BehavAuth

Human behaviour based authentication for Internet of Things

Open Call Project for TagItSmart: Grant agreement 688061

WP2: D2.4 Ethical review

Deliverable ID:	**D 2.4**
Deliverable Title:	**Ethical review**
Revision #:	**1.0**
Dissemination Level:	**Public**
Responsible beneficiary:	QBE
Contributing beneficiaries:	All
Contractual date of delivery:	28/02/2018
Actual submission date:	1/3/2018

Start Date of the Project: 1 September 2017 Duration: 6 Months

Contents

Section 1 - General Data Protection Regulation 6

1.1 Key changes from previous directive 6

1.2 Key entities 6

1.3 Data subject rights 7

1.3.1 Breach notification 7
1.3.2 Right to access 7
1.3.3 Right to be forgotten 7
1.3.4 Data portability 7
1.3.5 Privacy by design 7
1.3.6 Data protection officers 7

Section 2 - Personal data in BehavAuth 9

2.1 What personal data are used 9

2.2 In which way are they processed 10

2.3 For what purpose are they processed 11

Section 3 - Actions taken 12

3.1 Engaging privacy experts 12

3.1.1 Interview with privacy expert 12
3.1.2 GDPR Seminar – Penningtons Manches LLP 13
3.1.3 Internet of Things MeetUp – GDPR and Internet of Things 14

3.2 Risk analysis 14

3.3 Requirement analysis 17

Section 4 - Privacy-preserving system 19

4.1 Breach notification 19

4.2 Right to access 19

4.3 Right to be forgotten – Users consent 19

4.4 Data portability 20

4.5 Privacy by design 20

4.6 Data protection officers 20

Section 5 - Conclusion 22

Section 6 - References 23

Document history

Revision	Date	Modification	Authors
1.0	28/02/2018	Final version ready for review	N. Palaghias

Introduction

The main objective of this document is to provide the analysis and the outcome of an ethical review. This ethical review will analyse and discuss any privacy issues due to the use of personal data in the BehavAuth project.

Section 1 provides an introduction to the deliverable about the new EU regulation for data protection called General Data Protection Regulation. This section describes the key changes of the new regulation as opposed to the previous directive that was provided by the European Commission (EC). Then, a brief description of the key entities of the new regulation is provided in order to enable the reader to understand the terms used throughout this document. This document maintains consistency of the terms mentioned in order to allow the reader to understand the references from the GDPR. Furthermore, the data rights ensured by the EC through the GDPR are briefly described and analysed, to understand the requirements of the new regulation and in order to apply the new requirements to the BehavAuth system.

In **Section 2**, an analysis of the personal data included in the BehavAuth project is performed that will allow the derivation of the data protection requirements of the BehavAuth project. The section initiates with the description of the personal data and what type of personal data are collected through the BehavAuth project. Through this the reader will receive the understanding of what kind of personal data are collected. In the next subsection, the processing components are detailed for the reader to understand in which way the personal data are processed to derive knowledge. The last subsection provides information about the purpose of the processing of the personal data, so the reader understands why actually the personal data are required to be processed and stored.

Next, **Section 3** describes the actions taken from the BehavAuth project in order to receive a better understanding of the new personal data regulation. BehavAuth project considers personal data very important and has taken appropriate actions from the beginning of the project in order to ensure the users' privacy. The initial step to develop a privacy-preserving system was to engage from the beginning of the project, various stakeholders related to the users' privacy. The initial step was to engage experts on personal data such as Dr. Michele Nati from Digital Catapult (UK) [5], to receive the functional requirements that are needed to comply with the GDPR. The next step for the BehavAuth project, was to attend a GDPR Seminar that allowed BehavAuth to gain the appropriate knowledge to design the architecture of the system in a privacy preserving manner.

Given the knowledge about what is required from BehavAuth as a system related to GDPR and privacy as well as the analysis of the personal data required and the actions taken to better understand the GDPR detailed in the previous sections, **Section 4** provides an analysis of the design of the BehavAuth tool with respect to the users' privacy and the GPDR. Initially, an overview of the BehavAuth mobile app is provided with respect to privacy and how the mobile app tries to ensure in the best way possible, the users' privacy. Following, the document details the privacy-by-design applied to the backend platform and how the personal data are handled in the processing and inference engines.

Finally, this deliverable ends with the conclusion provided in **Section 5**.

European
Commission

Section 1 - General Data Protection Regulation

This section provides an introduction to the new EU regulation regarding personal data, which is called General Data Protection Regulation (GDPR) [4]. In particular, it details what are the requirements of the new regulation, which are the involved stakeholders and what are the rights of the people [1].

1.1 Key changes from previous directive

The new GDPR regulation for data protection has as sole purpose to protect all EU citizens from data breaches. The previous directive was defined in 1995, a time when the world was not so much data-driven. The key idea of the previous directive remains the same, however some additional features are added in order to protect citizens. Also, it should be noted that in 1995 a directive of the European Commission was released encouraging the organisations to follow certain strategies regarding the data protection. The GDPR constitutes a new regulation, meaning that the organisations need to apply this regulation otherwise particular fines will be issued. Following, the key changes from the previous directive are defined:

- *Increased territorial scope*. The new regulation applies to all companies that are processing personal data of data subjects residing in the European Union, without any dependence of where the company's location is. GDPR applies to the processing of personal data by controllers and processors in the EU, even though the processing is not taking place in the EU.
- *Penalties*. Through the new GDPR, organisations that will perform breach of the new regulation can be fined up to 4% of annual global turnover or €20 million. These fines constitute the highest penalties and will be enforced in worst cases where for example the organisation has not received consent from the user to collect data. In smaller penalties, where for example a company does not keep the records in order, the penalty is 2% of the annual global turnover. It should be noted that the rules apply to both controllers and processors.
- *Consent*. This constitutes an important aspect of the GDPR. In the former directive, organisations were able to provide a long "terms & conditions" statement that was accepted at the beginning of the user of the service. With the GDPR organisations should request formal consent through an intelligible and easily accessible form, using a clear and plain language with attached the purpose of the data processing. Also, it should be easy for the data subjects to withdraw the consent at any point.

1.2 Key entities

This section describes the key entities of the GDPR:

- *Personal data* is any information that relates to an identified or identifiable natural person i.e. data subject. This can be a name, an identification number, location data, an on-line identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.
- *Data subject* is any natural person.
- *Controller* is a natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data; where the purposes and means of such processing are determined by Union or Member State lay, the controller or the specific criteria for its nomination may be provided for by Union or Member State law.

- *Processor* is a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller.

The reader should pay particular attention on the above definition of these key terms, as they will be used throughout this document to describe processes and entities.

1.3 Data subject rights

This section details the rights of the data subject according to the new GDPR regulations.

1.3.1 Breach notification

In case of a data breach where it is likely to "result in a risk for the rights and freedoms of individuals", it is mandatory to notify the data subjects. The notification must take place within 72 hours from when the breach was noticed and no further delay should be added.

1.3.2 Right to access

Data subjects have the right to obtain data from the data controller, a confirmation regarding if personal data related to the data subject are being processed, if so where and for what purpose. Also, the data controller has the obligation to provide a copy of the personal data to the data subject, free of charge and in an electronic format.

1.3.3 Right to be forgotten

This right is called also Data Erasure, it provides the ability to the data subject to ask from the data controller to delete any personal data related to the person, to stop any exploitation of the personal data that is taking place as well as to force third-party companies also to stop the processing the personal data of the person. In case the data are no longer being processed for the initially agreed purpose, according to article 17 the data controller needs to cease the processing and erase the personal data. According to [1] the controller should compare the subject's right to "the public interest in the availability of the data" when considering such requests.

1.3.4 Data portability

Through data portability, data subjects have the right to ask from the data controller the personal data related to them. The data controller should provide them in a "commonly use and machine readable format". The personal data in the aforementioned format could be sent to another data controller.

1.3.5 Privacy by design

The privacy by design was included also in the previous directive, however through the GPDR it becomes part of the new regulation. The data controllers need to design the systems with privacy as a key principle of the architecture of the system. The data controllers should store and process only the absolutely necessary data in order to provide their corresponding service; this is called data minimisation. According to the Article 23: "The controller shall … implement appropriate technical and organisational measures ... in an effective way… in order to meet the requirements of this Regulation and protect the rights of data subjects".

1.3.6 Data protection officers

This part of the regulation refers to large public or private organisation that need to get a Data Protection Officer. In the former directive, each of these organisations had to assign a local DPA. Based on the new GDPR, the organisations do not need to submit notifications to each local DPA for the data processing activities and will not need to notify about transfers given the

Model Contract Clauses (MCCs). Regarding the new GDPR, there will be a need for the organisations to internally keep records. For those that perform large scale or special categories of personal data (including criminal convictions and offences) regular and systematic monitoring of personal data, they will need to appoint a DPO. Among the responsibilities of the DPO are:

- Appointed on the basis of professional qualities and expert knowledge on data protection law and practices
- May be a staff member or an external service provider
- Contact details must be provided to the relevant DPA
- Must be provided with appropriate resources to carry out their tasks and maintain their expert knowledge
- Must report directly to the highest level of management
- Must not carry out any other tasks that could result in a conflict of interest

Section 2 - Personal data in BehavAuth

This section provides a detailed description of the personal data used in the BehavAuth project, in which way they are processed and for what purpose the personal data are processed. Each of these subjects are detailed in the following sections in order to provide an analysis of the ethical review regarding the personal data used in the BehavAuth project.

2.1 What personal data are used

The BehavAuth project has as sole purpose to provide an innovative, seamless and secure authentication mechanism that discards the need for the old-fashioned username/password approach and introduces a behavioural based authentication approach. In essence, sensor data such as locations, Bluetooth signal strength and accelerometer provide contextual information to the system. The collected data are forwarded to the knowledge extraction component that through machine-learning models learns the behavioural patterns of the users and is able to authenticate them without requiring from the user to participate in the authentication process.

The above introduction to the system was provided to showcase the significance of the personal data in such project. The sensor data that are extracted from the smartphone of the user provide the appropriate information to the knowledge extraction mechanism to provide the answer regarding whether the user is the person claimed or it is an imposter. In order to perform the knowledge extraction, multiple parameters were taken into consideration and based on the use case and requirement analysis that engaged the potential stakeholders, a decision was made about which personal data should be included in the system. The decision was made taking into account the parameters such as privacy of the users, the performance of the system, the intrusiveness of the system, the easiness in collecting data etc.

Following the above process, BehavAuth project concluded on the following personal data being collected from the users smartphones:

- ***Location***. The BehavAuth mobile app collects the location of the user using the Android SDK. This location is a fusion of sensors i.e. the GPS, the cellular network and the WiFi-based location. Each of these sensors provide an estimation of the current location of the user with certain accuracy. Each of these location sensors has its own drawbacks and for that reason, a fusion technique is provided by the Android SDK to improve the performance of the location sensors.
- ***Bluetooth RSSI.*** The BehavAuth mobile app collects the Bluetooth information from devices in vicinity. In particular, among the information the mobile app collects is the Bluetooth name, MAC address and Bluetooth received signal strength indicator from the devices in vicinity as well as the timestamp at which point the device was discovered. All the information is provided by the Android SDK.
- ***Accelerometer.*** The BehavAuth mobile app collects the accelerometer data from the accelerometer sensor. The accelerometer sensor provides a 3-dimensional vector with the acceleration across 3-dimensions of the smartphone, as well as a timestamp at which point of the time were the data generated.
- ***Identifier.*** The BehavAuth mobile app uses an identifier in order to differentiate among the users of the platform. This identifier allows only the data collector to map and understand who is the actual user and to whom the behavioural data belongs. The BehavAuth platform operates as a processor in the GDPR terms, as it receives some personal data, performs some processing and the result is forwarded to the data collector who is able to identify and map the knowledge to the actual users.

2.2 In which way are they processed

This section details about the processing taking place at the BehavAuth platform. In particular, the previous section provided the details about what kind of personal data are collected by the BehavAuth platform. This section will analyse the way the collected personal data are analysed in order to derive knowledge about the user authentication. This knowledge allows the BehavAuth system to extract certain behavioural cues using machine-learning and then through artificial intelligence the system is able to infer about the authenticity of the user for example if the user is the person claimed or it is an imposter.

The BehavAuth project has developed various data processing mechanisms that leverage personal data from the users' smartphones in order to extract certain behavioural cues and then through the behavioural knowledge engine authenticate the user. The processing is divided into two stages as in most machine-learning approaches. In the first stage, the behavioural cues are extracted from the personal data and then in the second stage the behavioural cues are leveraged to derive the user authenticity knowledge. In the next paragraphs these two stages will be analysed to provide the appropriate knowledge and understanding on how the personal data are processed.

For the first stage the personal data detailed in Section 2.1, are analysed through data processing techniques in order to derive the following behavioural cues [2]:

- **Average location or location hubs.** For most people, because of static locations of work/school/university and home (even the gym), they tend to frequent the same geographic area. Even for many consultants, who travel more than the average office worker, client work tends to be delivered around set project outcomes and therefore, even though their average location for work may change more than usual, there is a habit/pattern of behaviour that occurs during that timeframe. Indeed, consultants and the freelance community could be identified by their lack of average location stasis – a behavioural trait in itself.

- **Travel to our average location.** For many people, they have a tried and tested route to work and home. So not only is there a pattern that could emerge around an average location, but *how* people travel to those average locations is a pattern too. Does their average location include a train station? They are probably taking the train. Does it involve taking the back streets or main roads? The directions people take to travel to their average locations could become a key pattern to identifying them.

- **Speed and distance within the environment.** How we move around our environment is another behavioural trait worth exploring. For some it's by bicycle, for others by car, train or even, their two feet. This largely determines how fast we move and how many steps on average, we take a day. Therefore, looking at how fast someone moves around their environment in addition to the relationship between speed and distance is important.

- **Average social interactions**. During the daily lives people interact with other people in different environments in order to exchange information as social beings. The social interactions taking place can be short-term (a few seconds) but also long-term (more an hour). The duration of the interaction could provide contextual information about the type of relationship among the people. Also, the location in which the social

interaction takes place provides additional contextual information and may allow the characterisation of the type of social relation among the people.

For the second stage of the behavioural pattern inference, the above behavioural cues are leveraged in order to learn the behavioural patterns of the users and further identify people based on these behavioural patterns. Artificial intelligence is used to combine the different information streams in to coherent and informative knowledge that will describe to the TagItSmart platform if the user is the person claimed or this is an imposter.

2.3 For what purpose are they processed

The sole purpose of the processing of the personal data mentioned in the above subsections is to authenticate the users to be able to access the TagItSmart [3] platform with the highest accuracy. All the collected data of the BehavAuth platform are only used for the authentication of the user and for no other purpose.

BehavAuth focuses on discarding the need of the old-fashioned username/password approach and introducing a novel and innovative way by learning the behavioural patterns of the users. In order to perform such task, BehavAuth requires collecting personal data about the user that will allow the accurate authentication of the user, however at the cost of privacy. This constitutes a trade-off between the accuracy of the authentication system and the privacy of the users that requires particular care in order to find the appropriate balance. This balance was achieved by considering many parameters that were analysed in the initial design of the system D1.1 [2].

Among the parameters processed in order to derive the authenticity of the user are:

- Average location or location hubs.
- Travel to our average location.
- Speed and distance within the environment.
- Average social interactions.

The parameters are computed through collecting data from the Android location sensor and the Bluetooth sensor. The data are fed into machine-learning models in order to derive the behavioural cues and then further forwarded into other machine-learning model that aggregates the behavioural cues and authenticates the user. Data minimisation has been applied in the data collection process in order to collect only the absolutely necessary data from the users, in order to authenticate them.

The data collection and processing that has been defined and included in the BehavAuth project has sole purpose to authenticate the user, by providing this authentication service.

Section 3 - Actions taken

This section provides a detailed analysis of the actions taken from the BehavAuth in order to evaluate, secure and preserve the use of personal data in the platform.

3.1 Engaging privacy experts

BehavAuth focuses on collecting sensor data from Internet of Things devices including smartphones in order to authenticate users to the TagItSmart platform based on behavioural patterns. The collected data that will allow the authentication of the users should be correlated with particular users, which will allow the differentiation among the users and further the identification of the users. Meaning that the data collected to identify the users are personal data. BehavAuth pays particular attention to the privacy and security of the users by recognising the existence of personal data in the authentication procedure. For that reason, from the beginning of the project, the team has defined some actions in order to preserve the privacy of the users. Among these actions are the engagement of privacy experts during the whole duration of the project who will provide feedback regarding the privacy aspects of the project.

3.1.1 Interview with privacy expert

As defined in the project plan and pointed out by the consortium of TagItSmart, BehavAuth engaged a privacy expert from Digital Catapult (UK) [5], Dr Michele Nati. Dr Nati is a Lead Technologist for Personal Data and Trust with a great focus on privacy from the point of the end-users. QBE established a collaboration with Dr Nati, in order to receive feedback about the potential privacy issues that may arise during the project. The BehavAuth project was presented to the privacy expert including an overview of the initial tool design. Dr Nati provided feedback and considering the new European Commission data privacy protection scheme (GDPR), he helped us shape the initial tool design from the perspective of users' privacy. A few points that where highlighted from the privacy expert which need to be taken into consideration

- **Consent**. For each biometric or data that will be collected for the authentication of the user, there is a need to receive explicit consent from the user. The consent must be requested in clear and plain language. For example, in case the tool uses three biometrics, an explicit consent in plain language must be requested from the user for each one of the biometrics.

- **Breach notification**. In case of a breach, the end user needs to be informed at most in 72 hour from the breach.

- **Right to access**. Users are allowed to request confirmation if personal data is collected, where, and for what purpose. Furthermore, users could ask for the personal data collected, thus BehavAuth shall provide a copy of the personal data, free of charge, in an electronic format.

- **Right to be forgotten**. The entity or organisation that controls the personal data collected from the users, needs to provide the ability to the user to delete any collected personal data, to cease any dissemination of the data and potentially have third parties halt processing of the data.

- **Data portability.** User can ask from BehavAuth to receive the personal data concerning them and have the right to transmit that data to another company.

- **Privacy by design**. BehavAuth should collect and store only personal data that are necessary for the completion of the authentication process (data minimization), and refuse personal data access to people that are not immediately connected to the authentication process.

3.1.2 GDPR Seminar – Penningtons Manches LLP

Following to the interview with the privacy expert from Digital Catapult, as part of the dissemination activities [7], the BehavAuth project participated in the GPDR Seminar provided by the Pennington Manches LLP law firm [6].

Penningtons Manches LLP is a leading UK law firm which provides high quality legal advice tailored to both businesses and individuals. The company addresses the specific objectives of each client through their technical expertise and offer a personal service founded on a strong team ethic.

The main speaker of the seminar was Joanne Vengadesan. Joanne is a commercial lawyer, specialising in technology transactions and data protection. Her work encompasses structuring and negotiating contracts for complex technology projects, IT agreements, intellectual property and technology licensing, e-commerce and consumer law issues. Joanne has particular expertise in the electric vehicle sector, where she has been assisting a client on its contracts with key partners in relation to electric vehicle charge point infrastructure, as well as on consumer law issues. As an expert in data protection laws, she is advising numerous clients on data protection compliance in readiness for the forthcoming General Data Protection Regulation. She also has expertise in brand protection, and she and her team manage intellectual property portfolios on behalf of a wide range of clients.

After attending the GDPR seminar, some important lessons were learned regarding the applicability of the new regulation. Among them are the following:

- ***Biometric data***. This is a new category of personal data that are added to the GDPR. Biometric data are considered the data leveraged to perform technical processing of personal characteristics to confirm identification. For the use of biometric data there is a need to acquire explicit consent from the user.
- ***Internationality***. The new GDPR regulation is applied to the member states of the EU. However, in case a company is not based in one of the EU countries and trades in one of the EU then there is a need to comply with the GDPR regulation. For example, given the foreseen Brexit where the UK with exit the EU, as BehavAuth focuses on trading also in the EU there will be a need by the BehavAuth to comply fully with the GDPR regulation.
- ***Breach management process***. According to the GDPR there is a requirement of mandatory notification of security breaches without undue delay and no later than 72 hours, if there is a risk to people's rights and freedoms. The data controller must document each breach, its effects and actions taken. Organisations should implement a data breach policy.
- ***Data Protection Officers***. Based on the GDPR data controllers and data processors must designate a DPO if the organisation is a public body or if core activities involve regular and systematic monitoring of personal data on large scale, or involve large

volume processing of special categories of personal data. At this points, this is not applicable for BehavAuth.

- ***Accountability and on-going compliance***. Businesses need to demonstrate that they comply through a series of actions. Among them are internal policies and processes that comply with GDPR, implementation of policies and processes into the business's activities, use of data protection impact assessments, appointment of a DPO and signing up to a Code of Practice or certification scheme. There is a need to pay attention to the Privacy Impact Assessments.

3.1.3 Internet of Things MeetUp – GDPR and Internet of Things

The final event that the BehavAuth project team attended throughout the duration of the project, was the MeetUp Internet of Things in Guildford [8]. The MeetUp was focused on guiding SMEs on how to enter the GDPR era. The MeetUp was organised by a privacy expert Dr Michele Nati from Digital Catapult [5]. The theme of the MeetUp was focused on businesses that collect and process personal and Internet of Things data from their customers and how the GDPR will affect these businesses.

The EU GDPR workshop at the MeetUp was focused mainly on the transparency introduced by the new regulation and also on the users consent and how to apply this in different Internet of Things businesses. People from different backgrounds attended the MeetUp enabling a holistic discussion among the participants.

Some of the key points that were derived from the workshop were:

- **GDPR and SMEs**. There are tools out there that enable SME to comply with GDPR without the need of incorporating a lawyer or a law firm in order to transform the business into GDPR compliant.

- **Consent Management Systems**. There are already solutions that provide consent management platforms in order to enable SME to become compliant with the GDPR in terms of data collection and processing consent. These enable SMEs to acquire and manage consent in an easy and GDPR compliant manner.

- **Transparency.** There is a need for businesses to be more transparent about the data collected and for what reason the data are collected and how they are processed.

- **AI and GDPR.** This constitutes an important issue as artificial intelligence is used widely in todays SMEs in order to derive knowledge. However, often the complexity of the knowledge extraction is very high and it remains a challenge how to explain to the end-user in plain text and in a concise manner how the data are processed and for which reason.

3.2 Risk analysis

This section provides a description of the risks that are related to the BehavAuth system [2]. The risk analysis that was performed had as sole purpose to assist the design of the system by understanding the use cases and requirements of the system, as well as, the risks that are related to the system. Table 3.1 provides a more detailed explanation of the potential risks that arose from the risk analysis. The first column provides a description of the risk, the second column shows the impact of that risk (High/Medium/Low), the third column present the

likelihood of that risk taking place (High/Medium/Low) and the last column provides the way how BehavAuth tries to mitigate that risk.

Table 3.1. Risk analysis

Risk	Impact	Likelihood	Mitigation
Malicious user tries to register to the system, (registering multiple ids)	Medium	Low	Third-party company can only initiate the registration of the user. This risk has been mitigated by the design of the system, where the mobile app itself cannot register a new user at the BehavAuth platform; only the third-party platform is able to perform a new user registration. This assumes the trustworthiness of the third-party company that wants to securely authenticate its users.
Malicious user tries to train the system	High	Low	Third-party company can only initiate the registration of the user. The training process is encapsulated in the registration process of a new user at the BehavAuth system. A new user is authenticated by the third-party platform through credentials, which have generated a token. This token is used to securely transfer the initial training biometric data from the device to the third-party platform and further to the BehavAuth platform. Then, the initialisation method is called by the third-party platform, including the training biometric data. The initialisation method creates and trains the initial models. Thus, the malicious users cannot train the models.
False positive authentication followed by training	High	Low	Multiple biometrics are proposed by the BehavAuth platform which has as sole purpose to increase the authentication rate (meaning reduce to minimum the likelihood) and reduce the rate of false positives. As the correct authentication rate will be very high, the likelihood of faulty accepting an authentication and including the faulty biometric data in the training process, will be very small.
Biometrics fail to authenticate user may not be able to access the service	Medium	Low	Use of secondary username/password credentials to sign in at the service. This constitutes a back-up mechanism in case all biometrics fail to authenticate and the actual user tries still to access the service. A limit of authentication tries will be implemented for the BehavAuth, to avoid attacks such as Brute Force.
Compromise communication channel between third-party platform &	High	Low	An HTTPS communication channel will be established between the third-party and the BehavAuth platforms. In addition, data transmitted between the two end-points will be

BehavAuth. Spoofing of the APP_KEY used in BehavAuth platform.			encrypted to strengthen even more the security among the two systems.
Each biometrics has a certain error of false positives.	Low	High	The use of multiple biometrics is the solution for mitigating this problem. It is a problem that is known in advance. However, the combination of multiple biometrics will allow each approach to compliment the error of other methods, leading to an authentication mechanism with very high accuracy i.e. minimal number of false positives.
Biometrics templates get exposed	High	Low	Encryption of the biometrics templates when created in the BehavAuth platform and stored at the database storage. This will enable the secure storage and retrieval of the biometric templates. The create/read/update/delete methods of the biometric templates are all implemented at the BehavAuth platform, which minimises the exposure of the data to malicious users.
Impersonate at the mobile app level, app sends wrong sensor data	Medium	Medium	An authentication token allows initial secure identification of the mobile app. Then, encryption hides the types of data that are sent to the platform.
Impersonate at user level	Medium	Low	Multiple biometrics constitute the solution for impersonating the biometric at the user level. For example, a malicious user might attempt to trick the voice recognition component, however BehavAuth combines multiple biometrics and behavioural patterns which constitute the impersonation practically impossible.
Reverse engineer the learning mechanism	Medium	Low	The learning mechanism is developed on the cloud and the mobile app implements only data collection process. For that reason, the behavioural pattern learning mechanism is highly unlucky to be exposed to third-parties as the code is preserved at a cloud infrastructure and is not accessible to others except from the BehavAuth developers.
Exposure of personal data in case of data breach	Low	Medium	Anonymous id is only registered. BehavAuth stores only a pseudonymised ID followed by the behavioural personal data. This implements the so-called *unlinkability* of the data, as it is not possible to reach the person for whom the data have been collected, unless there is a link between the pseudonym and the identification data. The identification data are stored at the third-party platform. In that way, the sensor data that are collected are unlikable.

User has requested data and data breach occurs at the end-users side.	N/A	Medium	It occurs outside the ecosystem of BehavAuth as from the point BehavAuth has provided the copy of the data to the user in a portable form and a breach occurs at the user's end, then BehavAuth is not related to that

3.3 Requirement analysis

This section describes the requirements derived from the use cases and the engagement of the various stakeholders such as end-users, psychologists, potential early adopters of the platform and privacy experts [2]. Table 3.2 provides a list of the requirements including category, requirement ID (REQ_ID) and description of each requirement.

Table 3.2. Functional requirements.

Category	REQ_ID	Description
Security	F01	The tool must provide a secure communication channel between the cloud and the mobile devices.
	F02	The platform must provide a secure communication channel between the cloud and the TagItSmart platform.
	F03	The platform must ensure the integrity of all the communications.
Privacy	F04	The platform must should store a pseudonym for each registered user.
	F05	Users' data should be processed only when explicit consent has been given by the users.
	F06	Users should have the ability to stop the data collection and processing at any time.
	F07	Only data required for the authentication mechanism should be collected.
	F08	Users should have the ability to delete the collected data and any extracted knowledge.
ID	F10	Only the pseudonym of each user will be stored, followed by the behavioural data.
	F11	The ID of each user will be anonymised.
	F12	The platform should allow the registration and deletion of users.
Authentication	F13	The authentication mechanism should be user-friendly.
	F14	The authentication mechanism should allow developers to easily integrate with TagItSmart and other third-party platforms.
	F15	The behavioural patterns should accurately identify and verify the users.

European
Commission

Section 4 - Privacy-preserving system

This section provides a detailed description of the privacy actions taken in the design of the BehavAuth system by taking into consideration the new EU regulations defined in the GDPR. The privacy of the BehavAuth system is analysed as a whole in terms of privacy and the actions taken are detailed as well as how these actions are reflected to the end user with respect to the GDPR guideline. The section is structured as follows; the first subsection provides an analysis of the design of the mobile app followed by a detailed description of the backend platform of the system.

4.1 Breach notification

In case of a data breach the users will be notified through email within 72 hours of acquiring the knowledge about the data breach.

4.2 Right to access

At this point the data collected though the mobile app, are processed locally on the device in order to infer if the user is the owner of the device or an imposter. After the knowledge extraction process the data are discarded from the device and are not stored somewhere. The result of the authentication is forwarded to the backend platform in order to inform the user repository about the fact if the user is authenticated or not.

4.3 Right to be forgotten – Users consent

Another important aspect of GPDR is users consent. In the previous EU directive regarding the data collection process, applications should provide a ‘terms & conditions” contract that should be prompted to the user at the first time that the application was launched. In this sense, the user provided the explicit consent and agreed to the data collection process. Based on the GPDR, the applications need to provide to the users in an easy to read and understand way as previously users were lost in the endless text of the terms and conditions. Also, users should be able at any point of time to withdraw their consent about the data collection and data processing.

BehavAuth focused also on the users consent. When the user deploys the mobile app, a particular framework was developed that focuses on receiving and withdrawing the users’ consent. Immediately when the users launch for the first time the mobile app, a specific screen wizard is prompted to the users in order to explain through one sentence the personal data to be collected and for what reason. BehavAuth team paid particular attention in describing the requested consent for personal data in plain English, with easy to understand language and without the use of technical terms. In more details the steps taken for receiving the users’ consent are described below:

- **First launch wizard**. A simple, easy-to-read and clear wizard was developed in order to request the users consent when the mobile app is launched for the first time. Screens are prompted to the users explaining initially the use of personal data and then for each type of personal data a screen is presented describing in a simple and easy way why that personal data is crucial for the app to be collected and next it requests the explicit permission from the user to collect that data. The user has the ability to accept or decline the collection and processing of various personal data.

- **Sensor permission**. The collection of personal data relates highly with the integrated smartphone sensors. This means that the users are able to withdraw any permissions given for data collection from the device sensors using the native Android interface. The mobile app is continuously monitoring the state of the sensor permissions in order to perform the data collection from the particular sensor interface. If the permission is withdrawn the mobile app continuous to operate without collecting and further processing the sensor data from the particular sensor interface.
- **Data collection permission**. In addition to the sensor permission, the user is able through the interface of the mobile app to withdraw the permission for data collection and processing at any point. This functionality has been implemented in order to comply with the requirements of GDPR, in order to allow the users at any point to initiate or stop the data collection as well as the data processing. User is informed that in case the data collection of a particular interface is stopped then the performance of the authentication mechanism will decrease.

4.4 Data portability

Through data portability, data subjects have the right to ask from the data controller the personal data related to them. The data controller should provide them in a "commonly use and machine readable format". The personal data in the aforementioned format could be sent to another data controller.

4.5 Privacy by design

According to GDPR, data processors and data controllers are obliged to collect only the data required for the particular process. Until now it is a common phenomenon that various applications are requested access to personal data that are not required by the actual application. For example, a mobile app that allows the users to have access to wallpaper images for the smartphone request multiple permission for access data such as photo gallery, contacts and location. The access to photo gallery is requested to enable users to save wallpapers to the photo gallery, however the contacts and location permissions are not relevant to the application's functionality.

For that reason, BehavAuth took particular care to ensure that only the absolute necessary data are collected by the mobile app and in general by the overall system. In detail, BehavAuth requires the following personal data in order to authenticate the user:

- *Location*. The location data are imperative in order to extract behavioural cues related to the points of interest, the average location, the average movement of the user.
- *Bluetooth information*. The Bluetooth data are required in order to identify the social interactions among the users as well as the interactions with various objects.
- *Accelerometer*. The accelerometer data are required in order to compute the average speed of the users.
- *Identifier (ID)*. An identifier is required in order to establish the ownership of the data. This identifier is anonymised through pseudonym.

4.6 Data protection officers

At this point BehavAuth is a small data processing entity and given the current GDPR regulation and the advice received from the dissemination activities related to GDPR, BehavAuth does not require a Data protection officer.

European
Commission

Section 5 - Conclusion

This deliverable described the ethical review of the BehavAuth tool based on the new EU regulation for protection of personal data, the analysis of the personal data leveraged in the BehavAuth platform, the actions taken to design a privacy preserving system and also the privacy by design implemented in the architecture of the system.

This document started by describing the new regulation for protection of personal data in the EU. The description started by inducing the key changes performed with respect to the previous directive. Then, the key entities defined in the new regulation were outlined and explained to provide an understanding to the reader about the terms used throughout this whole document. Lastly, the data rights for the users are discussed in a simple and easy to understand way for the reader to receive an initial overview of the GDPR regulation.

Next, the personal data used in the BehavAuth project are analysed and detailed regarding which data are collected, how they are processed and for which purpose. It is important for the reader to understand the existence of personal data in the BehavAuth project. Thus, the personal data collected from the smartphone sensors, were explained in detail in order to understand which exactly are the collected data. Then, the way the personal data are processed was detailed to the user and finally for which purpose the personal data are processed and then used.

Following, the action taken to preserve users' privacy were presented. The action taken to preserve users' privacy included the engagement of the stakeholders, the risk analysis and also the requirement analysis. For the engagement of the stakeholders BehavAuth from the beginning of the project initiated the contact with related privacy experts in order to shape the use cases and the requirements of the system in a privacy-preserving way. Among the privacy experts engaged were Digital Catapult and the law firm Penningtons Manches. Next, a risk analysis was presented that will enable the reader to understand the risks of BehavAuth in terms of privacy. Last, the requirement analysis with respect to the privacy of the users was presented.

Finally, an analysis of the privacy-by-design applied to the BehavAuth tool architecture was provided, taking into account the new GDPR, the personal data required by the tool, the actions taken in order to define a privacy-preserving system. The privacy-preserving architecture of the BehavAuth system is analysed based on the key GDPR rights of the data subject. This constitutes an ethical review of the BehavAuth architecture and given the use of the new GDPR regulation, the requirements set by privacy experts as well as the personal data collected by the BehavAuth platform.

Section 6 - References

1. EU GDPR - https://www.eugdpr.org/key-changes.html
2. BehavAuth D1.1 Initial design based on the use case and requirement analysis
3. TagItSmart EU project – www.tagitsmart.eu
4. General Data Protection Regulation - http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016R0679
5. Digital Catapult - https://digital.catapult.org.uk/
6. Pennington Manches - https://www.penningtons.co.uk/
7. BehavAuth D2.3 Impact and dissemination activities report
8. MeetUp Internet of Things Guildford - https://www.meetup.com/Internet-of-Things-Guildford/events/246235093/