility to the border control application. During the iBorderCtrl project, the platform will operate both in a Private Cloud environment (hosted on selected provider after SLA agreement is achieved), as well as in a Hybrid Cloud environment (hosted on European Dynamics developed infrastructure for the information that will not be exposed to external providers).

### 2.2.1.3 Open Technologies

Being open is about adopting the technology decisions that organizations have made and giving them the freedom to move across technologies, models and cloud providers. Systems composed of open technologies

<sup>5</sup> Open cloud: Just a buzzword or the future of infrastructure?, viewed November 9, 2017 <<http://goo.gl/1QRICY>>

provide the freedom to change environments and deliver a robust and secure experience extending existing IT to the cloud. Embracing an open cloud means there is no technology lock-in, no contractual lock-in and no service lock-in. It means providers don't dictate technologies and that competition is embraced. New, emerging standards will increase the portability and interoperability of systems across cloud service providers, and will reduce or eliminate this current barrier to cloud adoption.

**The iBorderCtrl approach:** The iBorderCtrl Platform will be built upon widely accepted open source technologies. Moreover, its architecture is a baseline for future extensions and modifications with the objective to allow developers to improve the way functions are implemented or to add new features not currently available.

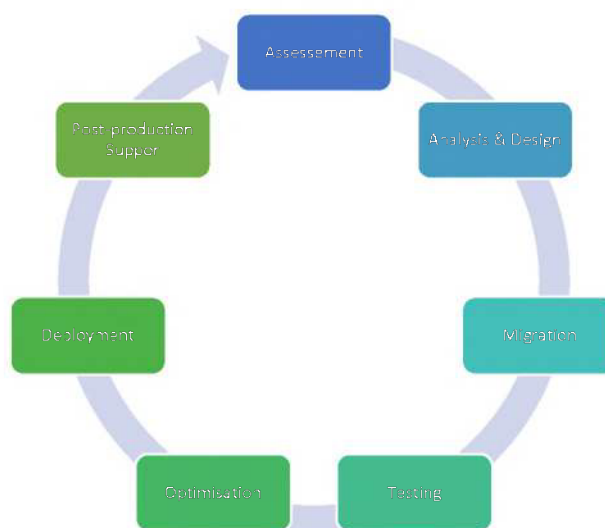
It will be developed using the following open source solutions:

- ✓ **OpenStack** for the implementation of the IaaS Layer. OpenStack is the most popular and most adopted open source IaaS solution.
- ✓ **LAMP** (Linux, Apache, MySQL and PHP) for the implementation of applications' VMs.
- ✓ **MySQL/MariaDB** and **PostgreSQL** database engines for the implementation of Database Services Module.
- ✓ **Gluster** for the implementation of file Sharing Service Module.
- ✓ **HAProxy** for the implementation of Load Balancer Module.
- ✓ **Zabbix** for the implementation of the Monitoring Module
- ✓ **phpMyAdmin** for the implementation of the MySQL Database Administration Module
- ✓ **phpPgAdmin** for the implementation of the PostgreSQL Database Administration Module
- ✓ **Duplicity** for creating the backups.

The implementation of iBorderCtrl platform on open source technologies will not lock the organisations that use it into a proprietary ecosystem and thus made it extremely hard to move their application to another provider. Another significant advantage of the use of open source solutions is the fact that the platform will continue to benefit from the improvements in the operability and security of OpenStack, Cloud Foundry and all the other tools. It can scale without significant development effort. Also, the selection of broadly adopted software packages guarantees the long-term support of the solution.

#### 2.2.1.4 Migration planning and application

Application migration is the process of redeploying an application, typically on newer platforms and infrastructure. Comprehensive planning, driven by a disciplined migration process will contribute greatly to a successful redeployment of the applications to a new cloud environment.



**Figure 13 A typical migration project life cycle**

During the migration process the following technical considerations must be taken into account<sup>6</sup>:

- The creation of a detailed inventory of the current application portfolio really helps in terms of understanding the scope of the migration effort. This includes capturing information regarding the

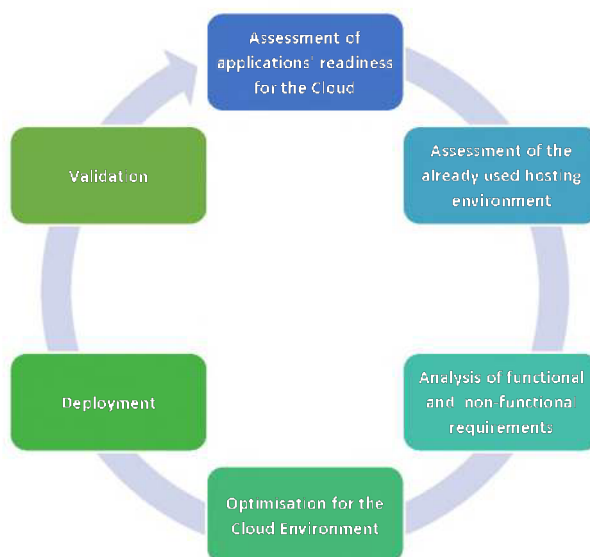
<sup>6</sup> Microsoft, 2014, Protecting Data and Privacy in the Cloud

number of software modules, scripts, and external interfaces involved. It also includes hardware and software configuration information, including operating system versions, database versions, features/functionalities in use, and similar information.

- A security audit of the application and its data is vital. The cloud service's security features may be very different from those of the in-house environment, and the security risks and the measures applied to counter them must be assessed carefully.
- Temporary subsystems can be established to facilitate migrations.
- Standardization and automation can help reduce the risk of migration errors. Virtual machine templates can be rapidly deployed and bring an environment online in a day. Automated data integrity and validation methods can be used to verify and validate data, databases and files during the initial synchronization.
- The creation of migration tools, which ensure a high level of automation along with accuracy in migration can result in less time spent in migration and testing.

**The iBorderCtrl approach:** iBorderCtrl will follow a multi-phase migration process; process will start with the assessment of each application regarding its readiness for the new Cloud environment, its architecture and its functional and non-functional requirements. Afterwards, the code and data will be deployed in the platform's Application and Data Service Layers, respectively. The process will be completed with the validation that the application is fully operational in the new Cloud environment.

The following diagram presents the iBorderCtrl migration process.



**Figure 14 iBorderCtrl migration process**

The iBorderCtrl migration process includes the following six steps:

**Step 1 - Assessment of applications' readiness for the Cloud**

The 1<sup>st</sup> step aims to evaluate if the services are ready for the cloud environment. Aspects such as customization, regulatory compliance, complex service architectures and service maturity are carefully investigated, as they would negatively impact the cloudification process. A crucial aspect is the availability of both the application's source code and documentation (installation manual, code dependencies, required software packages, etc.). Finally, the commitment of the application's development and support team should be ensured.

**Step 2 - Assessment of the already used hosting environment**

The 2<sup>nd</sup> step aims to analyse the environment used to host the services. The analysis covers both the network (e.g. configuration, connectivity requirements from the municipality premises to the cloud environment, and supplementary services such as SMTP, DNS and WWW) and architecture (e.g. use of resources, underlining technologies, licenses, and security mechanisms) of the service.

### **Step 3 - Analysis of functional and non-functional requirements**

The 3<sup>rd</sup> step aims to define the technical characteristics of the Virtual Machines that will host the applications on the new Cloud Environment. The analysis of the functional requirements covers technical details (e.g. Operating System, Scripting Language, Database, Web/Application Server, Data Formats, Frameworks/Libraries and External Services used), interoperability issues, and static characteristics such as hard-coded IP address and directory paths. Furthermore, the analysis of the non-functional requirements addresses issues related to the proper functioning of the application such as security, regulatory compliance, performance, availability, backup; privacy, reusability, and interoperability. An estimation of the use of resources regarding RAM, Disk Space, CPUs, Bandwidth, Hits/Month, Registered Users, Max On-line Users, and Average On-line Users contributes to the calculation of the expected workload per application. An important characteristic that should be examined in this step is if the application's design supports its deployment in multiple servers. In that case the application will take full advantage of the performance benefits that cloud offers.

### **Step 4 - Optimisation for the Cloud Environment**

The 4<sup>th</sup> step aims to solve the problems identified in the previous step, so the application to be ready for deployment in the new environment. Moreover, it includes modifications that enable the application to support natively the most prominent Cloud characteristics (e.g. high-availability and scalability). The latter is closely related to the available budget or the internal IT capabilities.

### **Step 5 – Deployment**

The 5<sup>th</sup> step aims to transfer the ready to be cloudified applications to the new Cloud environment. The deployment process includes the following actions:

- a) setup of the cloud environment that will host the selected services;
- b) launch the VM instances that will host the applications and their data (e.g. database and file sharing modules).
- c) migrated both the applications and their data to the Cloud environment

### **Step 6 – Validation**

The final step aims not only to ensure that the deployed applications are operational but especially that they meet the initial set of requirements regarding cloudification. The validation is made in collaboration with the municipalities and includes functional tests ensuring that the deployed application performs as designed.

## **2.2.2 Security of cloud infrastructure**

Cloud computing security is an evolving sub-domain of information security and refers to a broad set of policies, technologies, and controls deployed to protect data, applications, and the associated infrastructure.

There is a number of security concerns associated with cloud computing; they can be broadly classified in two categories, namely issues faced by Cloud Service Providers (CSPs) and those faced by Cloud Service Consumers (CSCs). Providers must ensure that their infrastructure is secure and clients' data and applications are protected; consumers, on the other hand, must ensure that their provider has taken appropriate security measures to protect their information.

Current Cloud delivery models (whether implemented on an Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS), or Software-as-a-Service (SaaS) model) are ruled by Service Level Agreements (SLAs) that normally define mutual supplier and user expectations and obligations. The central idea behind such models is that the consumer ought to trust the supplier. Venturing into a public cloud environment, especially via an IaaS model, security becomes a shared responsibility. Although there are certain measures, which a cloud provider will apply to ensure that Virtual Machines (VMs) stay secure, a considerable number of tasks are left in the hands of the tenant (cloud consumer).

The UK's National Technical Authority for Information Assurance, which provides advice on Information Assurance Architecture and cyber-security to UK government and the wider public sector and suppliers to UK

government, published 14 security principles to consider when evaluating cloud services, and why these may be important to an organisation<sup>7</sup>.

Cloud Security Principle	Description
1. Data in transit protection	Consumer data transiting networks should be adequately protected against tampering and eavesdropping via a combination of network protection and encryption.
2. Asset protection and resilience	Consumer data, and the assets storing or processing it, should be protected against physical tampering, loss, damage or seizure.
3. Separation between consumers	Separation should exist between different consumers of the service to prevent one malicious or compromised consumer from affecting the service or data of another.
4. Governance framework	The service provider should have a security governance framework that coordinates and directs their overall approach to the management of the service and information within it.
5. Operational security	The service provider should have processes and procedures in place to ensure the operational security of the service.
6. Personnel security	Service provider staff should be subject to personnel security screening and security education for their role.
7. Secure development	Services should be designed and developed to identify and mitigate threats to their security.
8. Supply chain security	The service provider should ensure that its supply chain satisfactorily supports all of the security principles that the service claims to implement.
9. Secure consumer management	Consumers should be provided with the tools required to help them securely manage their service.
10. Identity and authentication	Access to all service interfaces (for consumers and providers) should be constrained to authenticated and authorised individuals.
11. External interface protection	All external or less trusted interfaces of the service should be identified and have appropriate protections to defend against attacks through them.
12. Secure service administration	The methods used by the service provider's administrators to manage the operational service should be designed to mitigate any risk of exploitation that could undermine the security of the service.
13. Audit information provision to consumers	Consumers should be provided with the audit records they need to monitor access to their service and the data held within it.
14. Secure use of the service by the consumer	Consumers have certain responsibilities when using a cloud service in order for this use to remain secure, and for their data to be adequately protected.

**Table 6 Cloud Security Principles (Source <http://goo.gl/mUf5c2>)**

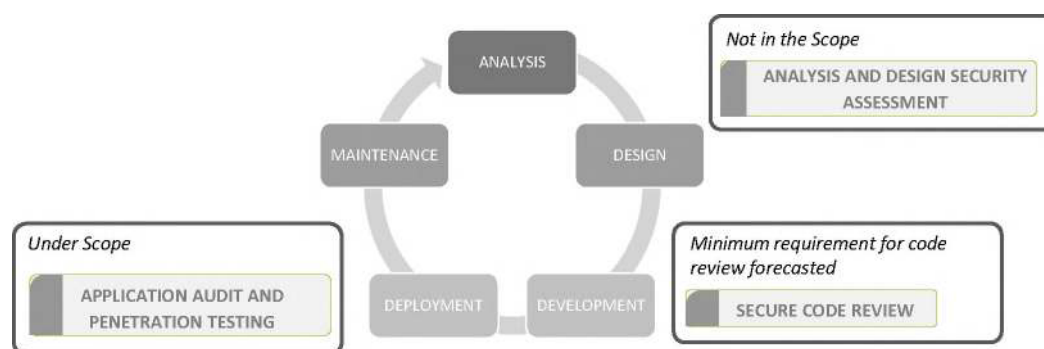
<sup>7</sup> OWASP Zed Attack Proxy Project, viewed November 10, 2015 <<https://goo.gl/A46kLn>>



Consumers of cloud services should decide which of the principles are important, and how much assurance they require in the implementation of these principles, while providers of cloud services should consider these principles when presenting their offerings to public sector consumers. This will allow consumers to make informed choices about which services are appropriate for their needs.

**The iBorderCtrl approach:** In order to achieve a clear understanding of the security requirements and protect accordingly the system a set of penetration tests will be performed to iBorderCtrl system to detect any possible vulnerability and act accordingly to address them.

The methodology that is going to be followed to perform this test has been developed by the everis ADS cybersecurity team, EVER-Audit, and combines the best practices defined by prestigious communities such as OWASP (Open Web Application Security Project) and standards such as NIST 800-115 or ISSAF (Information Systems Security Assessment Framework), but also offers flexibility and can be adapted to different applications in terms of criticality and functionality. This methodology cover the complete lifecycle of the software development thanks to three different methods, as depicted in the picture below:



**Figure 15 Security assessment methodology**

However, within the scope of this work package only the “application audits and penetration testing” and “secure code review” will be undertaken. All the test will be provided from the everis ADS SOC that groups a specialized penetration testing team.

Depending on the degree of information that is open and shared for each of the different systems developed in iBorderCtrl one of the following approaches will be selected to perform the penetration tests:

- **Black box**, not knowing the details of the application: we will only need information about the communication and connection. For example, the URL of the Application, or IP address.
- **Grey box**, when requested by legitimate users and when limited information is provided: for this approach, besides the information provided for a black box approach, we will need brief information regarding the target and credentials to access it. For example, the credentials to access a private area of a Web site application.
- **White box**, when requested by legitimate users and when full information is provided: for this approach, besides the information provided for a grey box approach, we will need all the information regarding the target and credentials to access it. For example, credentials to access a private area of a Web site application, manuals, design documents, requirement lists, ...

The methodology used to conduct an application penetration test analyses dynamically (during runtime) the security features of the applications that a user and/or machine interacts with. It is based either on **black box**, **on grey box** or **on white box**. The following picture shows the different phases we execute to perform an application penetration testing:



**Figure 16** phases in penetration test analysis

**Penetration planning:** during this phase it will be defined different aspects such as the scope of the penetration testing, limitations, test plan, etc. It contains two sub phases. The details obtained from these two subtasks will depend on the approach selected (black / grey / white):

- **Logistics:** the first set of actions regarded in the methodology will aim to specify every single organisational, technical and/or operating aspect required for the audit. Aspects and **limitations** (time), **connectivity** or **access** will be defined.
- **Test planning:** once the initial context for the audit and the potential limitations have been established, a plan with the tests that will be executed will be elaborated. This activity aims to define the **OWASP controls** (based on the OWASP test guide) and the **OWASP domains** that will be tested in the following phases. The definition of the specific test cases will take place during the "security test execution" phase. They will be included in a specific chart of the OWASP control table (in the case of a test per control), in order to facilitate the reading and tracking of the cases.

The rest of the phases are equal regardless of the approach selected.

**Security test execution:** during this phase the tests previously agreed upon will be performed in two different sub-phases:

- **Passive mode:** the tests executed during the "Passive Mode" sub task refer to those tests that aim to understand the application from a technical and functional perspective; always recurring to a non-intrusive approach (transparent for the application). The tests will be executed through methods such as **navigation** throughout the whole application and **research**.
- **Active mode:** during this phase the necessary test plans will be designed to execute the **OWASP controls** previously defined. Finally, the tests will be executed.

**Report elaboration and impact assessment:** during this phase, the team will report the tests in detail, as well as solutions and recommendations for the vulnerabilities found. It has three sub-phases:

- **Technical Report:** the targeted OWASP controls for the audit are detailed during this phase, establishing a specific chart for each of them, where all the information regarding the testing and results will be gathered. The aim is to centralise the information and facilitate the understanding, both for management teams interested in the report and for the development and/or system teams. The report will contain the necessary **notes and evidences** to understand and justify the vulnerability and also recommendations and **solutions** to solve the shortcomings.
- **Impact assessment:** this phase aims to assess the risks associated with the OWASPs controls that failed to pass the audit tests. For this purpose the OWASP assessment methodology will be used, which resorts to an approach to determine the attacker typology, the vulnerability hazard, as well as




the technical and business impact, in order to evaluate the actual risk posed by the vulnerability. The team will analyse five critical points:



- a. **Threat agent:** estimation of the potential success of an attack based on the specific attacker profile that the team considers that may try to take advantage of the vulnerability. The metrics for the assessment will include aspects such as skill level, motive, opportunity or dimension.
  - b. **Vulnerabilities:** Estimation of the chances of finding and exploiting a vulnerability. The metrics for this analysis will include finding easiness, exploitation easiness, awareness or intrusion detection.
  - c. **Technical impact:** estimation of the scope of the impact (severity) on the system if the vulnerability is successfully exploited. The metrics for this analysis will include confidentiality loss, integrity loss, availability loss and accounts compromised.
  - d. **Business impact:** estimation of the scope of the impact (severity) on the business if the vulnerability is successfully exploited, using relevant factors from the business management field to determine the effects. The metrics for this analysis will include financial damage, image damage, Noncompliance and Privacy violation.
  - e. **Risk:** risk estimation based on the likelihood of this event taking place, as well as the impact incurred. The likelihood estimations will be made based on the “threat agent” and “vulnerabilities” metrics, and other metrics will be used to calculate the impact.
- **Full Report or brief report:** after the impact assessment, the technical report (findings) will be completed with the necessary and relevant information. The final report will contain a **controls summary**, obtained **conclusions** and **recommendations** and an **executive summary**. Then, the report will be completed with the necessary chapters for a full report or a brief report, depending on the request and the needs of the project. The executive summary will focus on gathering a controls summary with the main conclusions drawn from the audit process for a full understanding of the situation.

**Shortcoming review:** this **optional phase**, is performed upon the implementation of solutions in the development equipment/systems, to ensure that everything is properly functioning.

- **Post audit:** the reports previously elaborated are submitted to the development and infrastructure teams in order to fix the vulnerabilities previously identified. After fixing them, we carry out a **shortcoming review** to make sure that the measures implemented succeeded in eliminating the vulnerabilities, and the result of these tests are included in a **post-audit report**.

The following table describes the different tools used and the benefits of each one of them.

Application audit tools (tools and benefits)	
 <b>Burp Suite</b>	Set of tools integrated in a platform that will allow the security team to execute complete web audit tests. The version that the security team uses allows the integration of plug-ins and other tools, for the team to specify the tests and obtain more accurate results. Their vulnerability database is frequently updated, letting us detect the newest vulnerabilities in the market.
 <b>Zap Proxy</b>	The security team uses Zap Proxy to complement the Burp functionality because it provides an exceptional fuzzer with a lot of dictionaries (WebScarab’s dictionaries) and other functionalities that help us to audit Ajax applications.
 <b>Acunetix</b>	This tools allows to schedule audit processes and thus monitors the status of vulnerabilities. In addition it counts with a great vulnerability database frequently updated. Moreover, it has a great vulnerability scanner for Web Services.

 <b>Nessus</b>	<p>Nessus is used to obtain information about the architecture of the solution (web server, public directories, type of machine, etc.). This data is very valuable to identify sensitive information about the application such as the HTTP server version or HTTP methods.</p>
 <b>soapUI</b>	<p>SoapUI offers a good interface to test Web Services. It allows us to modify parameters, structures and other aspects and we can watch the response very quickly.</p>

**Table 7 Penetration test audit tools**

### 2.2.3 Cloud Providers selection

The choice of a Cloud Service Provider (CSP) requires the evaluation of an extensive list of options. The principal elements to consider for almost every organisation are:

- Service Levels:** This characteristic is essential as the Public Authorities in most cases have strict needs regarding availability, response time, capacity and support. Cloud Service Level Agreements (CSLA) are an essential element to choose the right provider and establish a clear contractual relationship between a cloud service customer and a cloud service provider of a cloud service.
- Support:** The support is a parameter to consider carefully. It could be offered online or through a call centre, and in some cases, it could be necessary to refer to a dedicated resource with precise timing constraints.
- Security:** As already mentioned security is paramount. Although normally, the potential supplier should follow recognised security policies in line with industry best practice, Public Authorities have to formulate a number of relevant questions (i.e. what is the security level offered by the providers? which mechanisms are in place to preserve client’s applications and data? etc.) to evaluate this essential feature for the overall architecture.
- Privacy:** Particular attention has to be reserved to legal requirements for the protection of the personal data hosted in the cloud service. Public Authorities should understand the data privacy and retention policies too, as well as where the CSP’s data will be located, including any transborder data transfer, if applicable.
- Open Standards:** In order to avoid getting locked-in to cloud infrastructure that has restrictive contracts or proprietary technologies (technologies that are unique to the particular supplier), Public Authorities should prefer solutions that are implemented with fully open source technologies and open cloud standards. These technologies have an elegant escape hatch built into them by their design. Public Authorities can take the entire stack and host it on another CSP or in their premises without losing productivity or data. This backup plan protects them against legislative changes, company restructuring, and much more.
- Compatibility:** The requirement of the cloudified applications have to fit into the CSP’s existing pre-configured templates and may increase the cost of configuration. Moreover, the CSP’s architecture should meet scalability, availability, capacity and performance guarantees and should be sufficient for agency requirements.
- Pricing:** Although most cloud providers use the aforementioned “Pay per Use” model, each CSP has a different price system. Understanding how you pay for each service is essential for a meaningful comparison. Moreover, additional costs can still arise, for example through the use of extra features. Terms of the contract, payment methods and payment dates can be deciding factors as well.
- Redundancy:** The provision of duplicate or backup equipment that takes over the function of equipment that fails should be discussed at an early stage. The redundancy process and timeframe have to meet the agency’s requirements and especially its obligations to the citizens. Thus, adequate backup procedures and robust disaster recovery plans must be incorporated into the cloud offering.

- **Easy to use administration environment.** Make sure your potential provider has a user-friendly client portal. It should allow you to conduct admin tasks or add storage space or services quickly. Ask for a demonstration before you choose one CSP over another.

Most of the considerations mentioned above have already be analysed in separate sections of this document. Given this, we will put more emphasis into the evaluation of two essential elements: Cloud Service Level Agreements and pricing.

**The iBorderCtrl approach:** iBorderCtrl Platform is based on OpenStack. ED is participating in the **Cloud28+ initiative**<sup>8</sup>; Cloud28+ (<http://www.cloud28plus.eu/>) is an open community of Cloud Service Providers, Cloud Resellers, ISVs, Systems Integrators and government entities dedicated to accelerating enterprise cloud adoption across Europe, the Middle East and Africa. Cloud28+ maintain a catalogue of trusted, business cloud services that matches in-country or cross-border buyer and regulatory workload requirements. The initiative offers the following benefits:

- Find the right cloud service for your needs based on location of datacentres, price, SLA, certification level, or other workload criteria
- Enable your business to transform to fast, agile Hybrid IT
- Access the largest cloud services community and software developer network in the EU
- Learn about best practices and implementation success stories
- Maintain data sovereignty and feel secure with trusted certification
- Avoid proprietary technology lock-in, thanks to an open source service provider community

## 2.3 Security

As described in a thorough manner in D2.2 (Reference Architecture and Components Specifications) the iBorderCtrl system has been designed and developed following principles and requirements (constraints) stemming from the legal analysis of its concept. These have been transposed directly into technical specifications for the system. Therefore, security-by-design, privacy-by-design and data protection (in transit and stored) principles have guided the implementation of the IT infrastructure, including the backend but also the communication channels (interfaces between architectural components). As described in D2.2, the GDPR and Directive 680/2016/EU have been followed. To this extent, encryption is applied to all components whenever required, such as using “https” for the different web services, or securing the radio network using 802.11 transmissions and authentication using WPA2-PSK AES. Further measures to be implemented are in line with the participating parties corporate certified security services (e.g. ISO 27001 certification). More specifically, data isolation techniques are being used, according to which information remains private within a system, unless to be shared; sub-systems will not share internal information with each other. Anonymity is achieved, as a QR-code is exchanged between sub-systems, thus concealing any personal user information from being exchanged, unless it is required to do so. Database encryption is assured though MySQL “embedded” Transparent Data Encryption.

Additionally, a Data Protection Impact Assessment (DPIA) has been conducted (Section 7), as the process that can help identify and reduce the privacy risks of the system. DPIA is expected to enable all participating technology providers to systematically and thoroughly analyse how their particular implementations will affect the privacy of the individuals involved. DPIA can reduce the risks of harm to individuals through the misuse of their personal information and contributes to the design of more efficient and effective processes for handling personal data.

---

<sup>8</sup> The Cloud28+ initiative is not affiliated with the iBorderCtrl project.

## 3 Risk Based Assessment Tool (RBAT)

### 3.1 RBAT overview

iBorderCtrl's Risk Based Assessment Tool (RBAT), is a tool which will support the decision-making process of the Border Guard Authorities (Managers and Agents). RBAT will enable a common, harmonised model for risk management and implement a systematic process to stimulate compliance and prevent and/or treat the risk of non-compliance, including risk of fraud and any other risk which appears to threaten the Authorities objectives. RBAT will also identify cases that deserve further investigation, facilitating in this way better resources allocation for the Border Managers and Agents.

RBAT, which is used both at the pre-registration and the border crossing phase, serves three main purposes within the iBorderCtrl system:

- to calculate the overall risk of each traveller crossing the borders based on individual risk scores produced by other iBorderCtrl modules/tools
- to enable Border Managers to author rules based on previously identified "Risk objects" in order to automatically produce risk indicators
- to issue alerts to the Border Guards (based on the risk indicators accrued by the Border Managers' authored rules) to pay extra attention in specific cases or/and travellers.

The RBAT module provides a number of diverse techniques in order to identify a risk. The risk assessment process takes into account the following data:

- The risk scores provided by each iBorderCtrl module (during the pre-registration and the border crossing phases) which are included in the "Risk" database (presented in section 3.3.3) are used by RBAT's weight-based algorithm (presented in section 3.3.1) in order to produce the overall risk (for both phases). It should be noted that during the overall risk calculation, a "weight" is assigned to each module's produced risk score based on the Multi-Criteria Decision Analysis presented in section 3.2.
- The iBorderCtrl database fields presented in section 2.1.4.1 are converted by RBAT into "Risk objects" which are used by Border Managers during the rule authoring process as explained in section 3.3.2.
- All traveller related information gathered during the pre-registration phase (stored to the iBorderCtrl Database) is compared (for each traveller) to the generated rules in order to automatically produce "Risk Indicators" and alert the border guards accordingly.

RBAT is designed to treat risk management as an interactive process in which information is continuously updated, analysed, acted upon and reviewed.

#### Pre-registration

Once a traveller successfully completes the pre-registration phase, an .xml file containing a) all information stored to the iBorderCtrl database as entered by the traveller during the pre-registration phase and b) all risk scores provided by the other pre-registration modules (DAAT, FMT, ELSI) for this traveller and stored to the "Risk" database, is "pushed" to RBAT. The same procedure is followed for each traveller who completes the iBorderCtrl pre-registration procedure.

Once the RBAT engine receives the .xml file with the above mentioned data concerning a traveller will then:

- Produce the overall pre-registration risk for this traveller using the weight-based algorithm and store it to the "Risk" database.
- Produce Risk indicators by comparing the traveller related data to the existing rules authored by Border Managers to examine the possibility of a match. Regarding the pre-registration phase where there is some time margin until the traveller reaches the borders, the Border Managers are given the opportunity to evaluate the produced "Risk Indicators" that depict the likelihood of new risks and suggest further actions (asynchronous check).

- Based on the Border Managers “Risk Indicators” evaluation, RBAT will store these risk indicators to the “Risk” database so as to be available to the Border Guard for examination at the border crossing point

***Figure 17 – RBAT – Pre-registration Phase***

#### Border Crossing Point (BCP)

At the Border Crossing Point, once all checks have been completed by the border guard for a traveller, an .xml file containing a) all information stored to the iBorderCtrl database as entered by the traveller during the pre-registration phase and b) all risk scores provided by the other border crossing check modules (DAAT, BIO, FMT, HHD) for this traveller and stored to the “Risk” database, is “pushed” to RBAT. The same procedure is followed for each traveller going through checks at the borders.

Once the RBAT engine receives the .xml file with the above mentioned data concerning a traveller will then:

- Produce the overall risk for this traveller using the weight-based algorithm and store it to the “Risk” database.
- Produce Risk indicators by comparing the traveller related data to the existing rules authored by Border Managers to examine the possibility of a match. Regarding the border crossing phase where the border guard must take a direct decision for the traveller’s admission or entry refusal, and is necessary to have all data available, the Border Managers don’t have the time to evaluate the produced “Risk Indicators” (synchronous check).
- RBAT will, as a result, store all produced “Risk Indicators” to the “Risk” database so as to be available to the Border Guard (for examination).



*Figure 18 RBAT – Border crossing Phase*

### **3.2 Multi Criteria Decision Analysis (MCDA) - Steps for the “weight” determination of each iBorderCtrl module**

In order to calculate the overall risk of each traveller crossing the borders based on individual risk scores produced by other iBorderCtrl modules/tools, the significance and respective importance (“weight”) of each module to RBAT is going to be determined using the Multiple-criteria decision analysis (MCDA)<sup>9</sup> technique. MCDA is a sub-discipline<sup>10</sup> of operations research that explicitly evaluates multiple (sometimes conflicting) criteria in decision making. Using MCDA can be said<sup>11</sup> to be a way of dealing with complex problems by breaking the problems into smaller pieces. After weighing some considerations and making judgements about smaller components, the pieces are reassembled to present an overall picture to the decision maker. Most of MCDA methods deal with discrete alternatives, which are described by a set of criteria. Information could be determined exactly or could be fuzzy, determined in intervals.

#### **Establish the decision context**

This approach main purpose is the identification of the risk input tools/ parameters and the assessment of these options based on some generic identified criteria. The final aim is to prioritise the significance (weight) of each option regarding the final risk calculation.

All technical partners participated in the above procedure as all perspectives on the subject of the analysis should be covered and each one expertise (on their tool or technology) led to useful and significant contributions to the MCDA. Moreover, all agreed with the result of the prioritization analysis and weight definition process.

The MCDA is structured to:

---

<sup>9</sup> “Multi-criteria analysis: a manual”, Department for Communities and Local Government: London, 2009

<sup>10</sup> [https://en.wikipedia.org/wiki/Multiple-criteria\\_decision\\_analysis](https://en.wikipedia.org/wiki/Multiple-criteria_decision_analysis)

<sup>11</sup> “Multiple criteria decision-making techniques and their applications – a review of the literature from 2000 to 2014”, Abbas Mardani, Journal of Economic Research, vol.28, Sep 2015, pp.516-571.



- show the best way forward regarding the weight determination of the (risk input) options
- prioritise the options
- clarify the differences between the options

The whole MCDA procedure and specific steps followed is a scientifically proven methodology to facilitate the decision making and will contribute to minimise threats on the weight calculation for each risk score provided by each tool.

### **Identify the options to be appraised**

The identified options are all the iBorderCtrl modules which provide risk related information namely: DAAT, BIO (fingerprints), BIO (palm vein), FMT, HHD, RBAT, ELSI, BCAT and risk indicators. These options are placed as the first column of the traceability matrix presented in table 9.

The MCDA should be open to the possibility of modifying or adding options as the analysis progresses.

### **Identify objectives and criteria.**

In this step specific criteria are identified for assessing the consequences of each option. Furthermore, the criteria are organised in clusters under high-level and lower-level objectives in a hierarchy.

A hierarchical model of objectives and criteria, a value tree, has been developed as shown in Figure 19 below.



***Figure 19 Value tree***

The identified Objectives and respective Criteria are thoroughly explained below:





The objectives are placed as the second row of the traceability matrix presented in Table 8 Performance matrix.

***Table 8 Performance matrix***



**'Scoring'. Assess the expected performance of each option against the criteria. Then assess the value associated with the consequences of each option for each criterion.**

A score to each option should be assigned based on the criteria for each objective. Then the consistency of the scores on each criterion should be checked. For this purpose, a consequence table was created for each objective by placing the identified options as the first column of the consequence table and the respective criteria for each objective as the first row. The separate consequence tables for each objective have been filled (according to the ranges explained in section 3.2.3 above) by the respective partner of each module and presented in the tables below.

***Table 9 Consequence table for "Technology Maturity" objective***

The averaged outcomes of the above table (**green cells**) are placed to the second column of the performance matrix (Table 8)

***Table 10 Consequence table for "Accuracy and reliability" objective***

The averaged outcomes of the above table (**yellow cells**) are placed to the third column of the performance matrix (9)

***Table 11 Consequence table for “Performance” objective***

The averaged outcomes of the above table (**grey cells**) are placed to the fourth column of the performance matrix (Table 8)

***Table 12 Consequence table for “Universality” objective***

The averaged outcomes of the above table (**purple cells**) are placed to the fifth column of the performance matrix (Table 8)

### **Table 13 Consequence table for “Phase applied” objective**

The averaged outcomes of the above table (**pink cells**) are placed to the sixth column of the performance matrix (Table 8)

#### **Assign weights for each of the objectives to reflect their relative importance to the decision.**

The weight on a criterion reflects both the range of difference of the options, and how much that difference matters. So it may well happen that a criterion which is widely seen as ‘very important’ – say accuracy and reliability – will have a similar or lower weight than another relatively lower priority criterion – say universality. Any numbers can be used for the weights as long as their ratios consistently represent the ratios of the valuation of the differences in preferences between the top and bottom scores (whether 100 and 0 or other numbers) of the scales which are being weighted.

The proposed weights per objective have been placed on the performance matrix (**blue cells**). Their sum equals to 1.

#### **Combine the scores for each option to derive an overall value (“weight” of each iBorderCtrl module).**

In order to combine the scores for each option to derive an overall value, an option’s score on an objective is multiplied by the importance weight of that objective. The same is applied for all the options and then the products are summed up to give the overall preference score for that option. The process is repeated for the remaining options. The identified weights for each tool are calculated, as previously stated, as a weighted average for each option across the objectives and presented in the two last columns of the performance matrix (**red coloured cells**). In the last column, the resulting weights are normalised so they sum to 1.0 (but displayed as 100).

## Examine the results.

The way forward was agreed by all partners based on their recommendations.

## Sensitivity analysis.

The next step was to conduct a sensitivity analysis: do other preferences or weights affect the overall ordering of the options? There was a potentially useful role for sensitivity analysis in helping to resolve disagreements between interest groups. Subsequently, the advantages and disadvantages of selected options was reviewed in order to compare pairs of options (create possible new options that might be better than those originally considered). The above steps were repeated until a 'requisite' model was obtained.

## 3.3 System technical description

This section is going to describe how the RBAT engine works (weight-based algorithm, rule authoring environment, retrieve risk scores from the risk database) and how the communication with the other modules of the iBorderCtrl system is achieved.

### 3.3.1 Weight based algorithm for the final risk score calculation

The outcome of the RBAT, namely the overall provided risk score for the preregistration and the overall risk score (in terms of Admission, Refusal or second line check for the traveller) is determined through a weight-based algorithm. This algorithm takes into account both the individual risk scores provided by each respective iBorderCtrl and the weight of each tool (determined by the previous MCDA technique, section 3.2). Weight based is a flexible algorithm which can be defined through the user interface. The user is able to define both the limits of the algorithm (low, medium, high risk) and the objects that participate on it; in this case, all the iBorderCtrl modules related to the risk assessment procedure and provide feedback to RBAT.

The main purpose of RBAT is to support the final decision of the border guard: admission, second line, refusal and the assessment of that options by providing an overall risk score and Risk Indicators based on the risk scores provided by each tool and the rules that the Border Managers are able to author respectively.

RBAT's weight-based algorithm is structured to:

- show the decision maker (border guard) the best way forward
- prioritise the incoming risks
- help the key players to understand the situation better
- improve communication between parts of the iBorderCtrl system

The MCDA procedure presented in section 3.2 for the "weight" calculation (range: 0-100) of each tool will contribute to minimising threats on the final risk calculation by RBAT. The identified weights for each tool are presented in the two last columns of the performance matrix (**red coloured cells**).

Each tool is going to provide their own risk score. The risk score of each tool for RBAT will be expected in a range of 0-100 (value 0 represents 100% refusal to pass and value 100 represents 100% admission). In order to combine the weights and scores for each tool to derive an overall risk value, a weighted average of the risk score for each module is calculated giving the total risk score.

Concluding, RBAT offers two evaluation methods:

- **Weight Based Algorithm:** Calculate the **overall risk** of each traveller crossing the borders based on individual risk scores produced by other iBorderCtrl modules/tools and the weight of each tool
- **Rule Based Evaluation:** Produces risk indicators based on previously identified "Risk objects" and traveller's data

### 3.3.2 Rule Authoring environment

The Rule Authoring environment is a core functionality of the RBAT module. The Rule Authoring is a point and click graphical environment which enables the Border Manager to author rules through the use of structured, non-technical expression of logical interactions between previously identified "Risk objects" with aim to produce risk indicators. All the iBorderCtrl database fields are automatically translated into "Risk objects" which integrate exchanged information into a unified and familiar view of the underlying message-based infrastructure. An example of "Risk Objects" generated based on the iBorderCtrl fields of the Traveller table is presented in the figure below.



**Figure 20 – RISKS objects generation based on the Traveller table of the iBORDERCtrl database**

RBAT module enables the Border Manager to create complex rules and queries using the provided "Risk objects". The rules are structured, non-technical expressions of logical interactions between "Risk objects", resulting in a specific assessment (Risk Indicators). Hence, the Border Manager can define complex criteria based on received information per traveller, execute "what-if" scenarios and even test new rules and criteria.

A rule expresses a query combining the input data then and expresses the proposed action. The RBAT rule authoring environment provides a wide array of logical operators (e.g. equal to, not equal, start with, end with, contained in etc.) and comparisons options. Also it provides the ability to organise the rules into logical groups for efficient maintenance. An example of a very simple defined rule is presented in the figure below:



*Figure 21 Example of rule authoring*

The rule defines:

IF

EXISTS Traveller having surname contained in [REDACTED]

And

EXISTS Travel having Origin equal to "Italy"

And

EXISTS Border Control having vehicle human check equal to "positive"

THEN

Report as Risk Indicator

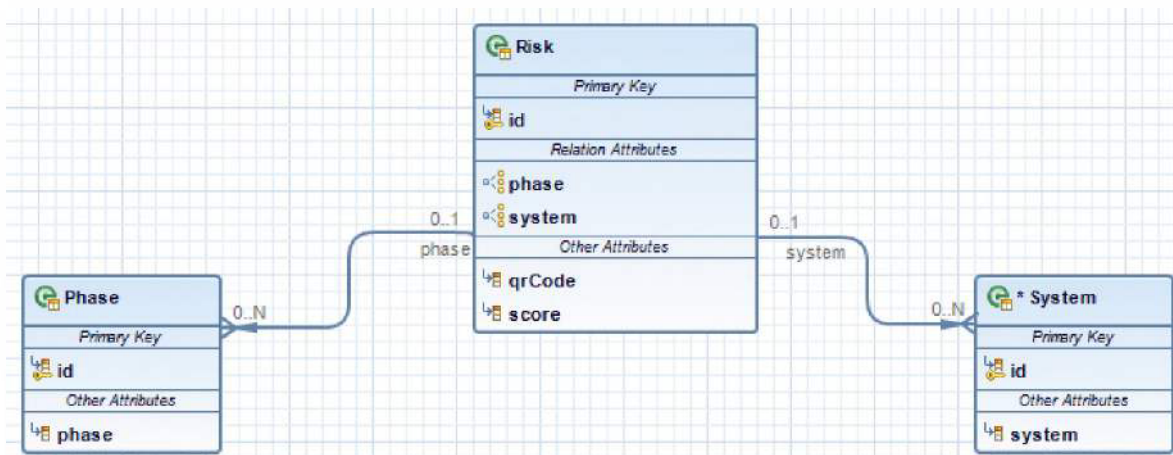
The rule in other words says that if a traveller has the name [REDACTED] and is coming from "Italy" and the vehicle human check was positive for his vehicle, then issue an additional Risk Indicator. If the rule is satisfied, then the produced Risk Indicator can be further evaluated by the Border Manager after the pre-registration phase is completed and before the traveller reaches the borders. If the Risk Indicator is identified during the completion of the border-crossing phase, it is displayed directly to the Border Guard for examination together with the overall risk.

The Border Manager is responsible for the maintenance of the rules that he/she created. The system provides a number of actions over the rules, such as add, edit, delete, copy allowing the user to perform any desired adjustment. The decision making process can be based on a combination of current and previous data sets.

By unifying all these disparate data sources and by providing a simple interface to combine them, the RBAT module allows the user to focus on the actual risk identification process rather than spending valuable time and resources on low level manipulation of data. Furthermore, additional actions can be assigned to each risk category which can either initiate a completely automated response or a mixture of automated responses and operator actions.

### 3.3.3 Risk Database

The Risk database stores the individual risk scores provided by each iBorderCtrl module along with any other information which might be of interest to the Border Guard for both phases: pre-registration and border crossing phase. The following figure presents the tables of the Risk database schema while the table below provides an explanation of each field.



**Figure 22 Risk database schema**

**Table 14 Risk database fields explanation**

Field	Explanation
<b>Phase</b>	Two options: Pre-registration and border crossing phases
<b>System</b>	The name of the system that provides the score (i.e. FMT, DAAT, BIO etc.)
<b>QR code</b>	Generated QR code with traveler_id and trip_id
<b>Score</b>	[json field [includes the risk provided by each system (useful for RBAT) and also other information which might be of interest to the border guard (not useful to RBAT)]]

### 3.3.4 Interfaces to iBorderCtrl system

An.xml file containing all fields from the iBorderCtrl database and the risk score fields per module from the Risk database will be “pushed” to RBAT. This data exchange will take place both at the pre-registration and the border crossing phase. The RBAT module will send the pre-registration and the overall calculated risk and possible identified risk indicators to be stored to the Risk database.

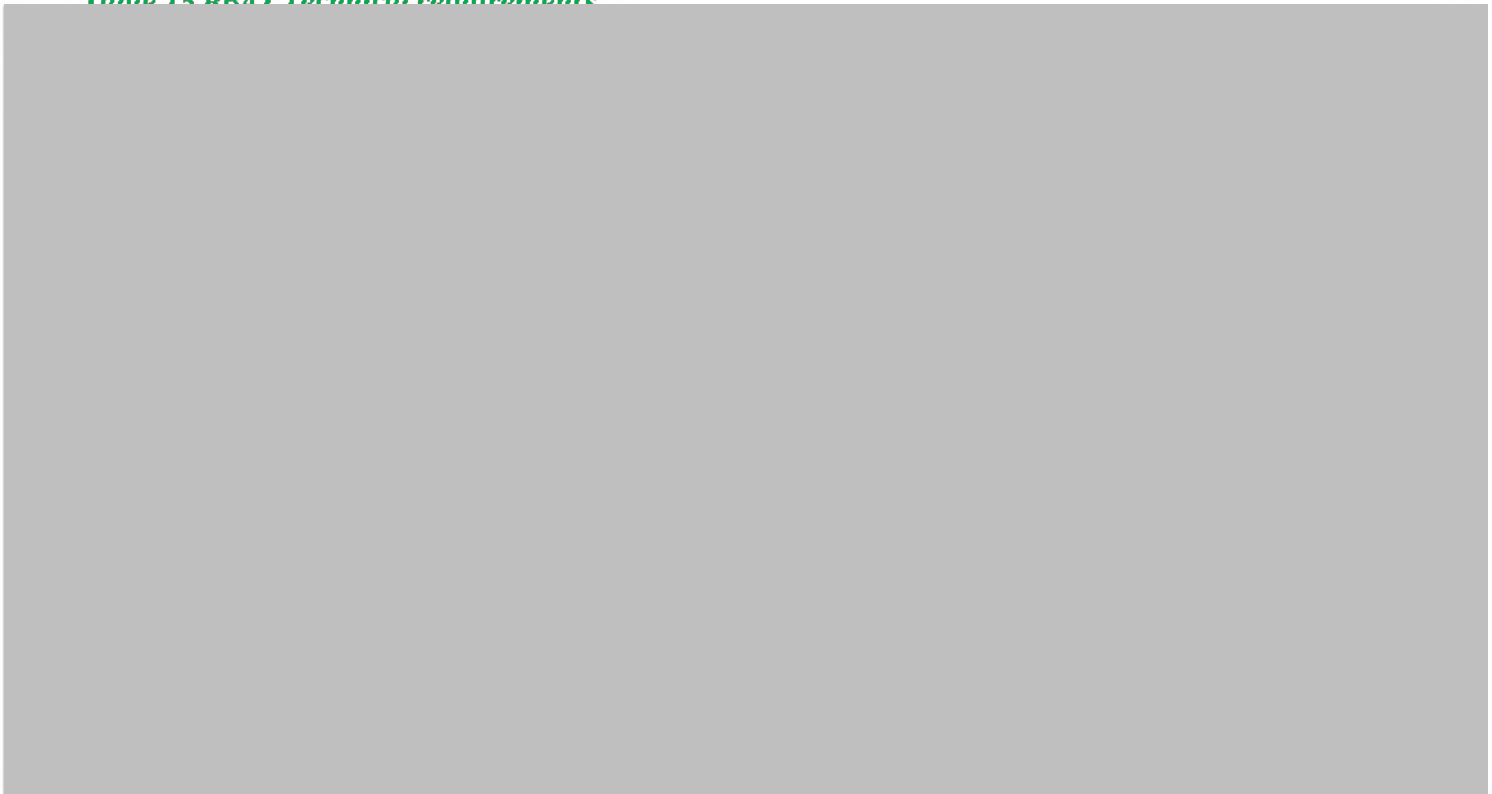


*Figure 23 – RBAT interfaces to the iBorderCtrl system*

### **3.4 RBAT technical requirements coverage**

In this section the way RBAT covers the requirements presented in deliverable D2.2 is going to be described. The next table will list the requirements related with the module and the implementation description.

*Table 15 RBAT Technical requirements*





## D4.1 First version of the iBorderCtrl software platform





## 4 External Legacy and Social Interfaces (ELSI)

In this section a thorough description of the progress on the development of the ELSI module is given. The main idea of ELSI system is to be used to crosscheck traveler information from legacy systems and on-line legally and ethically available information such as from social media. ELSI will both link with the external databases and will attempt to validate relevant to the specific traveler information; the outcomes of this process will update the iBorderCtrl system accordingly with risk scores and explanation checks.

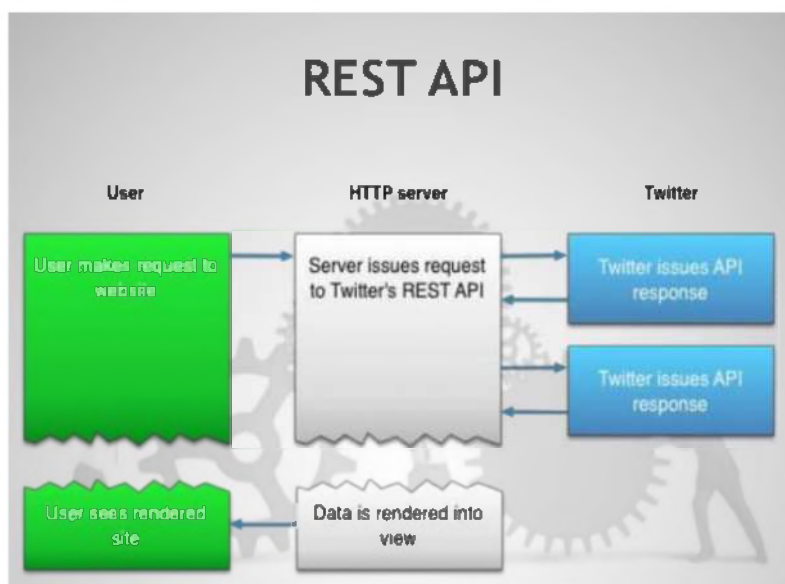
**External legacy databases** provide information of high reliability and will often lead to decisive conclusion, such as: the SIS database may link a traveler with an open arrest warrant, or the VIS database may help reveal potential forgeries, for example when the expiration date of a visa has passed but has been forged on the otherwise original and valid document.

**Publicly available databases** require the traveler to provide account information and consent to retrieve and process such information. This means that the traveler could choose whether he would like to provide this information or not use this functionality at all and potentially even prepare for the purpose of getting a low risk score social media accounts designed to portray them in a benign fashion. Thus, the weight of this analysis will be limited, and the analyses will focus primarily on utilizing the data to identify threats, rather than to validate expected behavior. An example would be the network of followers on twitter that may include specific accounts link to activities and posts of known criminal elements.

### 4.1 Publicly available information from social media platforms

All access to social media will be clearly documented in information provided to the traveler and these will only take place following the consent of the traveler to the analyses described below.

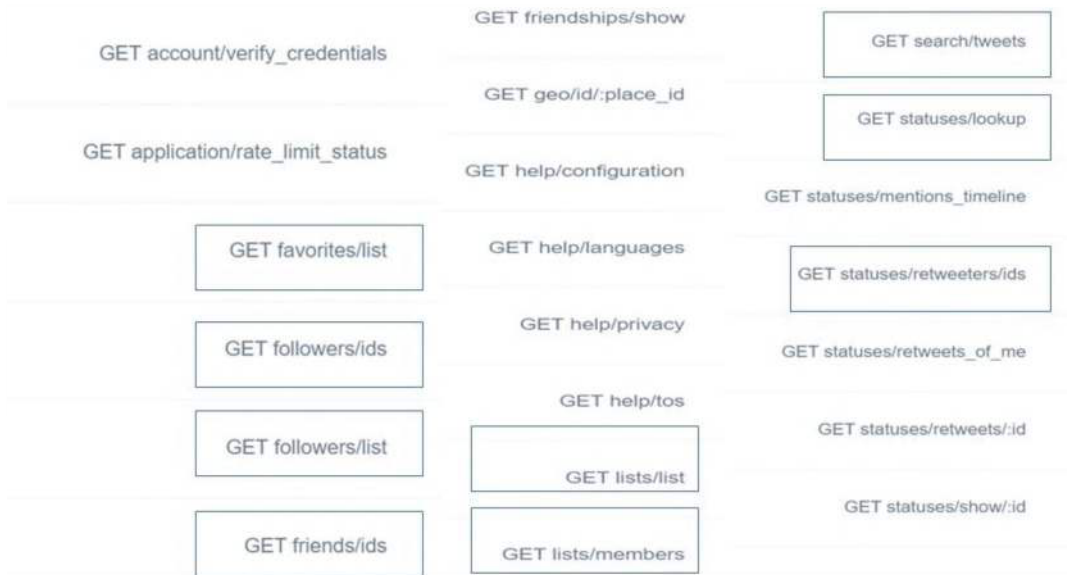
Extracted data from social media will be used in two ways, as an option for border guard to make an additional check and gain access to that information, for example to read the travelers recent tweets, or check followers rather than conclusive data. Amongst the elite of social media with more than 100 000 000 register users twitter is the social media software that allows you to interact with its data such as the tweets and networks of followers by using twitter APIs. A server side scripting language to make requests to twitter API is necessary and results are in JSON format that can be easily read.



**Figure 24 Twitter REST API for ELSI**

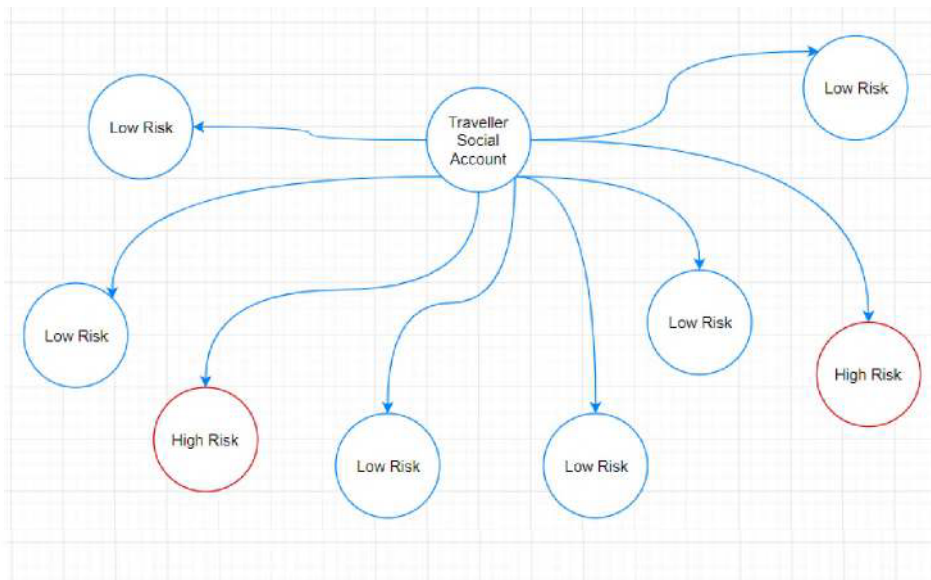
An example of calls and responses to the API was presented in D2.1. At the current stage of development of the ELSI system, twitter was selected as the social platform to use for piloting, whereas it has already been incorporated to the system. However, it's worth noting that legal and ethical challenges of using some the potential of such technology for border control are great and would likely require policy changes to enable their deployment in real life.

At the amount of information available through the twitter API is great, below on Figure 22 you can see some of the variables with some of the ones that ELSI will focus on identified in Squares.

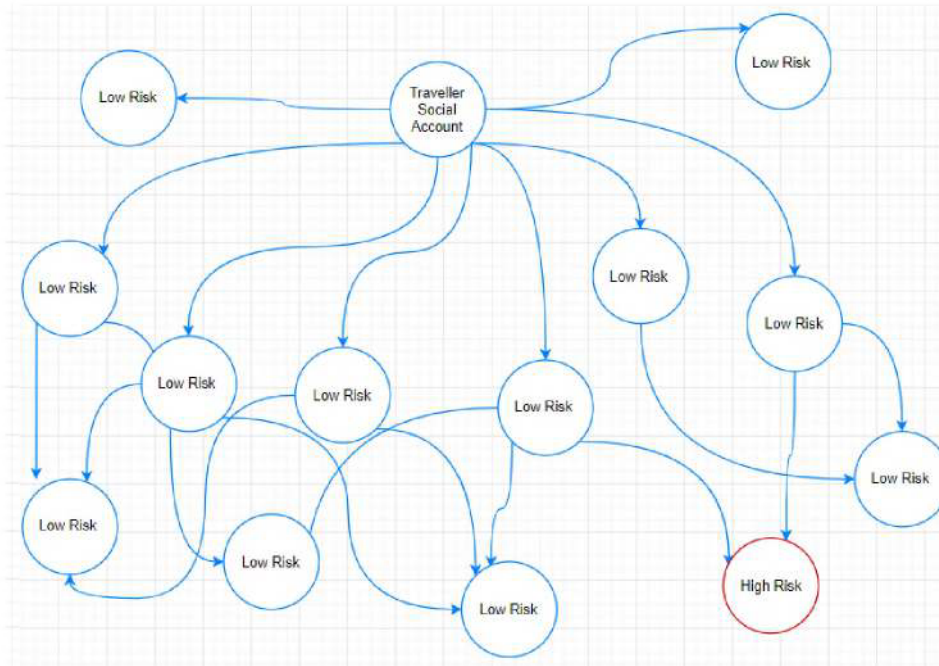


**Figure 25 Twitter API Variables**

One analyses planned for ELSI is the calculation of risk scores by comparing followers of users with known high risk accounts and providing schematics to demonstrating that to border agents so they can follow up on these individuals with targeted interviews when they get to the border. However legal and ethical review of these approaches are pending right now and will only be executed in pilots if they are approved. Figure 26 demonstrated how high risk individuals could be calculated based on pre-existing knowledge of high risk twitter accounts with 1 degree of separation from the traveller, and Figure 27 shows how this can be expanded to more degrees of separation.



**Figure 26 Direct Followers Risk Assessment ELSI**



**Figure 27 Multilevel followers risk assessment**

Less controversial approaches with regards to legal and ethical issues can be used when analyzing social media accounts to validate nationality, gender, and other personal information with the information on their social accounts, status, and place (geographical location). The risk estimation methodology is still undergoing development and its final form will take into consideration any legal and ethical issues identified by the legal expert partners in the group, as well as the ones at the pilot sites to ensure seamless execution of the pilots.

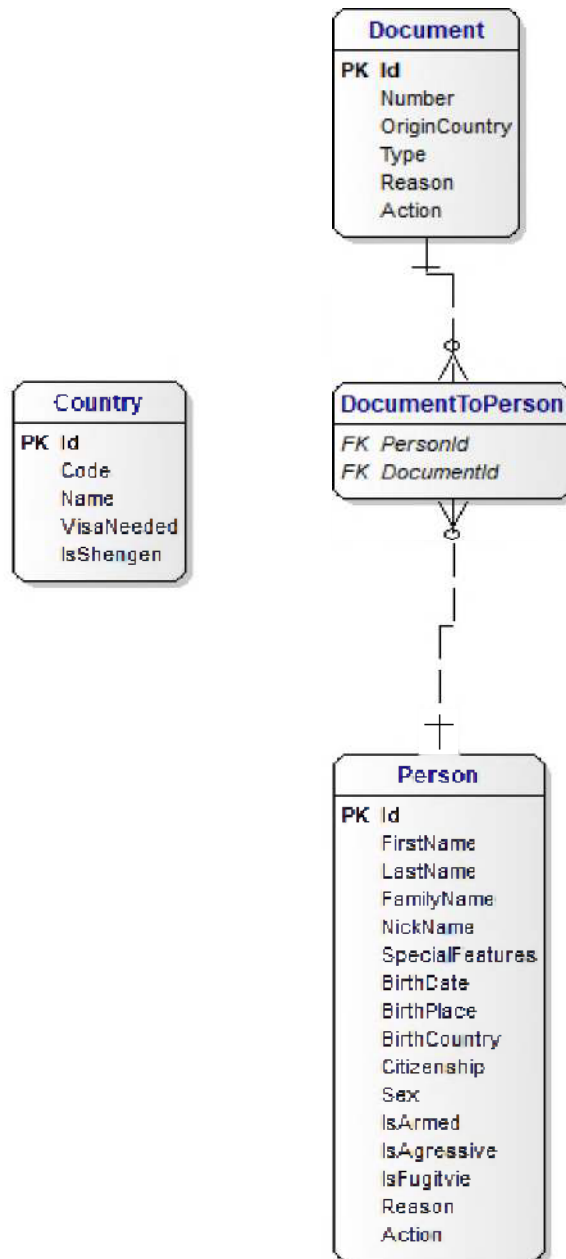
## 4.2 Border control consults system (VIS, SIS)

ELSI module is responsible for the connection with the legacy databases, SIS VIS and Entry Exit System Database. iBorderCtrl reviewed the process and attempted to get support to gain access to either the real databases, something that was found to be impossible due to confidentiality reasons, or to the mock-up databases generated for use by border control authorities. Based on the recommendation of partners in the project that represent border control agencies we identified as the only feasible solution for the pilots of this project to develop simulated versions of these databases and their respective API that would reflect in structure the real ones and populate these with dummy data that would be related to the subjects going through the pilots to ensure a realistic pilot deployment.



## 4.2.1 SIS, VIS, Entry Exist Databases

### 4.2.1.1 SIS Schema and Description



#### Country

Table stores information about countries.

- Id – Primary key
- Code – Country code
- Name – Country name
- VisaNeeded – Is visa required?
- IsShengen – Is the country a Schengen zone member?

## Document

Table stores information about documents.

- Id – Primary key
- Number – Document number
- OriginCountry – Identifier of the country signaling warning about the document
- Type – Identifier of the warning type
- Reason – Identifier of the reason for signaling warning
- Action – Action identifier

## DocumentToPerson

Table connects information about documents and persons.

- PersonId – Country Id
- DocumentId – VIS Id

## Person

Table stores information about reasons for arresting persons.

- Id – Primary key
- FirstName – First name of the arrested
- LastName – Last name of the arrested
- FamilyName – Family name of the arrested
- NickNames – Nicknames of the arrested
- SpecialFeatures – Special features
- DateBirth – Birth date of the arrested
- BirthCountry – Country of the birth
- Citizenship – Citizenship of the arrested
- Sex – Sex of the arrested
- IsArmed – Is the arrested person armed?
- IsAggressive – Is the arrested person aggressive?
- IsFugitive – Is the arrested person fugitive?
- Reason – Identifier of the reason for the person arrest
- Action – Action identifier

### 4.2.1.2 VIS Schema and Description

Country

Dictionary table storing information about countries.

- Id – Primary key
- Code – Country code

- Name – Country name
- VisNeeded – Is visa required?
- IsShengen – Is the country a Schengen zone member?

#### Photo

Table stores information about photos.

- Id – Primary key
- PhotoData – Photo in the binary format

#### Subject

Table stores information about fingerprints of the applicant.

- Id – Primary key
- SubjectId – Fingerprints identifier
- Template – Fingerprints in the binary format

#### VIS

Table stores information about reasons for arresting persons.

- Id – Primary key
- VisNumber – Primary visa number
- VisOtherNumber – Secondary visa number
- Issuingcountry – Country issuing visa
- StarDate – Date of issuing visa
- EndDate – Expiration date of visa
- LengthOfStay – Maximum allowable duration of the stay in the country
- State – Visa status
- Type – Visa type
- DocNumber – Document number
- DocEndDate – Document expiration date
- TripReason – Reason of entering the country
- AllowedTerritory – Is the entry of the specific country allowed?
- TravelLimit – Number of allowed entries to the specific country
- FirstName – First name
- LastName – Last name
- BirthDate – Date of the birth
- BirthPlace – Place of the birth
- BirthCountry – Identifier of the country of the birth
- Citizenship – Identifier of the country of origin
- Sex – Sex
- FingerPrints – Number of acquired fingerprints
- Photo – Photo identifier

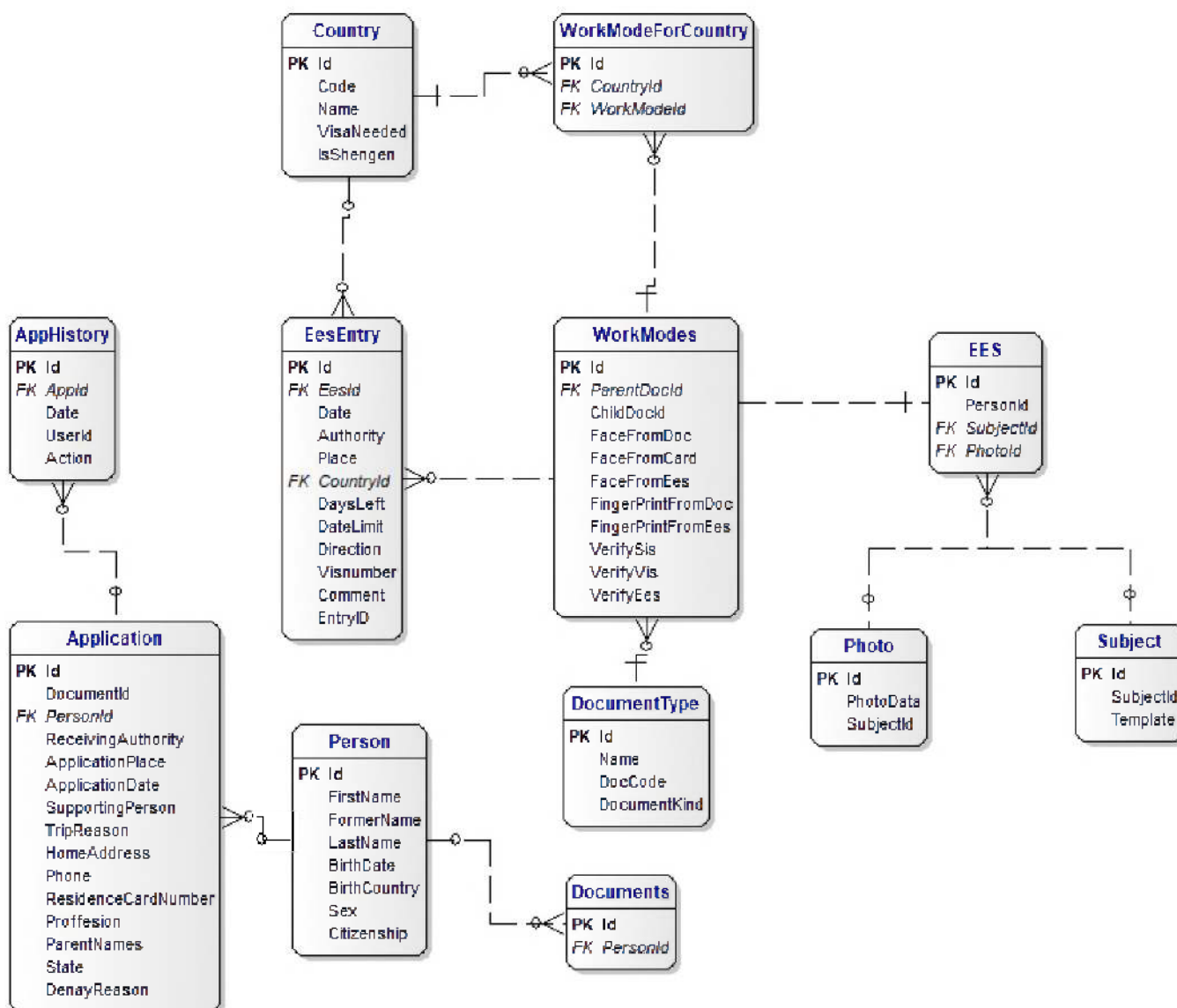
- Subject – Fingerprints identifier
- FingersNotNeeded – Are fingerprints required?

VisRestrictedAllowedTerritoryToCountry

Table stores information about countries, which can be entered using the particular visa

- Id – Visa identifier
- CountryId – Country identifier

### 4.2.1.3 EES Schema and Description



#### AppHistory

Table stores data about the application history.

- Id – Primary key
- AppId – Application identifier
- Date – Date of the application modification
- UserId – Identifier of the user modifying the application

- Action – Information about changes made in the application

## Application

Table stores information about the document type.

- Id – Primary key
- DocumentId – Document identifier
- PersonId – Person identifier
- ReceivingAuthority – Authority receiving the document
- ApplicationPlace – Place of filing the application
- ApplicationDate – Date of filing the application
- SupportingPerson – Person covering living expenses
- TripReason – Reason of the crossing the border
- HomeAddress – Home address of the applicant
- Phone – Phone number of the applicant
- ResidenceCardNumber – number of the residence card
- Profession – Applicant's profession
- ParentsNames – Names of applicant's parents
- State – Country of filling the application
- DenayReason -The reason of the application denied

## Country

Table stores information about countries.

- Id – Primary key
- Code – Country code
- Name – Country name
- VisNeeded – Is visa required?
- IsSchengen – Is the country a Schengen zone member?

## Documents

Table stores information about the document type.

- Id – Primary key
- PersonId – Person identifier
- DocType – Type of the document
- DocNumber – Number of the document
- DocIssueDate – Date of issuing the document
- DocEndDate – Date of the document expiration
- DocIssuingCountry – Country issuing the document
- DocIssuingAuthority – Authority issuing the document
- VisNumber – Number of the visa sticker

## **DocumentType**

Table stores information about the document type.

- Id – Primary key
- Name – Name of the document type
- DocCode – Code of the document
- DocumentKind – Is this a residence permission?

## **EES**

Table stores information about EES profiles.

- Id – Primary key
- PersonId – Person identifier
- SubjectId – Fingerprints identifier
- PhotoId – Photo identifier

## **EesEntry**

Table stores information about entry/leave history.

- Id – Primary key
- EesId – EES profile identifier
- Date – Date of crossing the border
- Authority – Authority entering the information about crossing the border
- Place – Location of crossing the border
- CountryId – Country identifier
- DaysLeft – Remaining number of days of permitted stay in the country
- DateLimit – Expiration date for permission of staying in the country
- Direction – Direction of crossing the border (entry of leave)
- VisNumber – Visa number
- Comment – Additional comment
- EntryId – Identifier of crossing the border

## **Person**

Dictionary table storing information about the document type.

- Id – Primary key
- FirstName – First name
- FormerName – Family name
- LastName – Last name
- DateBirth – Birth date of the arrested
- BirthCountry – Country of the birth
- BirthPlace – Place of the birth
- Citizenship – Citizenship of the arrested

- Sex – Sex of the arrested

### **Photo**

Table stores information about photos.

- Id – Primary key
- PhotoData – Photo in the binary format
- SubjectId – Photo identifier

### **Subject**

Table stores information about fingerprints of the applicant.

- Id – Primary key
- SubjectId – Fingerprint identifier
- Template – Binary representation of the fingerprints

### **WorkModels**

Table stores information about required verification, depending on the document type.

- Id – Primary key
- ParentDocId – Identifier of the stay document
- FaceFormDoc – Is the face verification from document required?
- FaceFromCard – Is the face verification from the residence card required?
- FaceFromEes – Is the face verification from the EES required?
- FingerprintFromDoc – Is the fingerprints verification from the document required?
- FingerprintFromEes – Is the fingerprints verification from the EES required?
- VerifySis – Is the verification in the SIS required?
- VerifyVis – Is the verification in the VIS required?
- VerifyEes – Is the verification in the EES required?

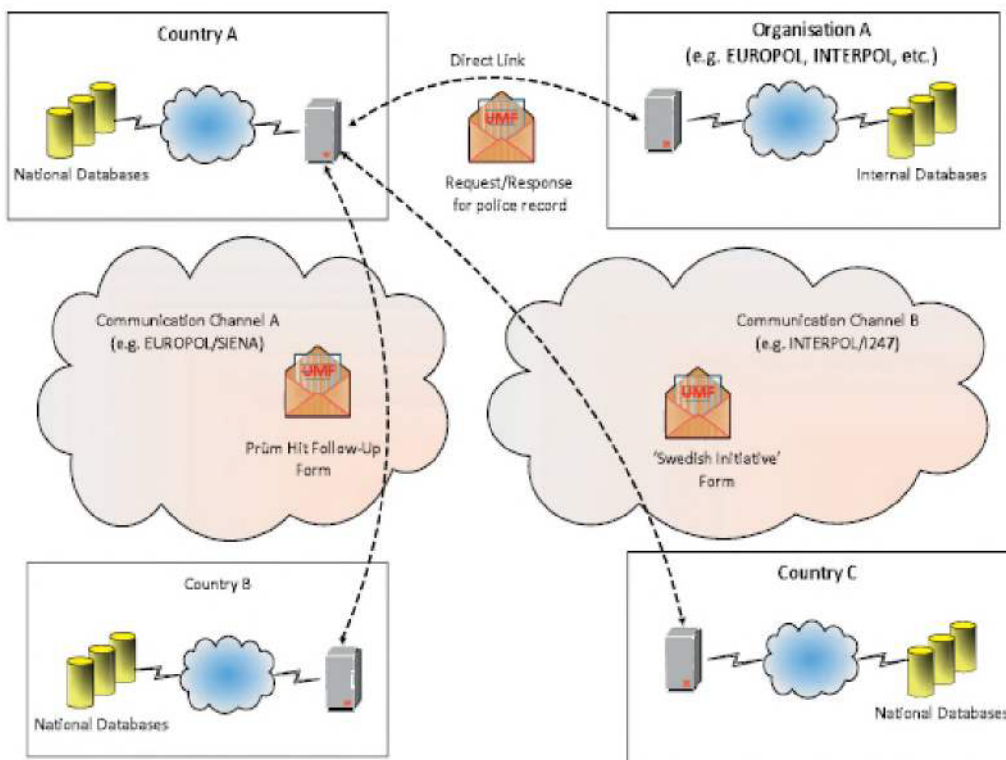
### **WorkModesForCountry**

Table stores information about required verifications for the particular country.

- Id – Primary key
- CountryId – Country identifier
- WorkModelId – Identifier of required modifications

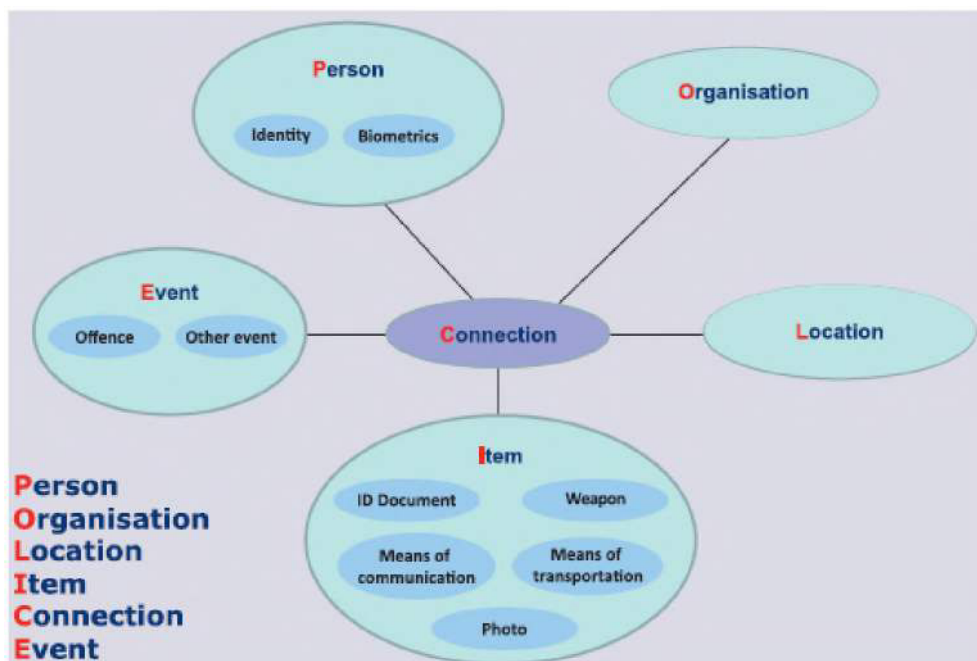
#### **4.2.1.4 Universal Messaging System (UMF)**

Communication between iBorderCtrl and SIS VIS and EES will be achieved through the Universal Message Format (UMF). UMF is defined by Europol a standard or agreement on what the structure of the most important law enforcement concepts when they are exchanged across borders should be. In other words, UMF is a set of concepts (building blocks) to construct standard data exchanges for interconnecting dispersed law enforcement systems. It must be emphasized that UMF is not the internal structure of systems/databases (you are not required to change your national systems, legislation or processes!) but rather an XML-based data format acting as a layer between them to be used whenever structured messages cross-national borders.



**Figure 28 UMF as a layer between systems (source 12)**

The UMF building blocks are the most important, usually exchanged, or cross-checked concepts from a law enforcement viewpoint: Person, Organisation, Location, Item, Connection, Event — creating the ‘POLICE’ Information Model.



**Figure 29 The POLICE information model (source 12)**

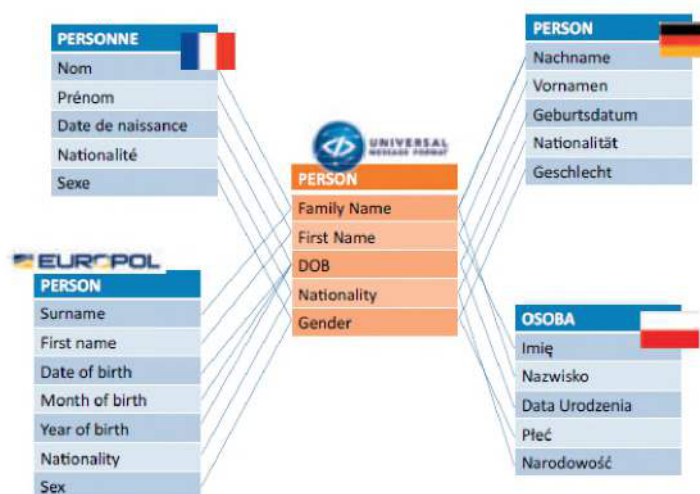
<sup>12</sup> UMF (Universal Messaging Format) Europol Programming Document 2017 – 2019 (The Hague, 17 January 2017 Europol Document)



As iBorderCtrl is an RIA project although the consortium made several attempts to acquire access to testing or real implementations of SIS VIS and EES from multiple countries these were unsuccessful, but the goals of the project could be achieved as well through simulated datasets made specifically for the purposes of the project. As such each system database has been developed for use in the pilots of iBorderCtrl with UMF as the messaging protocol between iBorderCtrl and the national system databases.

**UMF Mapping, enabling fast and cheap deployment**

iBorderCtrl is developed primarily by industry with the objective to one day move beyond research outcomes and onto real products on the market. To achieve that, focus was put on developing each aspect of the system in a way that allows for quick real world deployment following the completion of the funding period of iBorderCtrl. The Border Control Consults system achieves this by developing all messaging through the UMF and relying on a mapping build based on all public information available by Europol, and were needed extended following their stated rules. The mapping is the key enabler in the UMF that would allow a fast deployment to the real system as all that would be necessary to deploy the system are the mappings to each countries systems to enable messaging between them.



*Figure 30 UMF mapping (source 13)*

**4.3 ELSI technical requirements coverage**

Section 2.1.5 covers some of the requirements ELSI is responsible to cover and how it manages to do so. In this section we will describe some additional technical requirements as presented in deliverable D2.2. and how ELSI covers those requirements. The next table will list the requirements related with the module and the implementation description.

*Table 16 ELSI Technical Description*

<sup>13</sup> UMF (Universal Messaging Format) Europol Programming Document 2017 – 2019 The Hague, 17 January 2017 (Europol Document)



## 5 Border Control Analytics Tool (BCAT)

### 5.1 User Interface and pre-requisite knowledge required

This section is aiming to introduce the BCAT tool which is based on state of the art technologies and will be used to explore the iBorderCtrl Database data in order to discover actionable knowledge that otherwise could be missed. BCAT state of the art statistical algorithms are divided in specific categories based on the way they act to come to a conclusion, as it will be presented in the following paragraphs.

BCAT will be seamlessly integrated through the BMUA. To ensure that a user-friendly system is presented to the border manager and specific scientific workflows will be developed and will be made available with names and descriptions targeting the border manager with no analytic background. These will result in dashboard-like presentation of real-time information using visualizations and key metrics.

However, advanced options will also exist that will allow border managers with the knowledge and expertise to perform analyses by utilizing the full customizable power of the platform, these could be done for one-off analyses that aim to test specific hypothesis regarding border control data, and, were relevant, deploy these analyses as expansions to existing dashboards, or new ones to help their non-analyst colleagues make use of the new scientific workflows.

### 5.2 The underlying scientific analyses algorithms

BCAT will enable the metadata analyses of all travellers that have gone through the system and perform combinatorial analyses based on statistical modelling and data mining approaches. BCAT is key in enabling advanced post-hoc analytics that will help identify new patterns and knowledge allowing iBorderCtrl to adapt to new situations quickly and at the same time enabling the advanced exploratory analyses of the collected data to evaluate each model's performance.

It is based on a scientific workflow platform that includes:

- Univariate analysis tools: collect statistics on the performance of each individual tool
- Interaction analysis: tools to discover information interactively
- Feature selection, dimensionality reduction: to identify key high/low performing tasks
- Pattern discovery: to enable the automated discovery of key patterns in data, or results of primary analyses of travellers

BCAT is able to discover key patterns in the data associated with either false accept or false rejects of travellers. These can lead to better decision making at border control by expanding the patterns in RBAT, as well as identifying vulnerabilities in the security infrastructure of border control, enabling better focusing of resources to improve the performance of future system upgrades to address these. BCAT includes the state of the art on correlation analysis algorithms, on dimensionality reduction algorithms, on clustering algorithms and on clarification algorithms. The combination of those statistical approaches with state of the art machine learning on the outcome of each module will be the formula that will decide the effectiveness of each module based on their performance to the human border control. BCAT offers also the capability of providing key metrics relevant to expected traffic on specific planned travel dates, and to border guards traffic, and risk levels to allow for better planning. In this section we introduce the most important algorithms that will be included in BCAT. To demonstrate the potential impact of the tool and iBorderCtrl by extension we present for each algorithm a sample application where it would be used. Each application will be composed of an experimental design that includes clear indications of what data will be used and how, and what the output will be. A short discussion section will explain how to interpret the outcomes in each possible case. As the number of algorithms and the potential applications are in the thousands, we focus on key samples that will be performed at minimum as part of the funding period of the project to help meet the project's objectives related to the evaluation of the platform, as well as individual tools, and to determine and identify tools that when used together improve the overall performance of the platform by eliminating each other's vulnerabilities.

## 5.3 Evaluation of individual Border Control systems

The best way to evaluate providing convincing scientific evidence derived from empirical data collected through the pilots is to test specific a-priory defined hypothesis related to the efficacy of specific modules. These may be univariate, testing for risk related scores derived from a specific system to eventual outcomes of travellers, or they could be multivariate, linking the outcomes of multiple systems to test if jointly they offer an interaction effect that significantly improves their performance. Most methodologies were presented in D2.1. Some of those descriptions and schematics are also replicated here to assist the reader when reading the showcase scenarios to better comprehend the tools used from an analytical point of view.

### 5.3.1 Association testing

We will identify variables that are linked together, for example, some of the deception detection questions and outcomes of specific tools. The most appropriate algorithm will be selected based on the showcase. This will depend on the nature of variables involved (categorical, linear etc.) and the distribution of the collected data (linear or not to select parametric or non-parametric methods).

#### 5.3.1.2 Showcase Falsified Travel documents (VISA)

Visa related information collected by the user will be evaluated against the VIS database when possible. Specific hypothesis will be tested to see if there is a statistically significant association between the detection through the DAAT and FMT tools, and the user provided information at the pre-registration stage compared to the information in the VIS database. This showcase will focus on non-obviously falsified travel documents, such as:

- Stolen Identity: Valid visa but not belonging to the traveller, were the FMT tool should be able to identify it's a different person, by matching the passport picture.
- Manipulated original documents: Cases where a valid visa was once provided to the traveller but is no longer valid, and the traveller attempts to modify the visa to appear valid, such as changing specific dates on the document. The DAAT tool performance may be evaluated in these cases where through the VIS database we can identify the real expiration date of that visa.

## 5.4 Knowledge Discovery and pattern identification

### 5.4.1 Advanced exploratory analyses of the collected data

In the case of correlation analysis, the interest is in identifying any statistical relationship between two independent variables. Through that algorithm in BCAT, we can reveal correlation of any variable in the iBorderCtrl database that will lead to an estimation of the effectiveness of the system. More specifically, we can use iBorderCtrl system outcomes to see which systems behave similarly making joint deployment redundant, and cases where there is an interaction effect, making joint deployment beneficial.

### 5.4.2 Correlation Coefficients

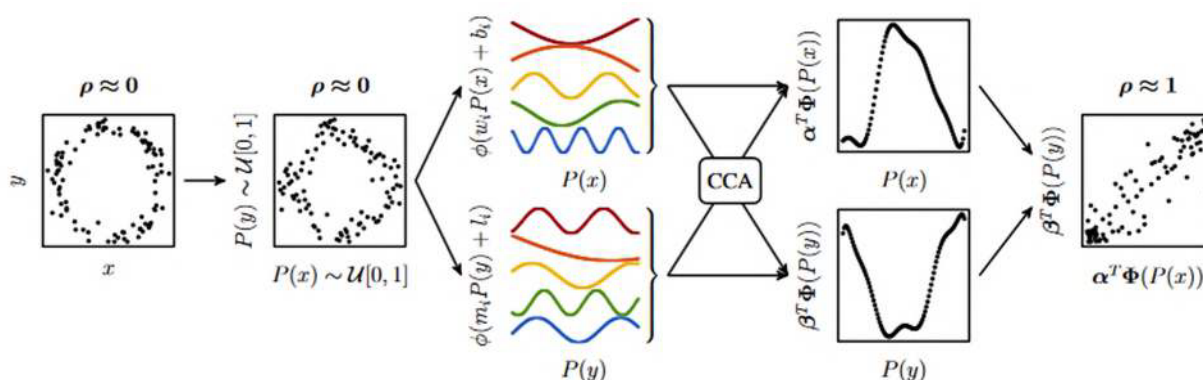
To obtain such a result there are various approaches, in BCAT though we will use Pearson's Correlation Coefficient an approach that measures the degree of linear correlation between two variables. The returned

correlation is in the range  $[-1,1]$ , with zero denoting no correlation among the variables, 1 a perfect increasing correlation (when there is an increase in the value one of the variables there is a linear increase in the value of the other variable as well), and -1 a perfect decreasing correlation. The measure for two random variables  $X$  and  $Y$  is calculated with the following formula:

$$\text{corr}(X, Y) = \frac{\text{cov}(X, Y)}{\sigma_x \sigma_y}$$

where  $\text{cov}$  is the covariance and  $\sigma$  is the standard deviation.

Another approach establishing the idea of correlation analysis is the Randomized Dependence Coefficient (RDC) (Lopez-Paz et al., 2013) is a measure of non-linear dependence between random variables based on the Hirschfeld-Gebelein-Renyi Maximum Correlation Coefficient. It is invariant to monotonically increasing transformations and it is ranged to  $[0,1]$ , with zero indicating no correlation and one indicating a complete correlation among the two variables. The RDC process for the variables  $x$  and  $y$  drawn from a noisy circular pattern is shown in the following figure. The samples are used to estimate the copula, and then are mapped with randomly drawn non-linear functions. The RDC is the largest canonical correlation between these non-linear projections.



**Figure 31 RDC process for a noisy circular pattern**

### 5.4.2.1 Correlation Coefficients Showcase

In this showcase, all pairs of tools producing risk scores could be tested. Priority will be given to those pairs where some overlap or potential combinatorial synergy is expected. As each tool in iBorderCtrl is independently developed and integrated in the platform, there's a need for analytical methodologies that make minimal assumptions with regards to the nature of the risk scores at each individual tool. Correlation analyses will allow testing for correlation between variables while addressing any inherent bias caused by the distribution of any one specific risk score.

In this showcase, we expect to identify pairs of modules that are compliant or replicate one another. For example, if DAAT and FMT applied at the pre-registration stage both perform perfectly, they may both identify the same subjects. Thus, revealing that either tool could supplement the other (this is a hypothesis to test, we do not expect this to be the case). In another example, it could be that when focusing on just subjects that are known to be illegal travellers, when looking at both tools together, where DAAT fails to detect high risk (for example because documents are real but belong to another person) FMT steps in and identifies the weak point of DAAT (real documents, but stolen identity). Having this analysis will determine the overall efficacy of the system when deploying both systems together and would reveal the pattern and relevant coefficients that could be used to derive weights to use in order to optimize the system based on previously collected data.

Machine Learning: Dimensionality Reduction, Principal Component Analyses and Random Forest Dimensionality Reduction can be used either for transforming the data into lower dimensions using the

information of all available variables/features, or selecting a subset of the available features, so as to remove the ones that are redundant or irrelevant. This can be used for visualizing data in lower dimensions and for helping creating simpler models, which are less prone to overfitting.

### 5.4.2.2 Principal Component Analyses

Principal component analysis (PCA) (Abdi and Williams, 2010) creates linearly uncorrelated variables (Principle Components) using orthogonal transformations. The resulting PCs are less than or equal to the number of the initial features. The largest variability in the data is explained by the first PC, then by the second and so on. PCA can be used for visualizing high dimensional data into lower dimension (e.g. 2 or 3) that can help in exploratory analysis. Additionally, it can be used for noise removal.

The HSIC Lasso (Yamada et al., 2014), is a feature-wise kernelized Lasso approach. Since it is feature-wise, the kernel is applied on each feature vector, which enables the identification of non-linear dependence between a feature and the response variable. To identify features that are not redundant and are associated with the response variable the Hilbert-Schmidt Independence Criterion (HSIC) is used. The HSIC Lasso is given in the following form:

$$\min_{a \in \mathbb{R}^d} \frac{1}{2} \left\| \bar{L} - \sum_{k=1}^d a_k \overline{K^{(k)}} \right\|_{Frob}^2 + \lambda \|a\|_1, \text{ s.t. } a_1, \dots, a_d \geq 0$$

Where  $\|\cdot\|_{Frob}^2$  is the Frobenius norm,  $\overline{K^{(k)}}$  is the centered Gram matrix computed for the k-th feature and  $\bar{L}$  is the centered Gram matrix computed from the target variable. The algorithm returns the selected features with non-zero coefficients.

### 5.4.2.3 Decision Trees with Random Forest

There are good reasons, and many methods for automating decision tree generation in the literature. However automated decision trees tend to fail unless there is a very large number of data points (in our case travelers) to analyze, and more specifically, a relatively large number of travelers matching each case we are trying to detect (meaning high risk travelers that succeed in fooling the system). There is however one method that is based on these concepts, and achieves reliable good results, and that is Random Forest. The Random Forest algorithm utilizes similar concepts to automating decision tree generation, but rather than focusing on the decision tree, it instead focuses on ranking variables in a regression or classification problem in a natural way. It works through an unsupervised learning algorithm; it allows mix variable types and is robust to outlying observations.

### 5.4.2.4 Showcase Dimensionality Reduction

Taking a snapshot at the schema of the iBorderCtrl database and it becomes clear that the number of variables is quite large. Coupled with BCAT's challenge that is to analyze all metadata collected at border control points across Europe with the goal of identifying complex relationships across multiple methods that may include millions of daily travellers across many border crossing points and the challenge in terms of the computational complexity is clear.

Dimensionality reduction helps by identifying the variables (Dimensions) that are less likely to provide useful information reducing the number of variables to analyse. This is a good step to apply when there's a need to reduce the number of variables either because we want to perform lower number of tests.

Principal Component Analyses on the other hand merges variables together into less dimensions (Principal Components) allowing complex analytical approaches to be performed with potentially less information loss compared to dimensionality reduction.

Through BCAT all data collected through the pilots will be used to run both methods. In the case of dimensionality reduction, we expect an identification of the less informative variables to be removed. Principal component will act as an enabler to allow complex analytics to take place on a smaller subset of variables that may improve the risk calculations at a per-traveller case.

### 5.4.2.5 Machine Learning Showcase

All methods mentioned in this section will be tested in BCAT on all variables. These methods are designed to take as input a large number of dimensions (Variables from the iBorderCtrl databases) and in the case of:

**Dimensionality Reduction:** Identify the dimensions (variables), that carry the least amount of importance based on our collected data

**Principal Component Analyses:** Identify a significantly smaller number of variables that are calculated based on the total number of variables (principal components) that may be used in analytics as a way to discover complicated patterns identified with specific sub-types of travelers (Bona fide for example).

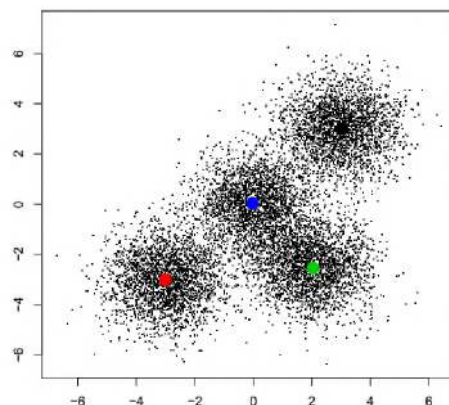
**Random Forest:** This algorithm will be used to rank the importance of all variables in order to identify the most interesting variables to focus on.

In reality, we expect these algorithms to be utilized in scientific workflows with the ones in other sections as a way to automate the handling of the high dimensionality nature of the database enabling a more intelligent way of performing the analyses that relies on automation thus allowing the border managers, who are non-experts from a computational sciences perspective, to perform complex analytics and increase the likelihood of actionable and accurate outcomes.

### 5.4.3 Clustering and Classification

Clustering is used for grouping data that have similarities together. The similarity between data is usually calculated using distance measures such as the Euclidean distance and the Hamming distance. Clustering can be used to group together travellers with similar patterns.

K-means (Hartigan and Wong, 1979) is a clustering algorithm that tries to partition data into  $k$  groups. The algorithm takes as input an  $M \times N$  matrix and  $k$  which is the number of clusters that will be created. Initially the centers of the clusters are randomly assigned in the data space. At each iteration of the algorithm, the subjects are assigned to the cluster whose mean is closest to them. The similarity measure used is usually the Euclidean distance. Then the means of each cluster are updated so that they are the mean of the subjects that are assigned in that cluster. This procedure is repeated either for a pre-specified number of iterations or until convergence. A drawback of  $k$ -means is that  $k$  needs to be defined a-priori. An example of a  $k$ -means clustering on data using  $k=4$  is shown in the following figure. Each coloured dot represents the center of the cluster, calculated by  $k$ -means.

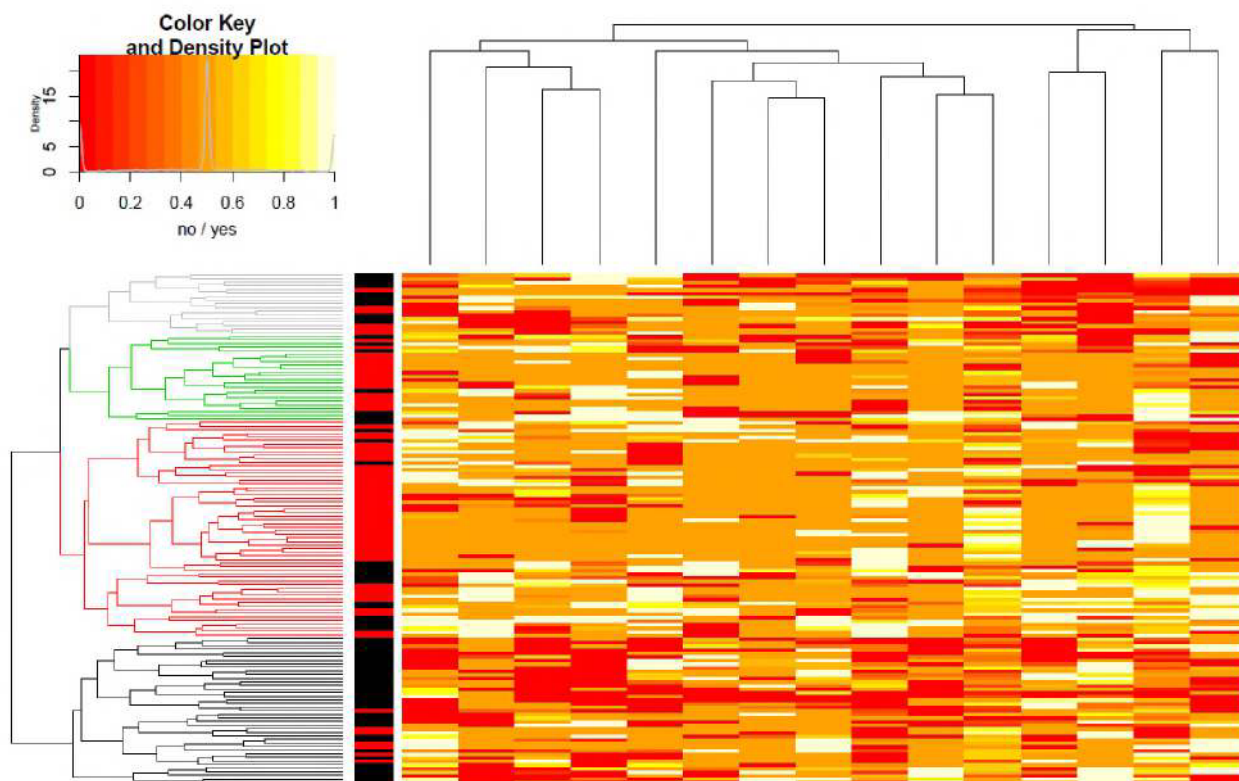


**Figure 32** A  $k=4$  clustering of data

Hierarchical clustering techniques do not require from the user to a-priori define the number of clusters (Hastie et al., 2009). In hierarchical clustering there are mainly two strategies, the agglomerative and the divisive strategy. The difference between these two strategies is that the first follows a bottom up approach, whereas the second one follows a top down approach.

In agglomerative clustering, at each step the two closest clusters get merged together until only one cluster is left. There are several methods for deciding whether two clusters are close. One approach is to consider the distance between two clusters as the distance between the closest pair among the two clusters also known as the nearest neighbour technique. An alternative of this method is to consider the distance between the furthest pair of the two clusters, known as the furthest neighbour technique. A compromise between these two methods is to consider the average distance between the two groups.

An example of hierarchical clustering is shown in Figure 33. Four clusters were created as shown with different colors on the left dendrogram. The heatmap shows the different values of the data in each cluster so as to visualize their differences.



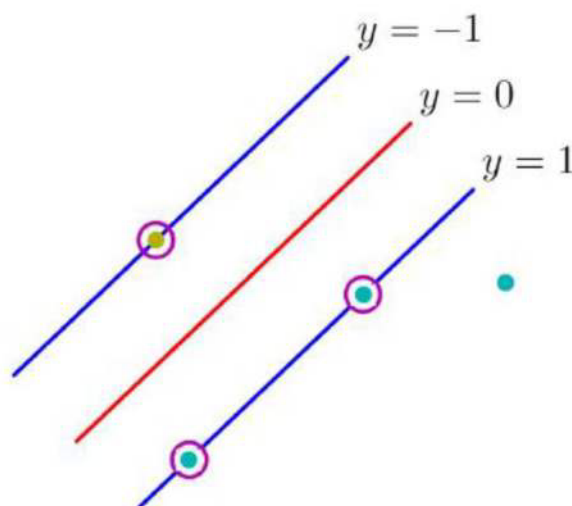
**Figure 33 Hierarchical clustering**

Classification algorithms try to predict the class of an instance using the knowledge they have from previously seen data. In BCAT classifiers will be used to aid the decision making in border control by modelling / identifying key patterns of travellers who should or should not pass into a country.

Random Forest (RF) (Breiman, 2001) is an ensemble of many de-correlated trees. Each tree is created using bootstrapping and the subjects not used are called the out of bag (OOB) samples, which are used for calculating the misclassification error. At each terminal node of the tree,  $m$  features out of  $p$  are randomly selected and the best feature out of them is used for further splitting the node. The trees are grown in full depth and no pruning is used. Majority voting is used for selecting the predicting class. The accuracy of the algorithm is calculated using the OOB error rate.

Support Vector Machines (SVMs) (Cortes and Vapnik, 1995) are also known as maximum margin classifiers. This is due to the fact that they try to find the decision boundary that has the maximum margin among the data points of the classes in the dataset. The margin is defined as the perpendicular distance among the decision boundary and the closest of the data points as shown in the following figure.





**Figure 31 Margin definition**

To cope with the non-linearly separable data, kernel functions are used for projecting the data in higher dimensions and converting non-linearly separable problems to linearly separable.

#### 5.4.3.1 Showcase Clustering and Classification for Pattern Discovery

Clustering and classification techniques can be utilized to identify clusters of travellers that exhibit common attributes. The aim is to identify clusters of patients that share either low or high risk consistently based on common factors between them. These clusters will then be used to identify the underlying patterns that characterize clusters of interest and will be used to expand the rule based applied on to the RBAT tool to increase the predictive power of future travellers. Hierarchical Clustering is a very powerful method that we intend to apply to this challenge through RBAT. It allows the simultaneous clustering across two axes, the travellers being one, and the second being the risk scores produced by each independent tool in iBorderCtrl. If the method is successful in identifying patterns, then it should be able to discriminate travellers (placing them with close hierarchical distance to each other) that share the actual border control outcome. Consider as an example a situation that due to changes in geopolitical forces a new type of illegal border crossing traveller arises. One that perhaps current systems fail to identify correctly at the border but are able to at a later time thanks to crimes, or other reasons allows the travellers to be revealed to the authorities. These methods could be used to test and see if the existing systems although they failed to identify these subjects individually, perhaps they behaved consistently among these travellers, and uniquely in their combined output compared to all other travellers. In this situation, the above-mentioned algorithms could be dissected to identify new patterns that could be incorporated in RBAT to allow the correct risk estimation calculation of these types of travellers for future crossings.

## 5.5 BCAT technical requirements coverage

Section 2.1.5 covers some of the actions BCAT is doing to meet the requirements. In this section we will describe some additional technical requirements as presented in deliverable D2.2. and how BCAT covers those requirements. The next table will list the requirements related with the module and the implementation description.



## 6 User interface

### 6.1 TUA implementation

#### 6.1.1 TUA interface implementation

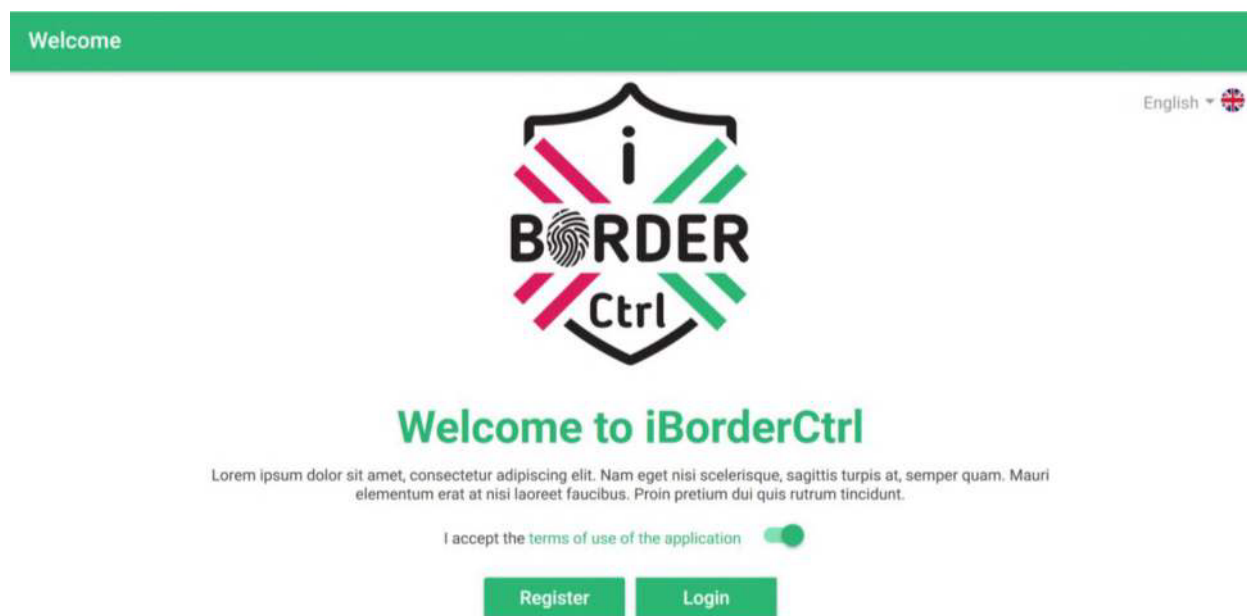
iBorderCtrl Traveller User Application is implemented in a common cross-platform client (both mobile and web) called Electron, featuring a number of REST services. Electron main advantage is that it operates independently of the Operating System and allows the creation of a web-application as well as a client application for Windows, Linux, Mac, mobile phones, etc. at the same time. More specifically, Electron is a framework for creating native applications with web technologies like JavaScript, HTML, and CSS and is combining **Chromium** and **Node.js** into a single runtime and its applications can be packaged for Mac, Windows, and Linux.

In the following paragraph, the main visualization aspects of the Traveller User Interface are presented in screenshots.

#### 6.1.2 TUA screenshots

In this paragraph, the screenshots of the developed Traveller User Interface are presented with respect to the TUA screens described in detail in Deliverable 2.2 section 7.1. It has to be noted that the avatar interview screen (Step 4/5) is under development


### Initial screen



*Figure 32 TUA Initial screen*

## Register screen

← Register

Please complete the following fields in order to register to the iBorderCtrl system English 

Firstname \*

Surname \*

Gender \*

Date of Birth \*

Nationality \*

Country of residence \*

Address \*

Mobile Telephone number \*

Email \*

Do you have single or multiple citizenship? \*

Twitter Account

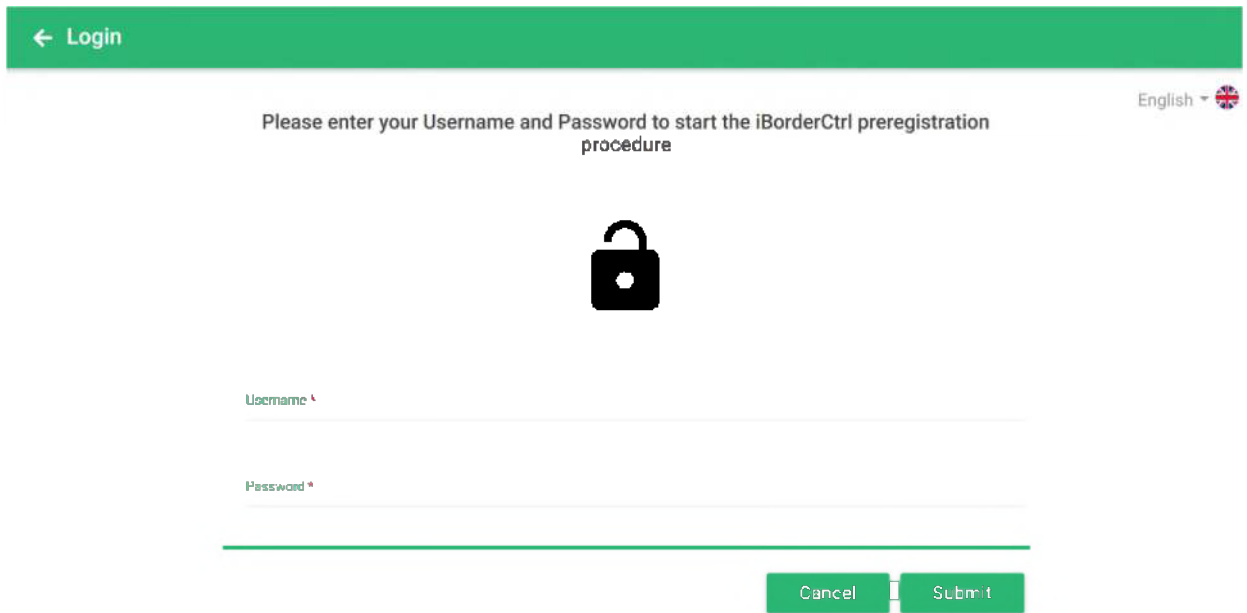
Login name \*

Password \*

Re enter your password \*

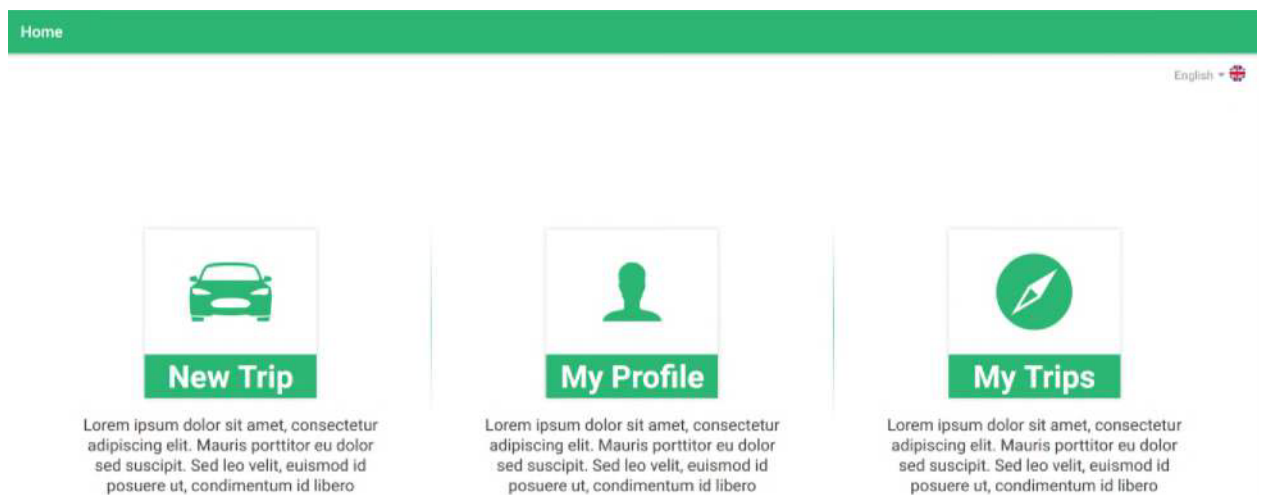
**Figure 33 TUA Register screen**

## Login screen



*Figure 34 TUA Login screen*

## Home screen



*Figure 35 TUA Home screen*

## Travel information screen (Step 1/5)

Wizard

English

Step 1/5

Please complete the following fields about your travel

Origin \*

Destination \*

Length of stay \*

Contact information

Hotel reservation

Purpose of Trip \*

Expected Date of Arrival \*

Expected Time of Arrival \*

Expected Date of Departure \*

Means of travel to your destination? \*

Who is paying for the expenses of your travel/stay? \*

Do you have health insurance? \*

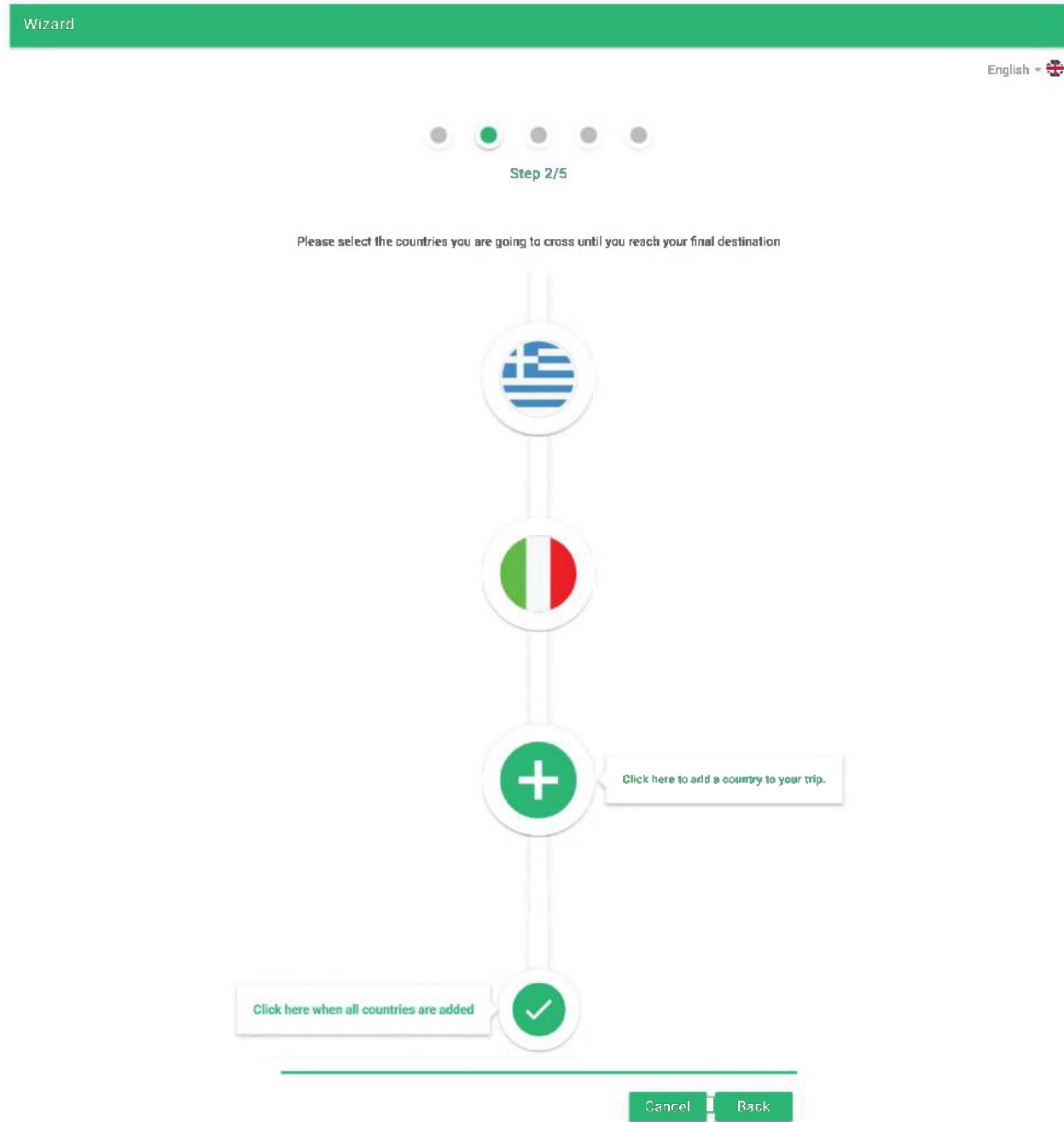
Have you ever been refused entry to the Schengen Area? \*

Have you ever been removed from the Schengen Area? \*

Cancel Next

**Figure 36 TUA Travel Information screen (Step 1/5)**

## Travel information screen (Step 2/5)



**Figure 37 TUA Travel Information screen (Step 2/5)**

## Document information screen (Step 2/5)

Wizard

English

Step 2/5

Please select the countries you are going to cross until you reach your final destination

**Country Data**

**Document Data**

Please complete the following fields about your Document

Will you travel for the current country using:

ID  Passport

Also complete using data for document:

Document number \*

Country issuing the document \*

Expiry date \*

Issued date \*

Document issuing office \*

Will you travel using visa for the current country?

Visa number \*

Country issuing the visa \*

Visa expiration date \*

Visa issuing office \*

Do you have a residence permit for the current country?

Residence Permit number \*

Country issuing the Residence Permit \*

Residence Permit Expiration Date \*

**Figure 38 TUA Document Information screen (Step 2/5)**



## Vehicle information screen (Step 2/5)

Wizard

English

Step 2/5

Please select the countries you are going to cross until you reach your final destination

Country Data ×

Vehicle Data

Please complete the following fields about your Vehicle

Will you travel using your own vehicle?

Auto-complete using data for vehicle: ▼

License plate \*

Insurance policy \*

Ownership \*

Driver license number \*

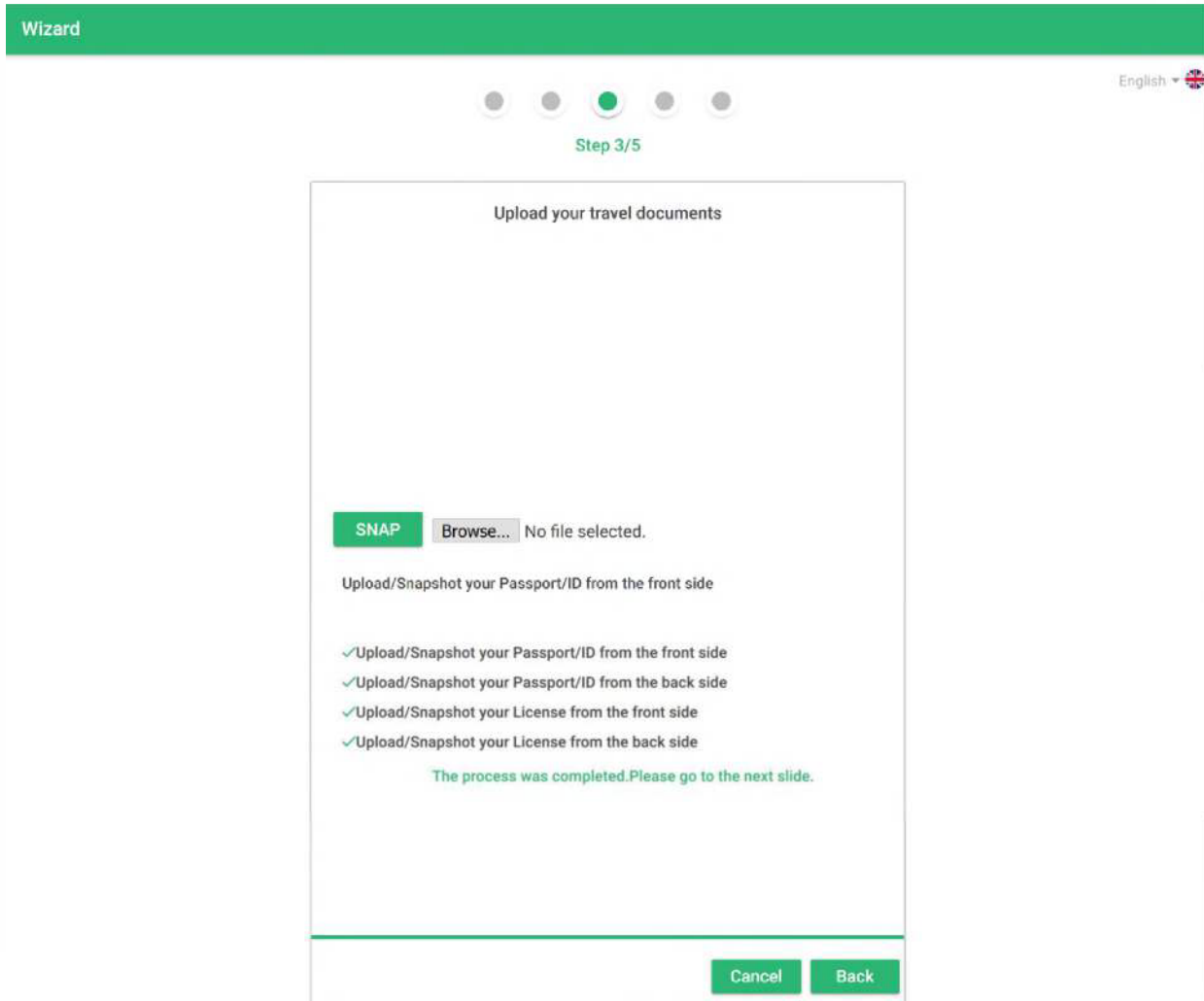
Issued country \* ▼

Will you exit from the border of that country with different vehicle?

Clear Done

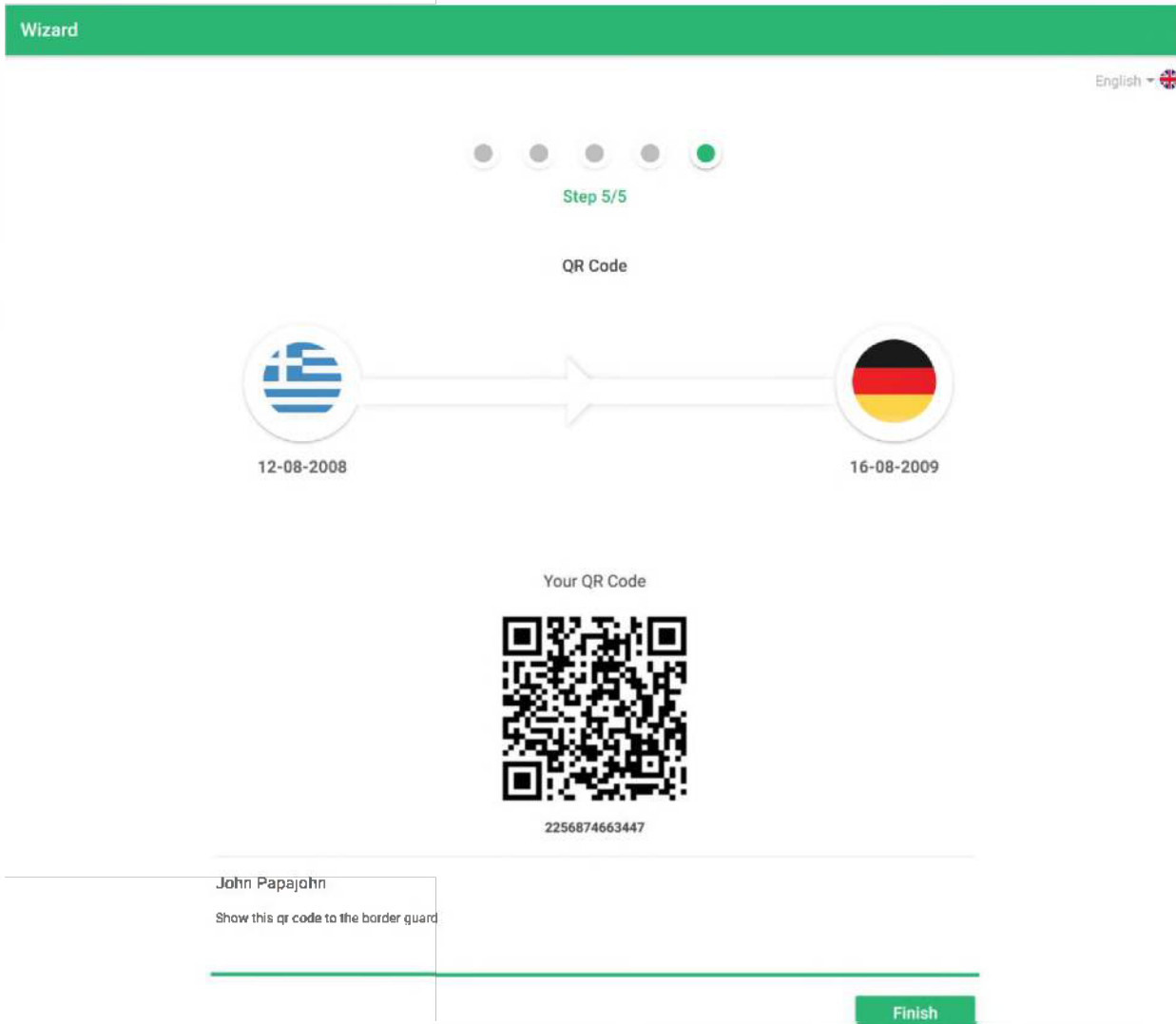
**Figure 39 – TUA Vehicle Information screen (Step 2/5)**

## Document/vehicle upload screen (Step 3 /5)



**Figure 40 – TUA Document/vehicle upload screen (Step 3/5)**

## QR code screen (Step 5/5)



**Figure 41 TUA QR code screen (Step 5/5)**

## My profile screen

← Profile

User Account English

Username\*  
johnDoe14

Password\*  
●●●●●●●●●●●●●●●●

Firstname\*  
John

Lastname\*  
Doe

Gender\*  
Male

Nationality\*  
American

Date of birth\*  
Jan 29, 2018

Date of issue\*  
USA

Address\*  
Reverly Hills, 90210

Telephone number\*  
+1 (310) 276-2251

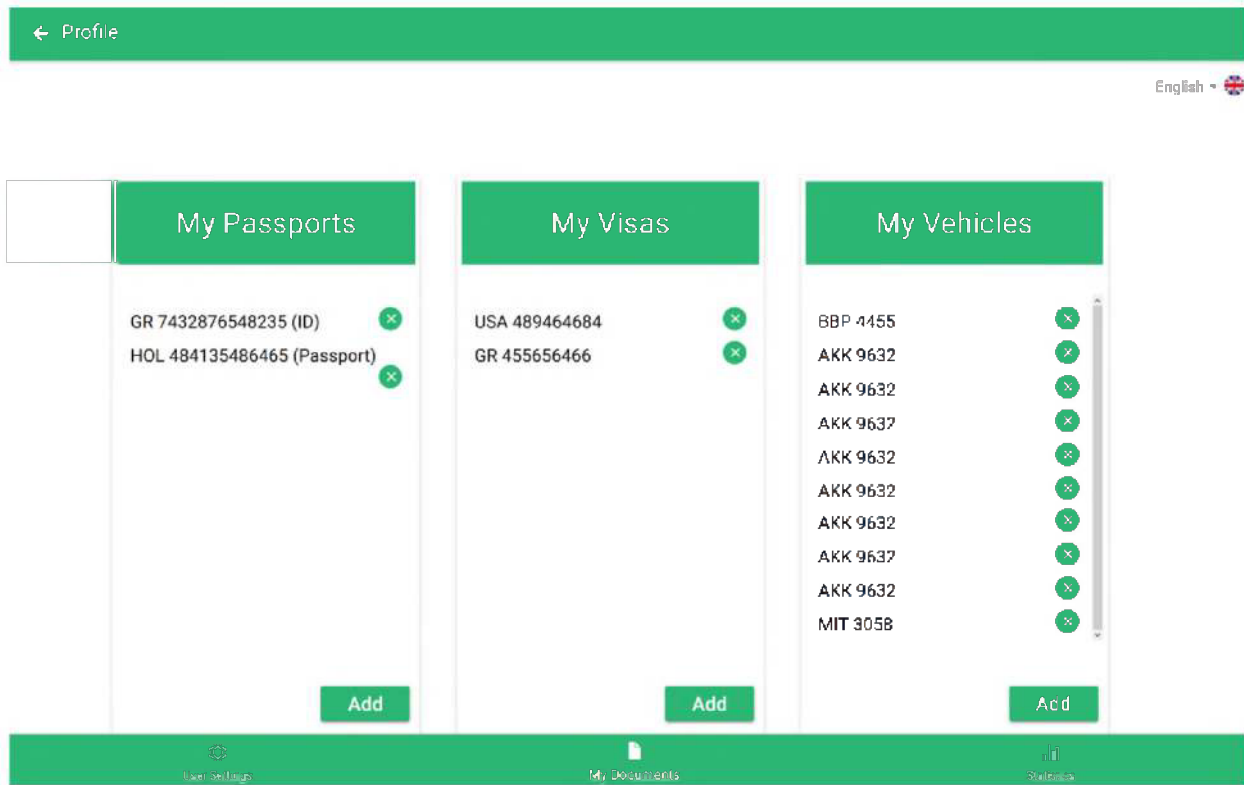
Email\*  
john.doe@rse.gov

Save

User Settings My Documents Statistics

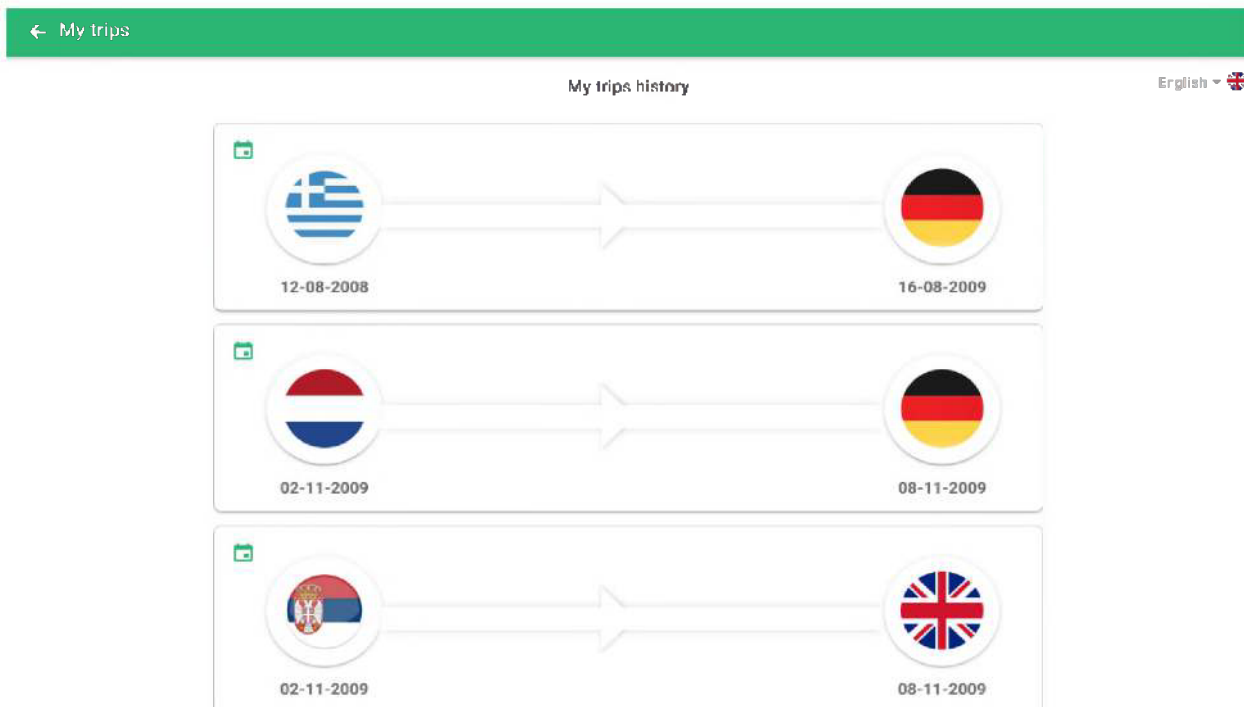
**Figure 42 TUA My profile screen**

## My documents screen



*Figure 43 – TUA My documents screen*

## My trips screen



*Figure 44 – TUA My trips screen*

## 6.2 BGUA implementation

### 6.2.1 BGUA interface implementation

The PU Border Guard Agent Interface (BGAI) workflow will be based on the “On-border” crossing point check general scenario that was described in the *WP2 Deliverable 2.2*. The iBorderCtrl officer is asked to enter his/her user name and password to get access to the application. All check screens are based on the chapter *7.2 Deliverable 2.2 User management and users interfaces – Agent User Interface For Border Guards*. On the top of the AUI, the application displays the following status information:

- access to the wireless networks,
- date, time and GPS data
- the current user login.

On the right side of the check screen, the border guard will be informed about the status of the BCP checks to be performed– the progress bar, connection status (between data acquiring devices and the tablet, as a processing unit) and battery charge level of each device. The active part of the traveler control is marked by the darker shade of gray. It is up to border guard whether and when the next part of the procedure should start. The process is triggered by the green START button.

All check screens instruct the Border Guard about the action to be performed. This is placed as the title in the middle of the screen in the following form: “Please, use XXX to scan the traveller’s YYY”, where “XXX” stands for the specific device and “YYY” stands for the specific data to acquire.

The middle part of the screen contains information regarding the acquisition result (Result Box) and the status of the process (like “Waiting for documents data...”). The border guard has to acquire the data within 10 sec (time will be counted down), after time run out the process could be restarted by clicking on retrieve icon.

The overall BCP screen offers the Border Officer two options to: proceed with the control or to turn back the traveler. The estimated risk information is displayed on the left side of the overall screen.

The current GUI of the application is presented in Fig. 6.3. It is the basic interface that will be expanded with the further development of the system. The window is divided into two parts:

- The main panel, responsible for presenting messages for the BG. Here all information about the current state of the application and commands for the BG will be presented (like encouragement to use the specific device, or the current risk assessment score).
- The control bar, containing three groups of objects. The first one is the information about the current user of the application (his/her name, login date and time), the status of devices connected to the PU (each device has its own icon, which can be red or green, depending on whether it is visible for the application or not), and configuration icons (used to perform the detection procedure and configuring the application).

The functionality of the software covers the following operations, which will be extended if needed:

- Detection of the state of all devices being part of the PU.
- Performing the data acquisition for each of available devices.
- Sending the acquired data to the local server, which will pass the information further, to each respective scanning processing module and to the central iBorderCtrl database when required.

The example of the data acquisition using the camera (in the current version the onboard tablet camera is used) is presented in the following Figure. The acquisition starts after clicking on the grey bar in the main part of the application.



**Figure 45 Main screen of the PU Application**



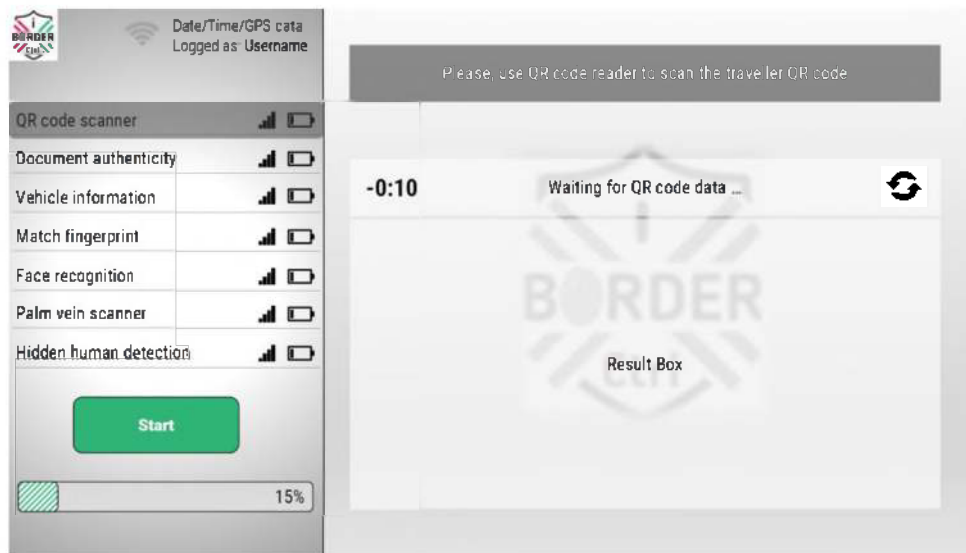
**Figure 46 Data acquisition screen for the camera**

### 6.2.2 BGUA Screenshots

In this paragraph, following the same structure as in the TUA ones, the screenshots of the presently under development Border Guard User Interface are presented with respect to the BGUA screens described in detail in Deliverable 2.2 section 7.2. Since, the development process is ongoing, the screenshots below, represent the current version which is a preliminary one to test the respective functionalities, while the final completed versions will be presented in full detail in the following Deliverable D5.2.

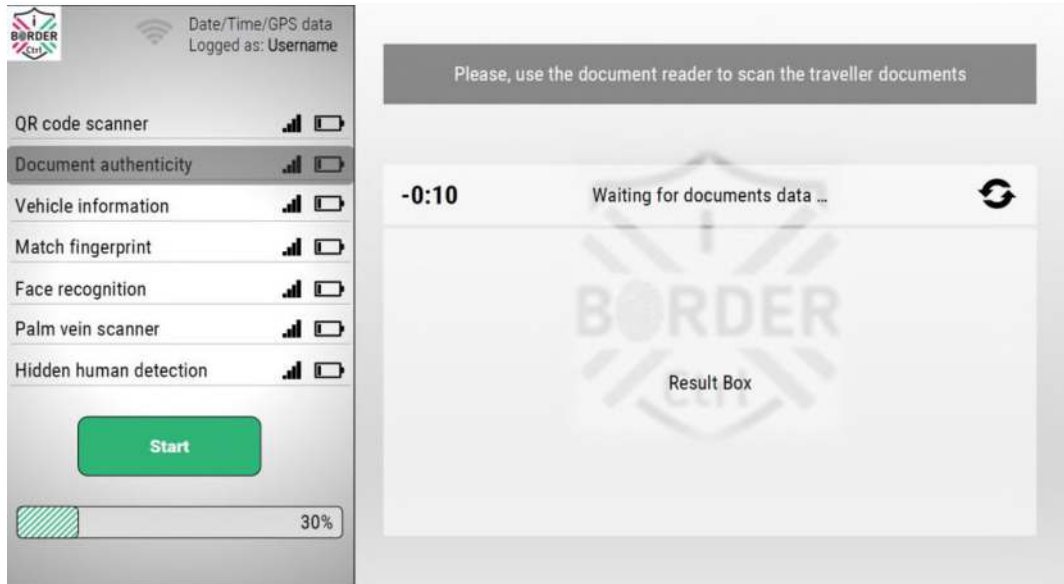


*Figure 47 Initial login screen of the PU Border Guard Agent User Interface*



*Figure 48 QR code check screen of the PU Border Guard Agent User Interface*

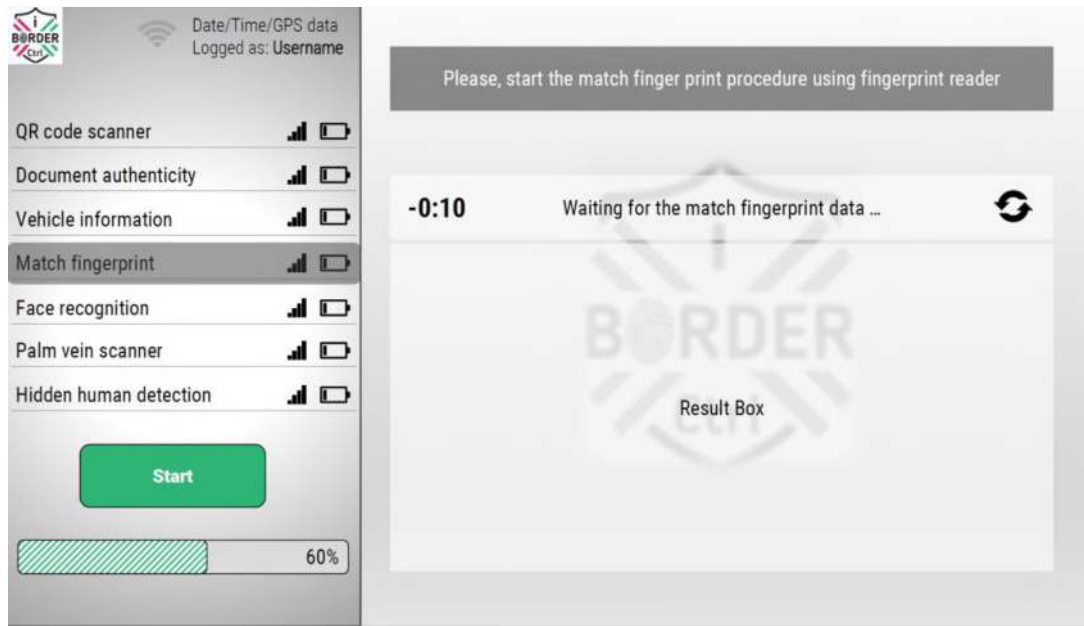




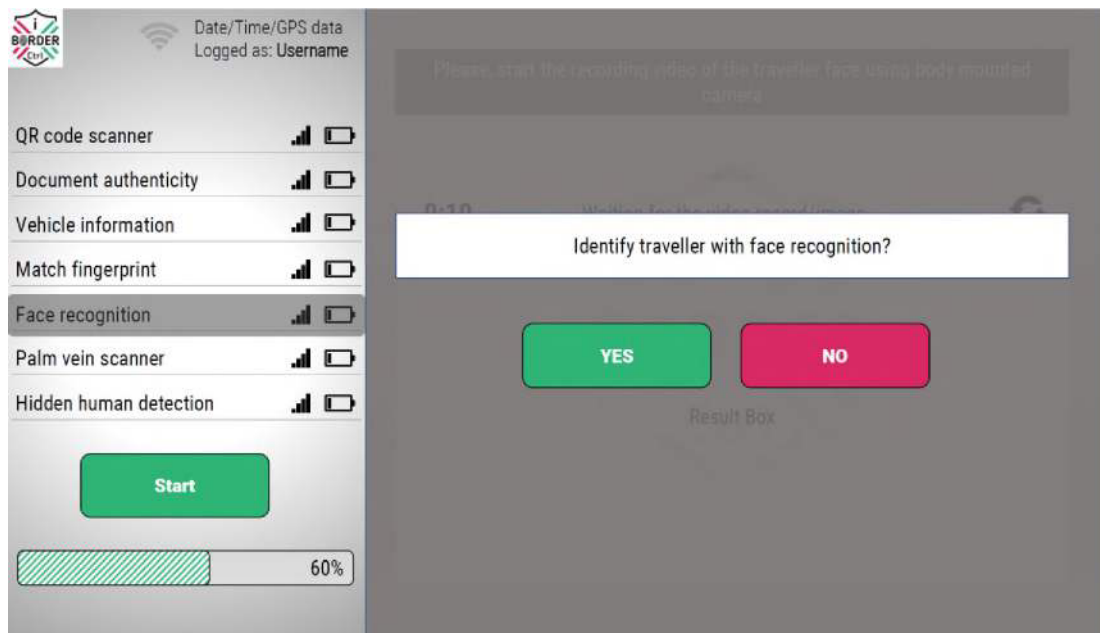
*Figure 49 Document authenticity check screen of the PU Border Guard Agent User Interface*



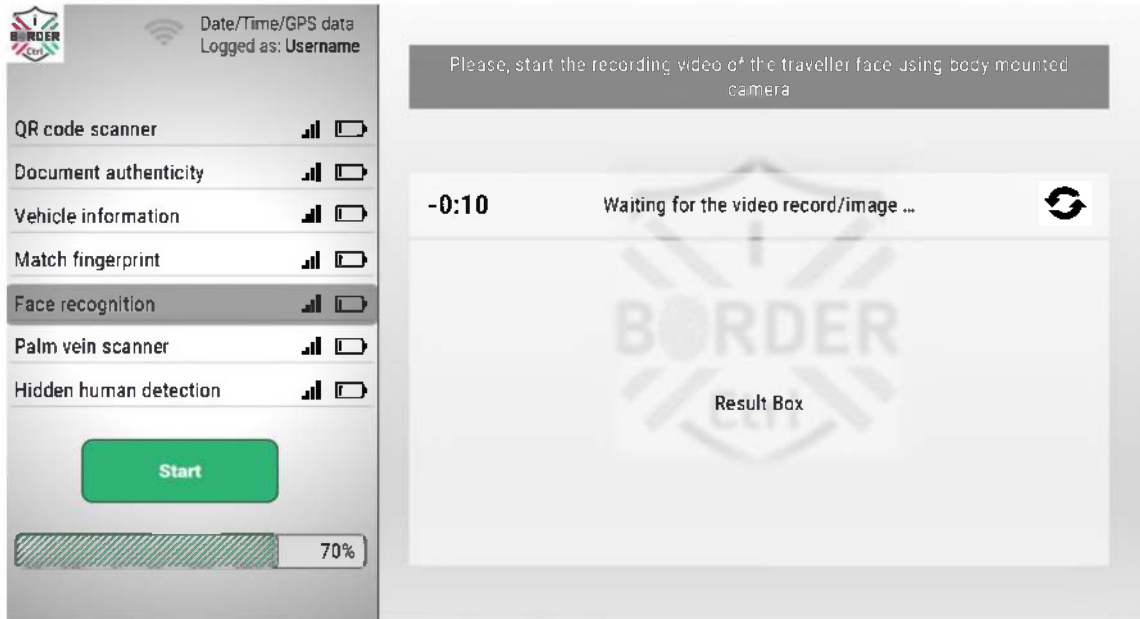
*Figure 50 Vehicle Information check screen of the PU Border Guard Agent User Interface*



**Figure 51 Match fingerprint check screen of the PU Border Guard Agent User Interface**



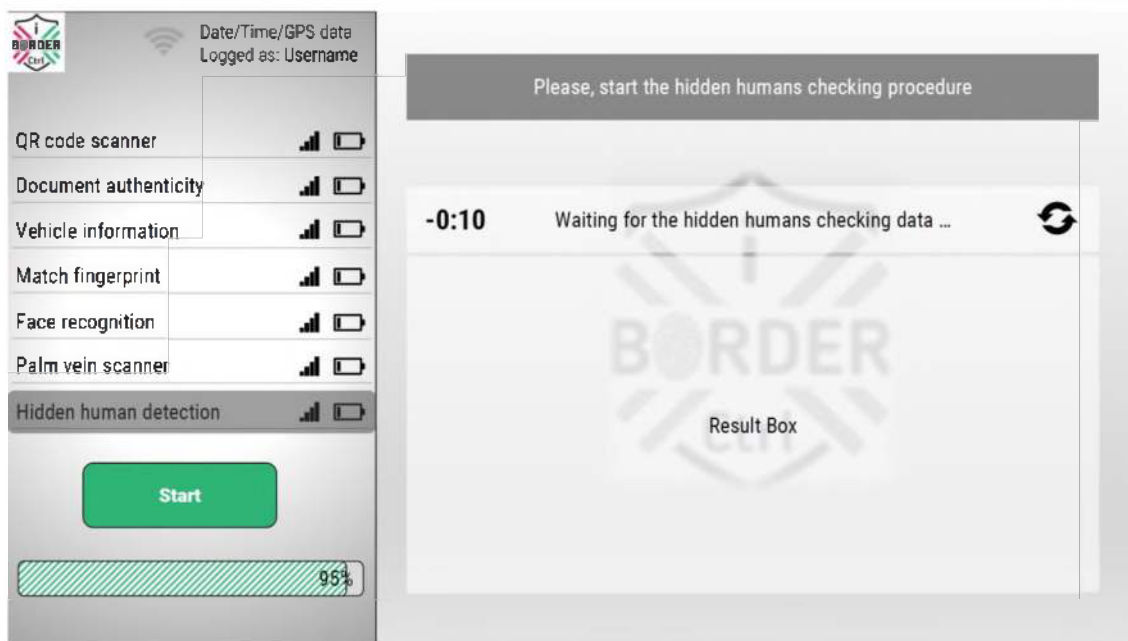
**Figure 52 Face recognition initial screen of the PU Border Guard Agent User Interface**



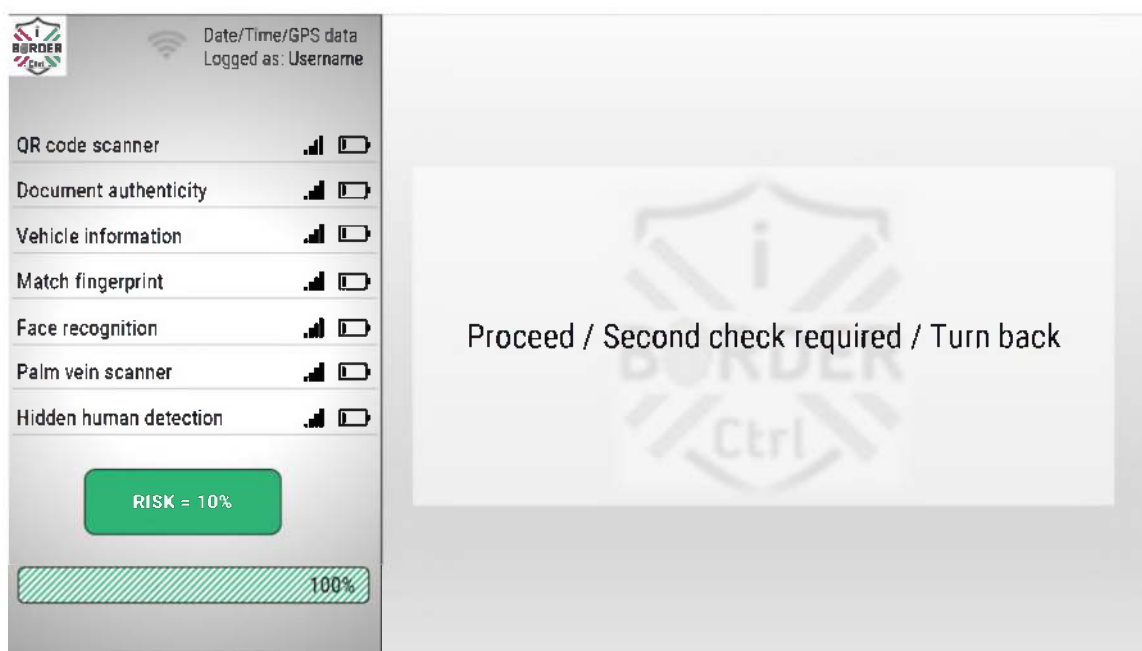
**Figure 53 Face recognition check screen of the PU Border Guard Agent User Interface**



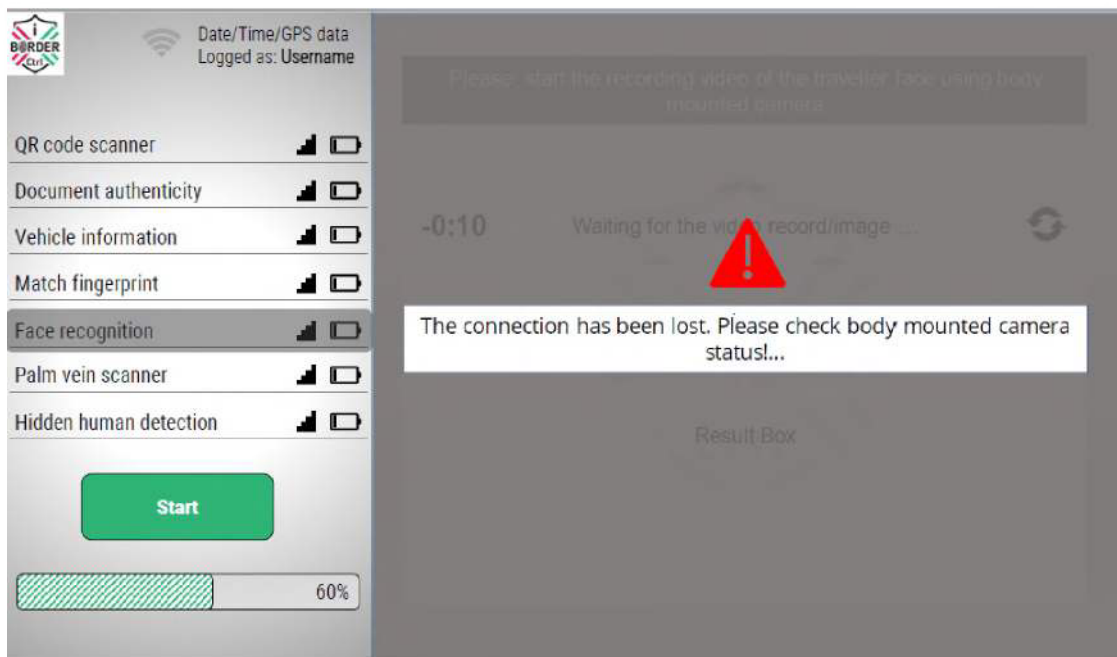
**Figure 54 Palm vein check screen of the PU Border Guard Agent User Interface**



**Figure 55** Hidden human detection check screen of the PU Border Guard Agent User Interface



**Figure 56** Overall BCP screen of the PU Border Guard Agent User Interface



**Figure 57** Lost connection between body mounted camera and tablet error screen of the PU Border Guard Agent User Interface

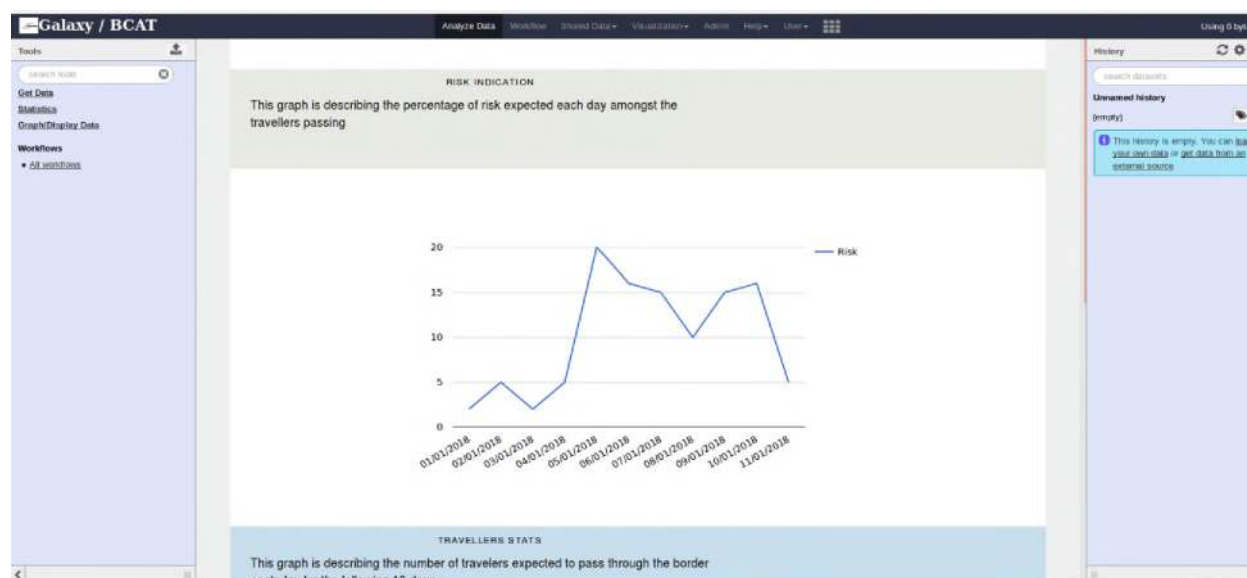
### 6.3 BMUA implementation

#### 6.3.1 BMUA interface implementation

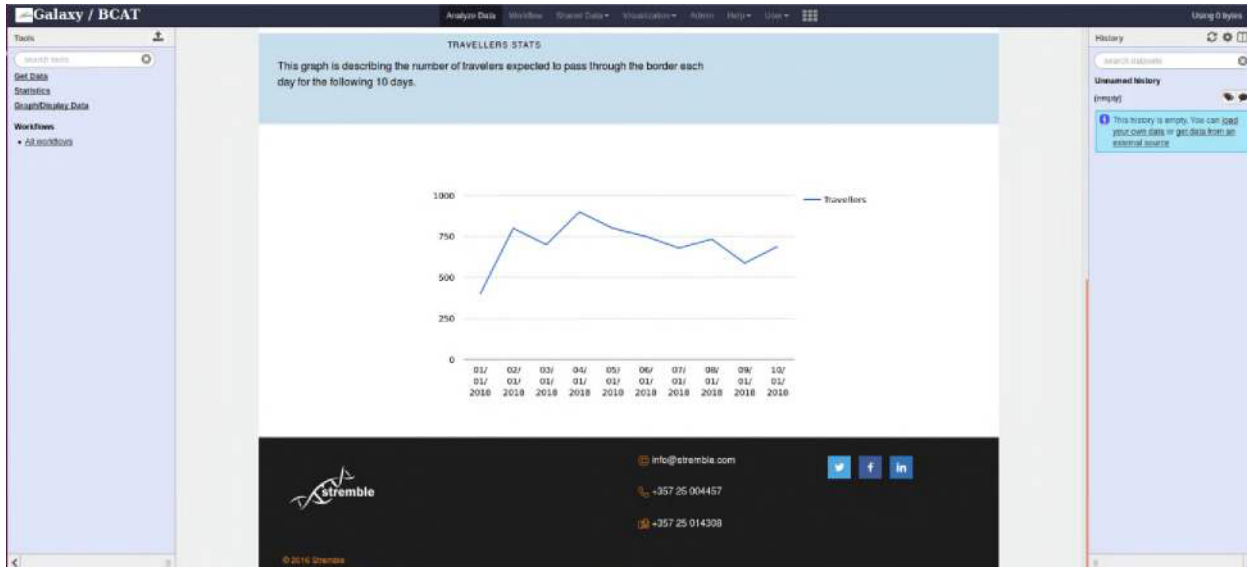
iBorderCtrl Border Manager User Application is implemented in a scientific workflow system called Galaxy. Galaxy is also a data integration platform and in our case will be receiving the information needs to be analysed. Galaxy also provides you with the ability to use or create your own statistical analysis tools and use a friendly interface to represent those results. By using python as the programming language to make the changes needed Galaxy will be the web server that will host iBorderCtrl border manager user application .

In the paragraph below you can view some screenshots of the application current progress.

#### 6.3.2 BMUA screenshots



**Figure 58** Expected risk dashboard graph for next days for border managers



**Figure 59 Expected Number of Travelers dashboard graph for border managers**

## 7 Data Protection Impact Assessment

### 7.1 Introduction

Regulation 2016/679<sup>14</sup> (GDPR) will apply from 25 May 2018, and its Article 35 introduces the concept of a Data Protection Impact Assessment (DPIA), which will be required of any data controller whose data processing is likely to result in a high risk to the rights and freedoms of the data subjects. A similar obligation is also contained in Directive 2016/680<sup>15</sup>.

A DPIA is a process designed to describe the processing, assess the necessity and proportionality of a processing and to help manage the risks to the rights and freedoms of natural persons resulting from the processing of personal data (by assessing them and determining the measures to address them). DPIAs are important tools for accountability, as they help controllers not only to comply with requirements of the GDPR, but also to demonstrate that appropriate measures have been taken to ensure compliance with the Regulation. In other words, a DPIA is a process for building and demonstrating compliance.<sup>16</sup> Non-compliance with DPIA requirements can lead to fines imposed by the competent supervisory authority. The term “Privacy Impact Assessment” (PIA) is often interchangeably and refer to the same concept. The DPIA should start as early as practicable during the design of the processing operation even if some of the processing operations are still unknown. This is also valuable for operationalising data protection by design as the results of the DPIA will point to areas that need privacy consideration and implementation. As the DPIA is updated throughout the lifecycle project, it will ensure that data protection and privacy are considered and promote the creation of solutions which promote compliance. It can also be necessary to repeat individual steps of the assessment as the development process progresses because the selection of certain technical or organizational measures may affect the severity or likelihood of the risks posed by the processing.

---

<sup>14</sup> Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).

<sup>15</sup> See Art. 27 of Directive 2016/680.

<sup>16</sup> Article 29 Working Party Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is “likely to result in a high risk” for the purposes of Regulation 2016/679, adopted on 4 October 2017

Chapter one: Screening questions

These questions are intended to help the decision whether a DPIA is necessary. Answering 'yes' to any of these questions is an indication that a DPIA would be a useful exercise. You can expand on your answers as the project develops if you need to. Please answer the following questions with a Yes (Y) or a No (N):

<b>Questions</b>	
1. Will the component involve the collection of new information about individuals?	
2. Will the component compel individuals to provide information about themselves?	
3. Will information about individuals be disclosed to components which have not previously had routine access to the information?	
4. Does the component use information about individuals for a purpose it is not currently used for, or in a way it is not currently used?	
5. Is the component using new technology that might be perceived as	



being privacy intrusive?  
For example, the use of biometrics or facial recognition.

6. Will the component result in making decisions or taking action against individuals in ways that can have a significant impact on them?

7. Is the information about individuals of a kind particularly likely to raise privacy concerns or expectations? For example, health records, criminal records or other information that people would consider to be private.

8. Will the component require to contact individuals in ways that they may find intrusive?

**Notes:**

- 1) If the answer to all the above questions is No for a specific component (e.g. HHID), then this component will not be further analysed in subsequent chapters
- 2) If the answer to any of the above questions for a component is Yes, then this component should be further analysed in subsequent chapters

Chapter two: Data protection impact assessment template

After the screening questions have identified the need for a DPIA, you can start to fill in details.

**Step one: Identify the need for a DPIA**

Explain what the module aims to achieve, what the benefits will be to the organisation, to individuals and to other parties.

Also, summarise why the need for a DPIA was identified (this can draw on your answers to the above screening questions).

**DAAT:** DAAT will be used both at the pre-registration and the border control procedures. DAAT will be responsible for the verification procedure of travel documents which the traveller provided himself during the pre-registration procedure and the border guards scanned during the border control check. The security features of travel documents (passport, visa) will be examined by DAAT against fraud characteristics. Thus, a matching score concerning the validation of the documents authenticity will be derived to facilitate the Border Guards to identify attempts of falsified document characteristics.

The DPIA is needed for DAAT because the component requires from the traveller to upload travel documents which contain Personally Identifiable Information (PII) about them, such as biometrics, photos or other personal information, thus raising privacy concerns.

<p><b>BIO (fingerprints):</b> The purpose of this module is to compare two sets of fingerprints in order to provide a match/no match response and will be used only during the border control procedures. The module will receive only the images of two fingerprints (one captured by the sensor in the portable unit with the collaboration of the traveller and the other retrieved by the system from the traveller documents) in order to compare them. No Personally Identifiable Information (PII) is used in the module and no information is stored in it. The result of the matching process is used by the border guard in order to decide if the traveller can cross or not the border and is also stored in the iBorderCtrl system in order to be used by the RBAT system.</p>
<p><b>BIO (palm vein):</b> It will be used for the traveller identification in the border control procedures. BIO will be responsible for the enrolment, verification procedure for palm vein recognition. The DPIA is needed for BIO, since a non-intrusive biometric authentication will be used for the identification of the traveller.</p>
<p><b>FMT:</b> This module will be used at both the pre-registration and the border control procedures with two different objectives. During the pre-registration, if is the first time the traveller uses the system it will create a biometric model using images captured during the Avatar interview. The FMT database will store this biometric model associated only with the traveller ID in the iBorderCtrl system. No other PII will be stored in the system. For the next visits of the traveller, the FMT will match the images of the new Avatar interview with the biometric model stored to assess the risk of the person doing the interview not being the registered traveller. No additional PII will be required or stored in the system when doing this validation.</p> <p>During the border control procedures, the FMT will match the images from the traveller (taken with the camera in the portable Unit) against the biometric reference stored during the pre-registration in order to assess the risk of not being the same person. It will also match those images with the passport image with the same purpose. Besides the HD image in the e-passport, no other information will be retrieved for the FMT module. Also during this procedure no new information will be stored in the system, only the risk assessments done will be stored for the RBAT system to be analysed.</p>
<p><b>ELSI:</b> ELSI will be used to correlate information provided by the traveller with either publically available information on the internet or with legacy border control consults systems such as the VIS, SIS. A DPIA is deemed mandatory as it processes traveller sensitive personal data such as visa card number, photos and other personal information.</p>
<p><b>TUA:</b> The traveller user application is the module responsible for managing the pre-registration procedure. The application backend will store the information provided by the user to the iBorderCtrl database (personal and travel information), and will provide simplified access to the stored data. The traveller user</p>

application will also include the traveller user interface, which is the presentation layer (visualisation) of the application

The DPIA is needed because the component requires from the traveller to enter personal and private data including travel documents that also contain personal, private and sensitive information. This information is also available to other components/modules. As the collected traveller information is going to be used for purposes not currently used in the standard border crossing procedures such as: deception detection, face matching, automatic travel document authentication check, analytics and statistics, cross-check with external legacy systems and risk calculation, there is an increased need for DPIA.

**BGUA:** The BG user application is the module responsible for managing the border control procedures and interaction with the BG officer. The application backend will be collecting (indirectly) all information provided by the traveller during the border control procedure. The module itself does not make value judgments about the traveller. All collected data will be sent to the local server database at the border crossing point. From there, modules such FMT or DAAT will be fetching traveller's data calculating the overall risk for the traveller and resending it to the PU. The BGUA is the tool that will help border guard officer to undertake critical decision whether a traveller will pass the border or not.

The DPIA is needed for BGUA because the component requires from the traveller to upload information such as traveller document data and biometric data that both contain Personally Identifiable Information (PII) about travellers such as different biometric patterns or address, place/date of birth.

**Portable Unit:** The PU is directly responsible for collecting all the information provided by the traveller during the border control procedure, such as travel documents and biometric patterns. The PU contains different sensors and devices that will be taking part in data collection procedures. The PU supplies the iBorderCtrl platform with the data that could be processed to calculate the overall risk for the traveller passing through the border. For the organisations, the benefits are the enhanced level of security in terms of crime detection.

The DPIA is needed for BGUA because the component requires from the traveller to upload the confidential information, such as the traveller's personal (from the documents) and biometric data that both contain Personally Identifiable Information (PII) about travellers, such as different biometric patterns or address, place/date of birth.

#### **Justification why DPIA is NOT needed for the HDD tool**

The collected data which are processed with the HDD tool sensors (radar and acoustic ones) are directly connected to the received electrical signal strength (received power); which are then processed and combined through advanced signal processing to provide the detection event. Thus, the whole detection and processing are directly connected to the electrical properties and parameters of the vehicles and subjects and not with recording or collecting human related data

(i.e. voice). Thus, it is hereby confirmed that neither voice recognition algorithms nor any kind of personal data collection for human recognition systems nor other related software, for matching human voice or motion with human identification are implemented; neither the nature nor the scope nor the addressed targets of the specific detection techniques require or dictate such kind of human related processing. Thus, the hardware and software techniques related to the above sensors are based on aggregate data analysis of the nature described above and are not at all connected with isolating or identifying human individuals and are not used to collect or process relevant data related to humans and thus do not raise any relevant data protection issues and consequently do not affect any kind of human privacy and personal data.

### **Step two: Describe the information flows**

Describe here the collection, use and deletion of personal data. It may also be useful to refer to a flow diagram or another way of explaining data flows. You should also say how many individuals are likely to be affected by the project.

(Describe here the following:

1. How will the information be collected and transferred to the project/module?
2. Who will have access the information?
3. Where will the information be held?
4. What will the information be used for?
5. How long will the information be retained? How will it be destroyed/deleted?
6. Who will be the owner of the information?
7. Will the information be shared with anyone? If yes, who?

If possible also add an information flow diagram.)



## D4.1 First version of the iBorderCtrl software platform



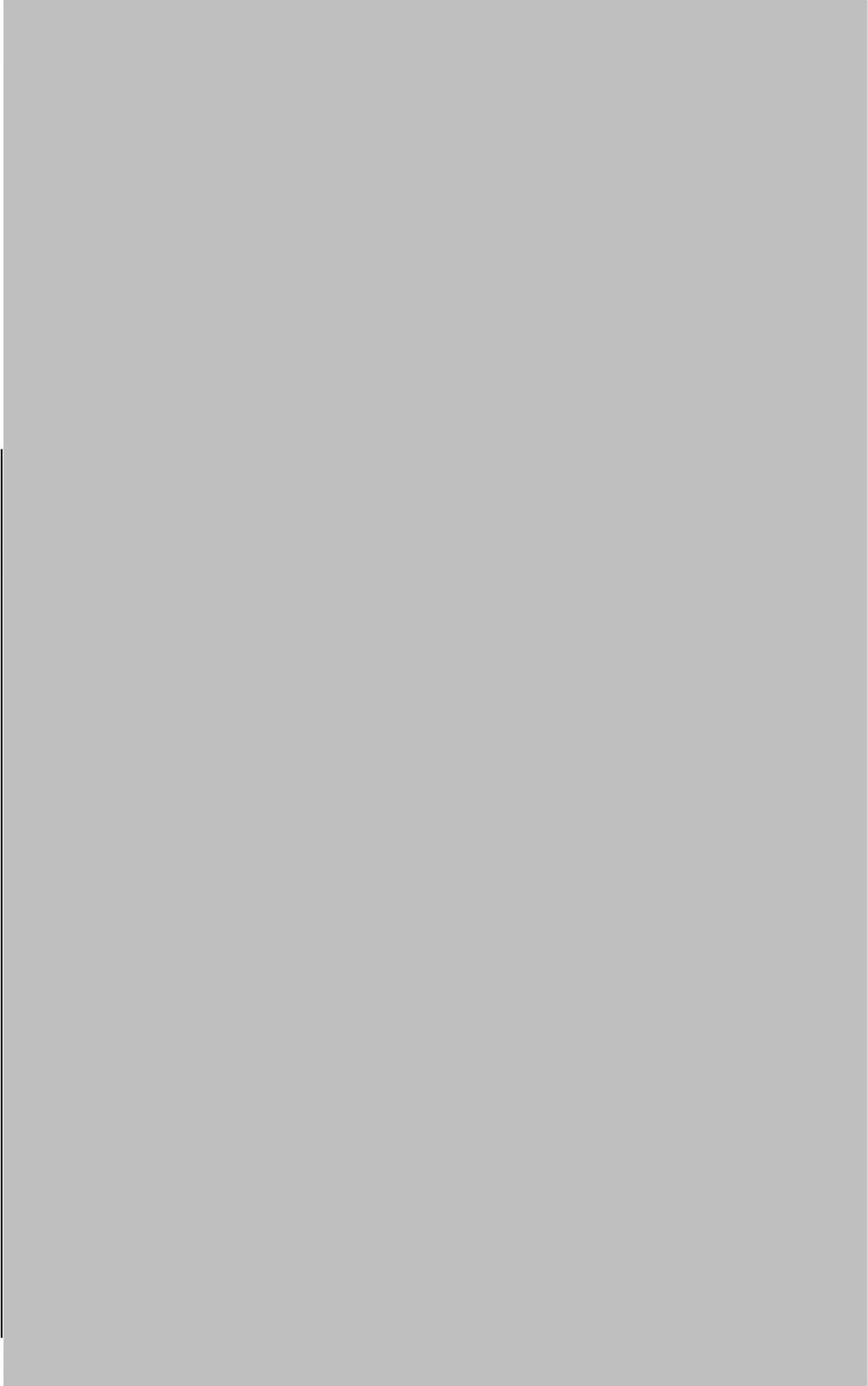


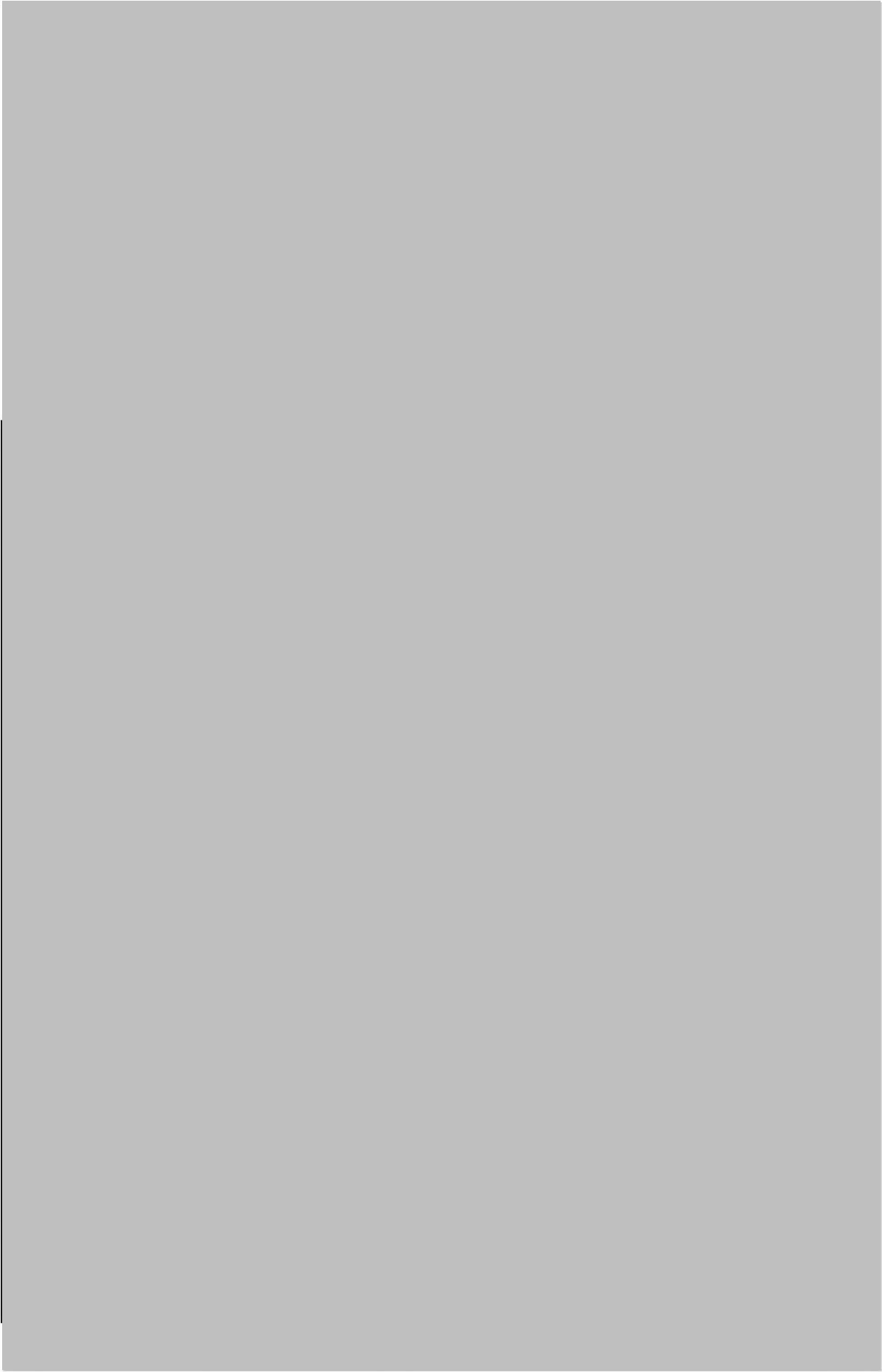
#### D4.1 First version of the iBorderCtrl software platform











### Consultation requirements

Explain what practical steps you will take to ensure that you identify and address privacy risks. Who should be consulted internally and externally about the privacy concerns/risks? How was the consultation carried out? How will any future consultations be carried out? You should link this to the relevant stages of your project management process. What were the outcomes of the consultation? What privacy concerns were raised internally/externally? How will you report on any future consultations?

You can use consultation at any stage of the DPIA process.

**DAAT:** Together with the database design team we have identified which personal/private information should be stored for supporting the module functions that also present privacy risks. Storage of this information was then discussed with the Information Security Officer. Internally data protection and security issues are

handled by the information security officer, complying with the policies and procedures as set by the ISO 27001:2015. Externally, the Project Co-ordination team and LUH will be consulted.

Future internal consultations will be carried through internal technical project meetings and together with the information security officer. However, due to the personal data that DAAT receives from the iBorderCtrl database and the document scanner at the border crossing point, this will result in a new DPIA document. Regarding external consultation, regulatory authorities, advocacy organisations, industry experts and academics can also be consulted.

As a result of this consultation process we have identified the privacy risks, allowing us to implement appropriate technical and organisational measures, ensuring that these risks are reduced.

**BIO (fingerprints):** Currently the only information stored in the iBorderCtrl database is the result of the fingerprint matching process. Hence no private neither sensitive personal information about the traveller is stored by the module.

**BIO (palm vein):** The palm vein system links the iBorderCtrl user\_ID to the palm vein template. Therefore the palm vein database will not contain any PII about travellers. The palm vein system has been audited by three national privacy authorities and confirmed that the biometric template is private information, not sensitive information. As a result of the audits we have identified the privacy risks, based on which IT measures have been taken to ensure reduce of this risk.

**FMT:** The security and integrity of the database in which the biometric record of the traveller is stored is handled by everis ADS information security officer and everis ADS cyber security team. Externally, the Project Co-ordination team and LUH will be consulted.

**ELSI:** The social media username is provided in from the traveller at the pre-registration phase. As such we expect from the legal team to ensure that inform consent is provided by the user who would knowingly provide access to this information being fully aware of what data and how it will be processed.

**TUA:** Together with the database design team we have identified which personal/private information should be stored for supporting the module functions that also present privacy risks. Storage of this information was then discussed with the Information Security Officer. Internally data protection and security issues are handled by the information security officer, complying with the policies and procedures as set by the ISO 27001:2015. Externally, the Project Co-ordination team and LUH will be consulted.

Future internal consultations will be carried through internal technical project meetings and together with the information security officer. Since TUA will handle different datasets coming from different sources, most of which are prone to privacy and security issues, this will result in a new DPIA document. Regarding external consultation, regulatory authorities, advocacy organisations, industry experts and academics can also be consulted.

As a result of this consultation process we have identified the privacy risks, allowing us to implement appropriate technical and organisational measures, ensuring that these risks are reduced.

**MGUA and Portable Unit:** JAS team will investigate secure wireless data transmission mechanisms, in order to protect from packet sniffing/capturing. Internal consultations will be carried through technical project meetings. Regarding external consultation the Polish border guard which cooperates with JAS, can also be consulted.

**Step three: Identify the privacy and related risks**

Identify the key privacy risks and the associated compliance and corporate risks. Chapter three can be used to help you identify the DPA related compliance risks.

Privacy issue <sup>17</sup>	Risk to individuals <sup>18</sup> (complete if appropriate to issue or put not applicable)	Compliance risk <sup>19</sup> (complete if appropriate to issue or put not applicable)	Associated organisation / corporate risk <sup>20</sup> (complete if appropriate to issue or put not applicable)

<sup>17</sup> what gives rise to the risk

<sup>18</sup> for example, misuse or overuse of their personal data, loss of anonymity, intrusion into private life through monitoring activities, lack of transparency)

<sup>19</sup> breach of the Data Protection Act (DPA)

<sup>20</sup> for example, failure of the project and associated costs, legal penalties or claims, damage to reputation, loss of trust of travellers or the public

<b>DAAT</b>	<b>DAAT-01</b> Personal data not securely stored	Individuals personal data are not protected against unauthorised or unlawful processing, accidental loss, destruction or damage	Non-compliance with: <ul style="list-style-type: none"><li>GDPR Art. 32 "Security of processing"</li></ul>
			<ol style="list-style-type: none"><li>1. Legal penalties or claims</li><li>2. May lead to public mistrust</li></ol>



					<p>3. May lead to sanction by the supervisory authority</p> <p>4. Unforeseen associated costs, for implementing appropriate technical and organisational measures to ensure a level of security.</p>
	<p><b>DAAT-02</b> Personal data kept longer than necessary for the purpose obtained</p>	<p>Increased risk that personal data are outdated, inaccurate, not securely stored</p>	<p>Non-compliance with:</p> <ul style="list-style-type: none"> <li>GDPR Art. 5 (1)(e) - "Principles relating to processing of personal data"</li> </ul>	<p>1. Legal penalties or claims</p> <p>2. May lead to public mistrust</p>	
	<p><b>DAAT-03</b> Collection of more information than needed for the purposes specified</p>	<p>Individuals share more personal data than needed</p>	<p>Non-compliance with:</p> <ul style="list-style-type: none"> <li>GDPR Art. 5 (1)(a) - "Principles relating to processing of personal data"</li> </ul>	<p>1. May lead to sanction by the supervisory authority</p> <p>2. May lead to public mistrust</p>	
<p><b>BIO (palm vein)</b></p>	<p><b>BIO-01</b> Information (result of the authentication, result of enrolment, biometric templates and event logs) not securely stored</p>	<p>Data created and stored by the module is not protected against unauthorised or unlawful processing, accidental loss, destruction or damage</p>	<p>Non-compliance with:</p> <ul style="list-style-type: none"> <li>GDPR Art. 32 "Security of processing"</li> </ul>	<p>1. Legal penalties or claims</p> <p>2. May lead to public mistrust</p> <p>3. May lead to sanction by the supervisory authority</p> <p>4. Unforeseen associated costs, for implementing appropriate technical and organisational measures to ensure a level of security.</p>	

FMT	FMT-01 Personal data not securely stored	Individuals personal data are not protected against unauthorised or unlawful processing, accidental loss, destruction or damage	Non-compliance with: <ul style="list-style-type: none"> <li>GDPR Art. 32 "Security of processing"</li> </ul>	<ol style="list-style-type: none"> <li>1. Legal penalties or claims</li> <li>2. May lead to public mistrust</li> <li>3. May lead to sanction by the supervisory authority</li> <li>4. Unforeseen associated costs, for implementing appropriate technical and organisational measures to ensure a level of security.</li> </ol>
	FMT-02 Personal data kept longer than necessary for the purpose obtained	Increased risk that personal data are outdated, inaccurate, not securely stored	Non-compliance with: <ul style="list-style-type: none"> <li>GDPR Art. 5 (1)(e) – "Principles relating to processing of personal data"</li> </ul>	<ol style="list-style-type: none"> <li>1. Legal penalties or claims</li> <li>2. May lead to public mistrust</li> </ol>
ELSI	ELSI-01 Collection of more information than needed for the purposes specified	Individuals share more personal data than needed	Non-compliance with: <ul style="list-style-type: none"> <li>GDPR Art. 5 (1)(a) – "Principles relating to processing of personal data"</li> </ul>	<ol style="list-style-type: none"> <li>1. May lead to sanction by the supervisory authority</li> <li>2. May lead to public mistrust</li> </ol>
	ELSI-02 Not open about the reasons for obtaining personal data and what we intend to do with it	Function creep	Non-compliance with: <ul style="list-style-type: none"> <li>GDPR Art. 5 (1)(a) – "Principles relating to processing of personal data"</li> <li>GDPR Art. 5 (1)(b) – "Principles relating to processing of personal data"</li> </ul>	<ol style="list-style-type: none"> <li>1. Legal penalties or claims</li> <li>2. May lead to public mistrust</li> <li>3. May lead to sanction by the supervisory authority</li> </ol>

<p><b>TUA</b></p>	<p><b>TUA-01</b> Individuals cannot access/read the consent form/ privacy notice</p>	<p>Individuals are not aware that their data is being collected and processed</p>	<p>Non-compliance with:</p> <ul style="list-style-type: none"> <li>GDPR Art. 5 – "Principles relating to processing of personal data"</li> <li>GDPR Art. 15 - "Right of access"</li> <li>GDPR Art. 21 "Right to object"</li> </ul>	<p>1. Legal penalties or claims 2. May lead to public mistrust 3. May lead to sanction by the supervisory authority</p>
<p><b>TUA-02</b> Individuals cannot access/read the consent form/ privacy notice</p>	<p>Individuals cannot withdraw their consent</p>	<p>Non-compliance with:</p> <ul style="list-style-type: none"> <li>GDPR Art. 17 "Right to erasure", also known as 'the right to be forgotten'</li> </ul>	<p>1. Legal penalties or claims 2. May lead to public mistrust</p>	
<p><b>TUA-03</b> Individuals cannot withdraw their consent, since it is not supported by the traveller application</p>	<p>Individuals cannot withdraw their consent</p>	<p>Non-compliance with:</p> <ul style="list-style-type: none"> <li>GDPR Art. 17 "Right to erasure", also known as 'the right to be forgotten'</li> </ul>	<p>1. Legal penalties or claims 2. May lead to public mistrust</p>	
<p><b>TUA-04</b> Individuals cannot access their personal data and other supplementary information</p>	<p>Individuals are not aware neither can verify the lawfulness of the processing</p>	<p>Non-compliance with:</p> <ul style="list-style-type: none"> <li>GDPR Art. 15 "Right of access"</li> </ul>	<p>1. Legal penalties or claims 2. Unforeseen associated costs, for dealing with a subject access request</p>	
<p><b>TUA-05</b> Individuals cannot update their personal data</p>	<p>Individuals personal data are inaccurate or incomplete</p>	<p>Non-compliance with:</p> <ul style="list-style-type: none"> <li>GDPR Art. 16 "Right to rectification"</li> <li>GDPR Art. 5 (1)(d) – "Accuracy"</li> </ul>	<p>1. Legal penalties or claims 2. May lead to public mistrust 3. Unforeseen communication costs, in case personal data are disclosed in third parties, since third parties should be informed while individuals should also be informed about the third</p>	

					parties to whom the data has been disclosed.
<b>TUA-06</b> Individuals cannot restrict personal data processing	Individuals cannot block or suppress the processing their personal data	Individuals cannot obtain or reuse their personal data for their own purposes across different services;	Individuals cannot block or suppress the processing their personal data	Not applicable since personal data are required to perform the task at hand	Not applicable
<b>TUA-07</b> Individuals cannot obtain their personal data in a structured, commonly used and machine-readable form	Individuals cannot: <ul style="list-style-type: none"> <li>obtain and reuse their personal data for their own purposes across different services;</li> <li>move, copy or transfer their personal data easily from one IT environment to another;</li> </ul>	Individuals cannot obtain or reuse their personal data for their own purposes across different services;	Individuals cannot obtain or reuse their personal data for their own purposes across different services;	Non-compliance with: <ul style="list-style-type: none"> <li>GDPR Art. 20 "Right to data portability"</li> </ul>	1. Legal penalties or claims 2. Unforeseen associated costs, for implementing personal data export mechanisms
<b>TUA-08</b> Individuals cannot object to processing for scientific/historical research and statistics purposes	Individuals cannot object on "grounds relating to his or her particular situation"	Individuals cannot object on "grounds relating to his or her particular situation"	Individuals cannot object on "grounds relating to his or her particular situation"	Non-compliance with: <ul style="list-style-type: none"> <li>GDPR Art. 21 "Right to object"</li> </ul>	1. Legal penalties or claims 2. May lead to public mistrust
<b>TUA-09</b> Personal data not securely stored	Individuals personal data are not protected against unauthorised or unlawful processing, accidental loss, destruction or damage	Individuals personal data are not protected against unauthorised or unlawful processing, accidental loss, destruction or damage	Individuals personal data are not protected against unauthorised or unlawful processing, accidental loss, destruction or damage	Non-compliance with: <ul style="list-style-type: none"> <li>GDPR Art. 32 "Security of processing"</li> </ul>	1. Legal penalties or claims 2. May lead to public mistrust 3. May lead to sanction by the supervisory authority 4. Unforeseen associated costs, for implementing appropriate technical and organisational measures to ensure a level of security.
<b>TUA-10</b> Security breach leading to the destruction, loss, alteration,	Individuals might lose more than personal data	Individuals might lose more than personal data	Individuals might lose more than personal data	Non-compliance with:	1. Legal penalties or claims

	<p>unauthorised disclosure of, or access to, personal data</p>		<ul style="list-style-type: none"> <li>GDPR Art. 33 "Notification of a personal data breach to the supervisory authority"</li> <li>GDPR Art. 34 "Communication of a personal data breach to the data subject"</li> <li>GDPR Art. 5 (1)(f) - "Principles relating to processing of personal data"</li> </ul>	<p>2. May lead to sanction by the supervisory authority</p> <p>3. May lead to public mistrust</p>
	<p><b>TUA-11</b> Personal data kept longer than necessary for the purpose obtained</p>	<p>Increased risk that personal data are outdated, inaccurate, not securely stored</p>	<p>Non-compliance with:</p> <ul style="list-style-type: none"> <li>GDPR Art. 5 (1)(e) - "Principles relating to processing of personal data"</li> </ul>	<p>1. Difficulty in responding to access requests for any personal data</p>
	<p><b>TUA-12</b> Collection of more information than needed for the purposes specified</p>	<p>Individuals share more personal data than needed</p>	<p>Non-compliance with:</p> <ul style="list-style-type: none"> <li>GDPR Art. 5 (1)(a) - "Principles relating to processing of personal data"</li> </ul>	<p>1. May lead to sanction by the supervisory authority</p> <p>2. May lead to public mistrust</p>
	<p><b>TUA-13</b> Not open about the reasons for obtaining personal data and what we intend to do with it</p>	<p>Function creep</p>	<p>Non-compliance with:</p> <ul style="list-style-type: none"> <li>GDPR Art. 5 (1)(a, b) - "Principles relating to processing of personal data"</li> </ul>	<p>1. Legal penalties or claims</p> <p>2. May lead to public mistrust</p> <p>3. May lead to sanction by the supervisory authority</p>
<p><b>BGUA and Portable Unit</b></p>	<p><b>BGUA PU - 01:</b> Traveller data not safely transmitted leading to the unauthorised disclosure of, or access to, personal data</p>	<p>Individuals might lose more than personal data</p>	<p>Non-compliance with:</p> <ul style="list-style-type: none"> <li>GDPR Art 32. "Security of Processing"</li> </ul>	<p>1. Legal penalties or claims</p> <p>2. May lead to sanction by the supervisory authority</p> <p>3. May lead to public mistrust</p>

				<p>Increased risk that personal data are outdated, inaccurate, not securely stored</p> <p>Function creep</p>	<p>Non-compliance with:</p> <ul style="list-style-type: none"> <li>GDPR Art. 5 (1)(a) – "Principles relating to processing of personal data"</li> </ul> <p>Non-compliance with:</p> <ul style="list-style-type: none"> <li>GDPR Art. 5 (1)(a, b) – "Principles relating to processing of personal data"</li> </ul>	<p>1. Legal penalties or claims</p> <p>2. May lead to public mistrust</p> <p>1. Legal penalties or claims</p> <p>2. May lead to public mistrust</p> <p>3. May lead to sanction by the supervisory authority</p>
		<p><b>BGUA PU - 02:</b> Personal data kept longer than necessary for the purpose obtained</p> <p><b>BGUA PU - 03:</b> Not open about the reasons for obtaining personal data and what we intend to do with it and where or if we intend to store it</p>				

**Step four: Identify privacy solutions**

Describe the actions you could take to reduce the risks, and any future steps which would be necessary (e.g. the production of new guidance or future security testing for systems).

Risk	Solution(s)	<p><b>Result:</b> is the risk eliminated, reduced, or accepted?</p>	<p><b>Evaluation:</b> is the final impact on individuals after implementing each solution a justified, compliant and proportionate response to the aims of the project?</p>
------	-------------	---	---



<b>DAAT-01</b>		IT infrastructure that meets relevant security, storage and data processing along with data privacy requirements	Reduced	Final impact is justified, compliant and proportionate – appropriate data protection and security measures will
----------------	--	--	---------	---

			safeguard against unauthorised or unlawful processing, accidental loss, destruction or damage of personal data.
<b>DAAT-02</b>	<p>Consortium will include in the consent form/privacy notice the data retention policy.</p> <p>Key issues that the consortium should consider when implementing a data retention policy will be identified, including: (a) location of where personal data is stored, (b) variable retention periods based on the type of data processed, the purpose of processing or other factors, (c) actions to be performed after the expiration of the applicable retention period and (d) other information obligations</p>	Eliminated	Final impact is justified, compliant and proportionate.
<b>DAAT-03</b>	Only the absolutely necessary traveller document information is going to be extracted.	Eliminated	Final impact is justified, compliant and proportionate
<b>BIO-01</b>	IT infrastructure that meets relevant security, storage and data processing along with data privacy requirements	Reduced	Final impact is justified, compliant and proportionate.
<b>FMT-01</b>	IT infrastructure that meets relevant security, storage and data processing along with data privacy requirements	Reduced	Final impact is justified, compliant and proportionate – appropriate data protection and security measures will safeguard against unauthorised or unlawful processing, accidental loss, destruction or damage of personal data.



<p><b>FMT-02</b></p>	<p>Consortium will include in the consent form/privacy notice the data retention policy.</p> <p>Key issues that the consortium should consider when implementing a data retention policy will be identified, including: (a) location of where personal data is stored, (b) variable retention periods based on the type of data processed, the purpose of processing or other factors, (c) actions to be performed after the expiration of the applicable retention period and (d) other information obligations</p>	<p>Eliminated</p>	<p>Final impact is justified, compliant and proportionate.</p>
<p><b>ELSI-01</b></p>	<p>Only the absolutely necessary traveller document information is going to be extracted.</p>	<p>Eliminated</p>	<p>Final impact is justified, compliant and proportionate.</p>
<p><b>ELSI-02</b></p>	<p>In the consent form, the reasons for obtaining the personal data and how they are going to be processed/utilised is clearly indicated. The rights of the data subjects is also be indicated.</p>	<p>Eliminated</p>	<p>Final impact is justified, compliant and proportionate</p>
<p><b>TUA-01</b></p>	<p>We will make sure that they read the consent form/privacy notice by placing it where they can easily find it (i.e. the user will not be able to check the consent box and continue unless he has scrolled down the whole consent form and not only the first pages)</p>	<p>Eliminated</p>	<p>Final impact is justified, compliant – individuals can access the consent form/privacy notice – and proportionate.</p>
<p><b>TUA-02</b></p>	<p>Consent form will explicitly state how an individuals can withdraw their consent.</p>	<p>Eliminated</p>	<p>Final impact is justified, compliant – individuals are aware how to withdraw their consent – and proportionate.</p>

<p><b>TUA-03</b></p>	<p>We will make sure that the users are able to revoke their consent at any time. Additionally all data stored for the specific individual will be removed from the iBorderCtrl database, including databases used by/from other services/components.</p>	<p>Eliminated</p>	<p>Final impact is justified, compliant – individuals can withdraw their consent – and proportionate.</p>
<p><b>TUA-04</b></p>	<p>The traveller will have access to their personal data through the Traveller User Application and will be able to modify them. Other supplementary information will also be available to the travellers.</p>	<p>Eliminated</p>	<p>Final impact is justified, compliant – individuals can access their personal data – and proportionate.</p>
<p><b>TUA-05</b></p>	<p>The traveller will have access to their personal data through the Traveller User Application and will be able to update them.</p>	<p>Eliminated</p>	<p>Final impact is justified, compliant – individuals can update their personal data – and proportionate.</p>
<p><b>TUA-07</b></p>	<p>All information stored in the iBorderCtrl database will be extracted in a structured, commonly used and machine readable form (i.e. .csv)</p>	<p>Reduced</p>	<p>Final impact is justified, compliant and proportionate – individuals can obtain their personal data in csv format.</p>
<p><b>TUA-08</b></p>	<p>The GDPR ensures that all data subjects have rights in relation to the processing of their personal data.  Member States can derogate from these obligations where data is processed for research/statistical/archiving purposes.  Consent form will explicitly state that personal data are collected for research purposes.</p>	<p>Eliminated</p>	<p>Final impact is justified, compliant and proportionate, since:</p> <ul style="list-style-type: none"> <li>• Data isn't processed to support measures or decisions with respect to particular individuals</li> <li>• Data is not processed in a way that would be likely</li> </ul>

	<p>We must have a clear understanding of potential security benefits and risks associated with cloud computing, and set realistic expectations from our CSPs. Attention must be given to the different service models (IaaS, PaaS or SaaS) as each model brings different security requirements and responsibilities.</p> <p>Scalable and robust cloud-based solution that will adequately meet the relevant security, storage and data processing along with data privacy requirements</p>	Reduced	to cause substantial damage or distress to individuals.
TUA-09	<p>Final impact is justified, compliant and proportionate – appropriate data protection and security measures will safeguard against unauthorised or unlawful processing, accidental loss, destruction or damage of personal data.</p>		
TUA-10	<p>According to regulation 5A of the Privacy and Electronic Communications (EC Directive) Regulations 2003 (PECR) the selected cloud provider should inform the consortium about the data breach and provide details of the breach log.</p> <p>A Clear SLA with cloud service provider should help.</p>	Reduced	Final impact is justified, compliant and proportionate – as long as the consortium is notified about data breaches we can notify individuals.
TUA-11	<p>Consortium will include in the consent form/privacy notice the data retention policy.</p> <p>Key issues that the consortium should consider when implementing a data retention policy will be identified, including: (a) location of where personal data is stored, (b) variable retention periods based on the type of data processed, the purpose of processing or other factors, (c) actions to be performed after the expiration of the applicable retention period and (d) other information obligations</p>	Eliminated	Final impact is justified, compliant and proportionate.
TUA-12	<p>Only the absolutely necessary traveller information which are going to be used by other components or the border guard will be stored to the iBorderCtrl database.</p>	Eliminated	Final impact is justified, compliant and proportionate
TUA-13	<p>In the consent form, the reasons for obtaining the personal data and how they are going to be processed/ utilised is thoroughly analysed.</p>	Eliminated	Final impact is justified, compliant and proportionate

<b>BGUA PU - 01</b>	End-to-end encryption will be done on the client side: no data will be sent to the server unencrypted; encryption keys stay at the user's side and never reach local BCP stations servers. Using industry-standard cryptography algorithms: AES-256, RSA with 4096 bit keys (or longer).	Eliminated	Final impact is justified, compliant and proportionate
<b>BGUA PU - 02</b>	The traveller data will be processed locally during the border control procedure, and erased as soon as these data are forwarded to relevant modules (DAAT, FMT, BIO).	Eliminated	Final impact is justified, compliant and proportionate
<b>BGUA PU - 03</b>	In the consent form, the reasons for obtaining the personal data and how they are going to be processed/utilised is thoroughly analysed.	Eliminated	Final impact is justified, compliant and proportionate

**Step five: Sign off and record the DPIA outcomes**

Who has approved the privacy risks involved in the project? What solutions need to be implemented?

Risk	Approved solution	Approved by (Project Manager/Main contact)
------	-------------------	---



<b>DAAT-01</b>	<p>Include redundancy and diversity within the network, to safeguard against “accidental loss or destruction of, or damage to, personal data”.</p> <p>Protection of object storage using LUKS, the standard for Linux hard disk encryption.</p> <p>Appropriate protection policies, describing how, when and where personal data will be processed. Protection policies should be consistent throughout the lifecycle<sup>21</sup> of the data and should only change when explicitly and authoritatively requested. If the sensitivity of the data changes during the lifecycle, mechanisms should allow for the protection level to adapt accordingly, without any gaps in the protection enforcement during these policy changes. Ideally the data should also be used in its encrypted form, using techniques such as ‘homomorphic encryption’.</p> <p>We should avoid:</p> <ul style="list-style-type: none"><li>• Web server with directory browsing enabled, since the entire contents can be indexed within minutes and therefore its contents made available.</li><li>• Using a unique URL parameter as part of a dynamic website without any additional access controls (such as log in), since editing the URL manually can make personal data accessible to others.</li><li>• Easily-guessable directory names that don’t have specific access restrictions.</li></ul> <p>Instead we should:</p> <ul style="list-style-type: none"><li>• Set up an isolated network or use internal firewall policies, if applications need to process particularly sensitive information.</li><li>• Ensure web servers are exposing only the intended content.</li></ul>	Solution approved by the Information Security Officer and the development team manager of ED, LUH and the Project Coordination Team
----------------	---	---

<sup>21</sup> Includes creation, modification, aggregation, distribution, archive, and destruction

	<ul style="list-style-type: none"> <li>• Provide awareness of personal data and where and how it should be stored. This can be achieved with good training. To be effective training should be tailored to different roles in the organization.</li> </ul>	
<p><b>DAAT-02</b></p>	<p>Organisational measures include data retention policy within the consent form/privacy notice, taking into consideration: (a) location of where personal data is stored, (b) variable retention periods based on the type of data processed, the purpose of processing or other factors, (c) actions to be performed after the expiration of the applicable retention period and (d) other information obligations.</p> <p>Technical measures include:</p> <ul style="list-style-type: none"> <li>• review how long we keep personal data;</li> <li>• consider the purpose or purposes we hold the information in order to decide whether (and for how long) to retain it;</li> <li>• securely delete information that is no longer needed for this purpose or these purposes; and</li> <li>• update, archive or securely delete information if it goes out of date</li> </ul> <p>To avoid the problems arising from keeping personal data for too long it is good practice to regularly review personal data and delete anything no longer needed. Information that does not need to be accessed regularly, but which still needs to be retained, should be safely archived or put offline.</p> <p>If the amount of personal data we hold is:</p> <ul style="list-style-type: none"> <li>• more than small, it is good practice to establish standard retention periods for different categories of information; a system for ensuring that we follow these retention periods in practice, and for documenting and reviewing the retention policy is also advisable;</li> <li>• modest, then we might not need a formal data retention policy; However, in order to comply with the protection principles we should conduct a regular audit and check through the records we hold to make sure that we are not holding onto personal data for too long, or deleting them too early.</li> </ul> <p>According to<sup>22</sup> appropriate personal data retention period is also likely to depend on:</p> <ul style="list-style-type: none"> <li>• The purpose for which it was obtained and its nature. Since personal data should be kept for research purposes, then we may keep data indefinitely as long as it is not used in connection with decisions affecting particular individuals, or in a way that is likely to cause damage or distress. Data should be immediately by removed when it is no longer needed for these purposes.</li> </ul> <p>On the other hand, information with only a short-term value may have to be deleted within days.</p> <ul style="list-style-type: none"> <li>• The surrounding circumstances. If a traveller stops using the service, we must decide what personal data to retain and what to delete. We may need to keep some information so that we can confirm that the relationship existed – and that it has ended – as well as some of its details. For example contact details might be useful to keep so that we can deal with any complaints they might make about the provided service.</li> </ul>	<p>Solution approved by the Information Security Officer and the development team manager of ED, LUH and the Project Coordination Team</p>

<sup>22</sup> "The Guide to Data Protection, How much do I need to know about data protection?", UK Information Commissioner's Office, 1 June 2010.

	<p>In that sense, we may need to keep personal data so we can defend possible future legal claims. Unless there is some other reason for keeping it, personal data should be deleted when such a claim could no longer arise.</p> <p>At the end of the retention period, the platform will flag data entries for review or delete them after a pre-determined period, especially useful since many records of the same type are held.</p> <p>However, there is a significant difference between permanently deleting a record and archiving it. [22] Archived data can reduce the risk of misuse or mistake at the expense of availability. Moreover, archived data must comply with the data protection principles and be readily accessed by their owner. If it is deemed necessary to delete data from the platform, it should also be deleted from any back-up made by the platform.</p>	
<b>DAAT-03</b>	<p>We should hold personal data about an individual that is sufficient for the purpose we are holding it for in relation to that individual and we should not hold more information than needed for that purpose.</p> <p>When it comes to sensitive personal data, we should make sure we maintain only the minimum amount of information needed.</p> <p>If particular information regarding certain individuals only is needed, we should collect it ONLY for those individuals, since the information is likely to be excessive and irrelevant in relation to other people.</p> <p>We should NOT hold personal data just in case that it might be useful in the future.</p> <p>In order to access whether we are holding the right amount of personal data, we must first be clear about why we are holding and using it. If collected personal data is not adequate to process the purpose in question, we may collect more personal data than originally anticipated.</p>	<p>Solution approved by the Information Security Officer and the development team manager of ED, LUH and the Project Coordination Team</p>
<b>BIO-01</b>	<p>Solution will be based on approved solution for DAAT-01.</p>	<p>Solution approved before by the Hungarian National Data Protection Authority, [REDACTED]</p>
<b>FMT-01</b>	<p>Include redundancy and diversity to safeguard against "accidental loss or destruction of, or damage to, personal data".</p> <p>Encryption of the hard disks to avoid possible theft or misused of the personal information stored.</p>	<p>Solution approved by everis ADS Information Security Officer and the development team manager of everis ADS, LUH and the Project Coordination Team</p>



	Ensure web servers are exposing only the intended content and only to the intended systems within iBorderCtrl systems.	
<b>FMT-02</b>	Solution will be based on approved solution for DAAT-02	Solution approved by everis ADS Information Security Officer and the development team manager of everis ADS, LUH and the Project Coordination Team
<b>ELSI-01</b>	Solution will be based on approved solution for DAAT-03.	Solution approved by the development team manager of STREMBLE
<b>ELSI-02</b>	At the consent form, the reasons for obtaining the personal data and how they are going to be processed/utilised is thoroughly analysed.	Solution approved by the development team manager of STREMBLE
<b>TUA-01</b>	We will make sure that they read the consent form/privacy notice.	Solution approved by the Information Security Officer and the development team manager of ED, LUH and the Project Coordination Team
<b>TUA-02</b>	Consent form will explicitly state how an individual can withdraw hers/his consent.	Solution approved by the Information Security Officer and the development team manager of ED, LUH and the Project Coordination Team
<b>TUA-03</b>	TUA application will allow travellers to revoke their consent at any time.  Additionally all data stored for the specific individual will be handled according to the consortium data retention policy.	Solution approved by the Information Security Officer of ED, LUH and the Project Coordination Team
<b>TUA-04</b>	The traveller will have access to their personal data both through: <ul style="list-style-type: none"> <li>• the Traveller User Application;</li> <li>• commonly used electronic format, in case the request is made electronically. The copy will be provided free of charge. However, the consortium can charge a 'reasonable fee' when a request is manifestly unfounded or excessive, particularly if it is repetitive. Consortium might also decide to charge a reasonable fee to comply with requests for further copies of the same information;</li> <li>• remote access to the secure self-service system, thus providing the individual with direct access to his or her information. However, we should</li> </ul>	Solution approved by the Information Security Officer of ED, LUH and the Project Coordination Team

	ensure that this access does not adversely affect the rights and freedoms of others.	
<b>TUA-05</b>	<p>Regarding keeping data up to date, it depends on what the information is used for. Since all collected traveller information relies on it remaining current, it should be kept up to date.</p> <p>Regarding the reasonable steps taken to ensure accuracy, we might have to get independent confirmation that the data is accurate. Again this is on a case-by-case basis and it depends on the nature of the personal data and what it will be used for.</p> <p>The traveller will have access to their personal data through the Traveller User Application and will be able to update them.</p>	Solution approved by the Information Security Officer of ED, LUH and the Project Coordination Team
<b>TUA-07</b>	All information stored in the iBorderCtrl database can be extracted in a structured, commonly used and machine readable form (i.e. .csv)	Solution approved by the Information Security Officer of ED, LUH and the Project Coordination Team
<b>TUA-08</b>	Consent form will explicitly state that personal data are collected for research purposes and that the results of the research don't identify individuals.	Solution approved by the Information Security Officer of ED, LUH and the Project Coordination Team
<b>TUA-09</b>	<p>Additionally to the approved solution for DAAT-01 we will:</p> <ul style="list-style-type: none"> <li>• identify which security certifications are important to us and insist from our Cloud Service Provider (CSP) to demonstrate their conformance. Cloud computing services certification is an important aspect since it provides assurance that our critical security requirements are being met;</li> <li>• check the terms of service of the selected CSP to determine if we are allowed to run security tests (auditing) on the CSP infrastructure, even when we are only targeting our own machines only. Unfortunately the majority of providers have taken a security through obscurity approach, meaning that they don't talk about what security controls they've put in place. Even in that case the CSA Security, Trust and Assurance Registry (STAR) Program can be used to identify and describe the security they have implemented inside their environment;</li> <li>• implement secure APIs. Anonymous access and/or reusable tokens or passwords, clear-text authentication or transmission of content, inflexible access controls or improper authorizations, limited monitoring and logging capabilities<sup>23</sup> will be avoided at all times. Best practices on securing APIs will be followed including: documentation, parameter validation, explicit threat detection, rigorous authentication and authorization, protection against the OWASP top 10 list of common application security flaws, penetration tests and vulnerability assessments;</li> <li>• apply security updates and general software updates, using configuration and patch management and deployment tools;</li> <li>• secure the hypervisor, the environment, and the VMs. To secure the hypervisor we should regularly check for new updates and apply them accordingly. To secure the VMs we should differentiate the traffic coming in and going out from a VM on the same physical host;</li> <li>• automate the process of verifying that a large number of security controls are satisfied for a given system configuration, with the use of</li> </ul>	Solution approved by the Information Security Officer of ED, LUH and the Project Coordination Team

<sup>23</sup> "Before you move to the cloud", InfoSec Institute, April 2013

	<p>security auditing tools. The auditing tools will highlight deployment concerns, while the configuration management tools will simplify the process of changing each system to address the audit concerns;</p> <ul style="list-style-type: none"> <li>• set clear expectations for service between us and the service provider, in a Cloud service agreement (CSA).</li> </ul>	
<b>TUA-10</b>	<p>Since CSPs take ownership of their environment but not the data placed in their environment, it is good practice to contractually bind the CSP from denying responsibility if there is a data breach within its environment. At a minimum we should:</p> <ul style="list-style-type: none"> <li>• Contractually ensure CSPs comply with the European Incident Management guides, as stated here: <a href="https://www.enisa.europa.eu/activities/cert/support/incident-management">https://www.enisa.europa.eu/activities/cert/support/incident-management</a> and</li> <li>• Contractually held CSPs accountable for incident responsiveness, including providing specific time frames for restoration of secure services in the event of an incident.</li> </ul>	Solution approved by the Information Security Officer of ED, LUH and the Project Coordination Team
<b>TUA-11</b>	Same as approved solution for DAAT-02.	Solution approved by the Information Security Officer of ED, LUH and the Project Coordination Team
<b>TUA-12</b>	Only the absolutely necessary traveller information which are going to be used by other components or the border guard will be stored to the iBorderCtrl database	Solution approved by the Information Security Officer of ED, LUH and the Project Coordination Team
<b>TUA-13</b>	At the consent form, the reasons for obtaining the personal data and how they are going to be processed/utilised is thoroughly analysed.	Solution approved by the Information Security Officer of ED, LUH and the Project Coordination Team
<b>BGUA PU - 01</b>	Implementation of data encryption standards and processes of authenticating and authorizing in the backend of the BGUA used for the purpose of traveller data transmission.	Solution approved by the development team manager of JAS
<b>BGUA PU - 02</b>	Implementation of a dataflow procedure that ensures processing of only essential traveller data and deletion (from the local server) as soon as these data are forwarded to relevant modules (DAAT, FMT, BIO).	Solution approved by the development team manager of JAS
<b>BGUA PU - 03</b>	At the consent form, reasons for obtaining the personal data and how they are going to be processed/utilised is thoroughly analysed.	Solution approved by the development team manager of JAS

### Step six: Integrate the DPIA outcomes back into the project plan

Who is responsible for integrating the DPIA outcomes back into the project plan and updating any project management paperwork? Who is responsible for implementing the solutions that have been approved? Who is the contact for any privacy concerns that may arise in the future?

Ref No.	Action to be taken	Date for completion of actions	Responsibility for action	Contact point for future privacy concerns
---------	--------------------	--------------------------------	---------------------------	---





<b>DAAT-01</b>	Include redundancy and diversity within the network	August 2018	Project coordination team to liaise with integrator	Information security officer
	Protection of object storage using LUKS	August 2018	Project coordination team to liaise with development team	Development team manager
	Define appropriate protection policies	August 2018	Project coordination team together with LUH to liaise with development team	Development team manager
<b>DAAT-02</b>	Formal data retention policy within the consent form/privacy notice	August 2018	Project coordination team together with LUH	Project coordination team
	Review how long we keep personal data	August 2018	Project coordination team with LUH to liaise with development team	Project coordination team
	Consider the purpose or purposes we hold the information in order to decide whether to retain it	August 2018	Project coordination team with LUH	Project coordination team
	Securely delete information that is no longer needed	August 2018	Project coordination team to liaise with development team	Development team manager
	Update, archive or securely delete information if it goes out of date	August 2018	Project coordination team to liaise with development team	Project coordination team
	Regularly review personal data and delete anything no longer needed	Periodic	Project coordination team with LUH to liaise with development team	Project coordination team
	Establish standard retention periods for different categories of information	August 2018	Project coordination team with LUH to liaise with development team	Project coordination team
<b>DAAT-03</b>	Collect the absolutely necessary traveller document information	August 2018	Project coordination team with LUH to liaise with development team	Project coordination team
	Do NOT hold personal data just in case that it might be useful in the future	August 2018	Project coordination to liaise with development team	Project coordination team
	Be clear about why we are holding and using traveller document information	August 2018	Project coordination team with LUH	Project coordination team
<b>BIO-01</b>	Define appropriate protection policies	August 2018	Project coordination team together with LUH to liaise with development team	Development team manager

<b>FMT-01</b>	Include redundancy and diversity within the network	August 2018	Project coordination team to liaise with integrator	Information security officer
	Protection of object storage	August 2018	Project coordination team to liaise with development team	Development team manager
<b>FMT-02</b>	Same action as for DAAT-02	August 2018	Project coordination team together with LUH	Project coordination team
<b>ELSI-01</b>	Collect the absolutely necessary traveller information	August 2018	Project coordination team with LUH to liaise with development team	Project coordination team
<b>ELSI-02</b>	Consent form to thoroughly analyse the reasons for obtaining the personal data and how they are going to be processed/utilised	August 2018	Project coordination team with LUH	Project coordination team
<b>TUA-01</b>	User will not be allowed to select the consent box unless she/he has scrolled down the whole consent form.	August 2018	Project coordination team to liaise with development team	Project coordination team
<b>TUA-02</b>	Consent form that explicitly states how an individual can withdraw hers/his consent	August 2018	Project coordination team with LUH	LUH
<b>TUA-03</b>	Allow travellers to revoke their consent at any time	August 2018	Project coordination team to liaise with development team	Project coordination team
	All data stored will be handled according to the consortium data retention policy	August 2018	Project coordination team with LUH to liaise with development team	Project coordination team
<b>TUA-04</b>	Access personal data through the application UI	August 2018	Project coordination team to liaise with development team	Project coordination team
	Provide a copy of the personal data as a csv file	August 2018	Project coordination team to liaise with development team	Project coordination team
<b>TUA-05</b>	Update personal data through the application UI	August 2018	Project coordination team to liaise with development team	Project coordination team
<b>TUA-07</b>	All information stored in the iBorderCtrl database about an individual will be extracted in a csv file	August 2018	Project coordination team to liaise with development team	Project coordination team

<b>TUA-08</b>	Consent form to explicitly state that personal data are collected for research purposes and that the results of the research don't identify individuals.	August 2018	Project coordination team with LUH	Project coordination team
<b>TUA-09</b>	Identify which security certifications are important	August 2018	Project coordination team to liaise with information security officer	Information security officer
	Check the terms of service of the selected CSP to determine if we are allowed to run security tests (auditing) on the CSP infrastructure	August 2018	Project coordination team with LUH	Project coordination team
	Implement secure APIs	August 2018	Project coordination team to liaise with development team	Development team manager
	Apply security updates and general software updates	Continuous	Project coordination team to liaise with information security officer	Information security officer
<b>TUA-10</b>	Contractually bind the CSP from denying responsibility if there is a data breach within its environment	August 2018	Project coordination team with LUH	Project coordination team
<b>TUA-11</b>	Same action as for DAA1-02			
<b>TUA-12</b>	Collect the absolutely necessary traveller information	August 2018	Project coordination team with LUH to liaise with development team	Project coordination team
<b>TUA-13</b>	Consent form to thoroughly analyse the reasons for obtaining the personal data and how they are going to be processed/utilised	August 2018	Project coordination team with LUH	Project coordination team
<b>BGUA PU - 01</b>	Implementation of the most effective data transmission regarding the portable device security	August 2018	Project technical and developers team	██████████ ██████████
<b>BGUA PU - 02</b>	Implementation of the dataflow and structures in which data will be kept, informing the traveller why data are collected.	August 2018	Project technical and developers team liaise with Polish Border Guard Consultant	██████████ ██████████
<b>BGUA PU - 03</b>	Consider the ways of informing the individuals about data collection purposes such as consent forms or declarations including the legal basis.	August 2018	Project technical and developers team liaise with Polish Border Guard Consultant	██████████ ██████████



## 8 Conclusions

The implementation of the iBorderCtrl software platform is progressing as planned and there are no currently foreseen deviations. The progress on the Secure IT infrastructure that hosts the applications (TUA, BMUA and BGUA), the Risk based Assessment tool (RBAT), the External Legacy and Social Interfaces capabilities (ELSI) as well as the three user interfaces are presented in this document. Adherence to technical requirements relevant to each are ensured and presented in this document through a tabular view with short explanations. Data flows, and states of current implementation as well as explanations of any decisions taken have been justified and presented. This deliverable is an interim report, Screen shots, schematics and visualizations are used throughout the document to demonstrate the current state of development, with the expectations that at the end of each task of WP 4 (month 24) when Deliverable 4.2 will be presented these will be presented in their final form. A thorough Data Protection Impact Assessment (DPIA) is also performed in this WP and presented in this deliverable to ensure adherence to Data protection requirements of the project. The DPIA is based on the overall system design and not just the state of the prototype implementation therefore we expect the continuation of adherence to it. However, monitoring of the evolution of each aspect of the system will continue to ensure the end iBorderCtrl platform adheres to all data privacy requirements.

This deliverable demonstrates that a great portion of the development of the iBorderCtrl software platform and related interfaces are complete, adheres to all technical as well as data protection requirements and is on track to be completed on time at M24. Furthermore, considerations of the system -as a whole- were considered to minimize any risk to the in parallel ongoing integration phase of the project.