



Persistent Personal Data Vaults Empowering a Secure and Privacy  
Preserving Data Storage, Analysis, Sharing and Monetisation Platform

## D6.2

# Pilot Scenarios and Implementation Plan

<b>Editor(s)</b>	Nikos Achilleopoulos
<b>Lead Beneficiary</b>	MAGGIOLI
<b>Status</b>	Final
<b>Version</b>	1.00
<b>Due Date</b>	30/06/2021
<b>Delivery Date</b>	30/06/2021
<b>Dissemination Level</b>	PU



DataVaults is a project co-funded by the European Commission under the Horizon 2020 Programme (H2020-ICT-2019-2) under Grant Agreement No. 871755 and is contributing to the BDV-PPP of the European Commission.

<b>Project</b>	DataVaults – 871755
<b>Work Package</b>	WP6 - Multi-Layer Demonstrators Setup, Operation and Business Value Exploration
<b>Deliverable</b>	D6.2 - Pilot Scenarios and Implementation Plan
<b>Editor(s)</b>	MAGGIOLI - Nikos Achilleopoulos
<b>Contributor(s)</b>	OSFP PRATO Suite5 ANDAMAN7 MIWENERGIA PIRAEUS UBITECH
<b>Reviewer(s)</b>	ATOS Darío Ruiz López FRAUNHOFER Yury Glikman

<b>Abstract</b>	This deliverable presents in detail the implementation plan for each demonstrator to be executed in the DataVaults project. These plans have been carefully designed to be executed in a consecutive order, following the development releases of the platform, which is expected to be rolled out in three phases during the project to be tested and evaluated, towards the final release that is scheduled for the end of the project.
<b>Disclaimer</b>	<p>The information and views set out in this publication are those of the author(s) and do not necessarily reflect the official opinion of the European Communities. Neither the European Union institutions and bodies nor any person acting on their behalf may be held responsible for the use which may be made of the information contained therein.</p> <p>© Copyright in this document remains vested with the DataVaults Partners</p>

## Executive Summary

This deliverable presents in detail the implementation plan for each demonstrator to be executed in the DataVaults project. These plans have been carefully designed to be executed in a consecutive order, following the development releases of the platform, which is expected to be rolled out in three phases during the project to be tested and evaluated, towards the final release that is scheduled for the end of the project.

As such, in each release, a different set of features/functionalities shall be made available which will be tested by different demonstrators, while at the same time the audience who will evaluate each version of the DataVaults platform shall also evolve, starting from a small set of individuals that will test the platform in closed-group experiments, to opening up to the real Data Owners who are the stakeholders engaged in each organisation who acts as a demonstrator.

For all the above in each demonstrator a set of different scenarios has been defined, whereas each such scenario is split into different phases that are mapped to the features to be introduced in each release of the platform. Moreover, for each such scenario, the key audience is also identified, alongside with other relevant information, such as the ways to reach this audience, or the envisaged compensation to be offered to Data Owners, and the identification of Data Seekers/collaborators which might be interested to join the platform at the final stages of testing, in an effort to holistically describe the operational conditions of each such demonstration.

For the above scenarios, the demonstrators have also worked with the technical partners that support them in WP6, in order to also define and plan the implementation of the software development and integration activities that should be performed in order to be able to connect to data sources that are necessary for the execution of the scenarios, and for providing some extra features which are specific to the demonstrators and are not part of the core DataVaults platform to be developed and integrated under WP5.

Finally, a set of business relevant KPIs is also identified, aiming to capture the business benefits springing from the execution of each demonstrator, as there is a need to evaluate the business benefits of the operation of DataVaults for Data Seekers, in parallel with the technical evaluation to be performed. It is noted that technical evaluation and validation, will be performed under WP5, and thus the scenarios that defined are deliberately not including aspects of technical testing regarding the main functionalities of the DataVaults platform.

## Table of Contents

1	Introduction .....	9
1.1	Relation to Other Tasks And WPs .....	9
1.2	Document structure .....	9
2	DataVaults Demonstration Baseline Activities .....	11
2.1	Demonstration Evolution Plan .....	11
2.2	Demonstration High Level Execution Plan .....	12
3	Demonstrator #1 - Sports and Activity Personal Data .....	13
3.1	Business Description and the Need for DataVaults .....	13
3.2	Target Audience Description – Data Owners and Data Seeker .....	14
3.2.1	Data Owners.....	14
3.2.2	Data Seekers.....	14
3.2.3	Means to Reach other Data Seekers .....	15
3.3	Demonstrator Scenarios.....	15
3.3.1	Scenario A - Club Fans and Members Personal Data Marketplace.....	15
3.3.2	Scenario B - Athletes Sports and Activity Data Sharing .....	18
3.4	Demonstrator Technical Development Needs .....	21
3.5	Compensation Methods Considered.....	21
3.6	Demonstrator’s Ethics Considerations .....	21
3.7	Impacts Expected .....	22
3.8	Demonstrator’s Activities Timeline .....	25
4	Demonstrator #2 - Strengthening Entrepreneurship and Mobility .....	28
4.1	Business Description and the Need for DataVaults .....	28
4.2	Target Audience Description – Data Owners and Data Seeker .....	29
4.2.1	Data Owners.....	29
4.2.2	Data Seekers.....	29
4.2.3	Means to Reach other Data Seekers .....	29
4.3	Demonstrator Scenarios.....	29
4.3.1	Scenario A - Smart Mobility Services for Individuals .....	29
4.3.2	Scenario B - Empowering local entrepreneurship .....	31
4.3.3	Scenario C - Services for Personalized cultural and touristic experiences.....	34
4.4	Demonstrator Technical Development Needs .....	36
4.5	Compensation Methods Considered.....	36
4.6	Demonstrator’s Ethics Considerations .....	36

4.7	Impacts Expected .....	37
4.8	Demonstrator's Activities Timeline .....	38
5	Demonstrator #3 - Secure Healthcare Data Retention and Sharing .....	41
5.1	Business Description and the Need for DataVaults .....	41
5.2	Target Audience Description – Data Owners and Data Seeker .....	42
5.2.1	Data Owners .....	42
5.2.2	Data Seekers .....	42
5.2.3	Means to Reach other Data Seekers .....	43
5.3	Demonstrator Scenarios .....	44
5.3.1	Scenario A - Get data from Andaman7 and provide to health sector .....	44
5.3.2	Scenario B - Data collection and Andaman7 improvements .....	46
5.4	Demonstrator Technical Development Needs .....	47
5.5	Compensation Methods Considered .....	48
5.6	Demonstrator's Ethics Considerations .....	48
5.7	Impacts Expected .....	49
5.8	Demonstrator's Activities Timeline .....	50
6	Demonstrator #4 – Smart Home Personal Energy Data .....	52
6.1	Business Description and the Need for DataVaults .....	52
6.2	Target Audience Description – Data Owners and Data Seeker .....	52
6.2.1	Data Owners .....	52
6.2.2	Data Seekers .....	53
6.2.3	Means to Reach Data Seekers .....	54
6.3	Demonstrator Scenarios .....	55
6.3.1	Scenario A - PV installation design for self-consumption .....	55
6.3.2	Scenario B - Improve profiling of the clients to enhance energy efficiency .....	60
6.3.3	Scenario C - Energy consumption patterns with personal preferences .....	63
6.4	Demonstrator Technical Development Needs .....	66
6.5	Compensation Methods Considered .....	66
6.6	Demonstrator's Ethics Considerations .....	66
6.7	Impacts Expected .....	67
6.8	Demonstrator's Activities Timeline .....	68
7	Demonstrator #5 - Personal Data for Municipal Services and the Tourism Industry .....	71
7.1	Business Description and the Need for DataVaults .....	71
7.2	Target Audience Description – Data Owners and Data Seeker .....	71

7.2.1	Data Owners.....	71
7.2.2	Data Seekers.....	72
7.2.3	Means to Reach other Data Seekers .....	72
7.3	Demonstrator Scenarios.....	73
7.3.1	Scenario A - Access to personal data for the analysis of mobility solutions.....	73
7.3.2	Scenario B - Access to personal data for the improvement of cultural offer in the city	76
7.3.3	Scenario C - Access to personal data for the delivery of personal certificates....	79
7.4	Demonstrator Technical Development Needs.....	81
7.5	Compensation Methods Considered.....	81
7.6	Demonstrator's Ethics Considerations .....	81
7.7	Impacts Expected .....	82
7.8	Demonstrator's Activities Timeline .....	84
8	Conclusions and Next Steps .....	87

## List of Figures

Figure 1: Demonstration High Level Execution Plan .....	12
Figure 2: Data owners data integration in DataVaults.....	53
Figure 3: Design of a photovoltaic system .....	55
Figure 4: Annual electricity consumption analysis.....	56
Figure 5: Hourly energy consumption profile analysis.....	56
Figure 6: Monthly energy consumption profile analysis.....	56
Figure 7: Energy efficiency class approximation, Spain .....	61

## List of Tables

Table 1: Impact Target KPIs for Demonstrator #1.....	23
Table 2: Execution Timeline for Demonstrator #1 .....	27
Table 3: Impact Target KPIs for Demonstrator #2.....	38
Table 4: Execution Timeline for Demonstrator #2 .....	40
Table 5: Impact Target KPIs for Demonstrator #3.....	49
Table 6: Execution Timeline for Demonstrator #3 .....	51
Table 7: Description of data needs for Demonstrator 4 - Scenario A.....	55
Table 8: Data Availability for Demonstrator 4 - Scenario A .....	59
Table 9: Data Needs for Demonstrator 4 - Scenario A.....	60
Table 10: Description of data needs for Demonstrator 4 - Scenario B .....	60
Table 11: Data Availability for Demonstrator 4 - Scenario B .....	63
Table 12: Data Needs for Demonstrator 4 - Scenario B .....	63
Table 13: Description of data needs for Demonstrator 4 – Scenario C .....	64
Table 14: Data Availability for Demonstrator 4 - Scenario C .....	66
Table 15: Data Needs for Demonstrator 4 - Scenario C .....	66
Table 16: Impact Target KPIs for Demonstrator #4 .....	67
Table 17: Execution Timeline for Demonstrator #4 .....	70
Table 18: Impact Target KPIs for Demonstrator #5 .....	83
Table 19: Execution Timeline for Demonstrator #5 .....	86

## Terms and Abbreviations

<b>API</b>	Application Programming Interface
<b>CRM</b>	Customer Relationship Management
<b>EHR</b>	Electronic Health Record
<b>GDPR</b>	General Data Protection Regulation
<b>KPI</b>	Key Performance Indicator
<b>MVP</b>	Minimum Viable Product
<b>QoL</b>	Quality of Life
<b>SME</b>	Small Medium Enterprise



# 1 INTRODUCTION

The successful implementation of WP6 in terms of effectively testing the DataVaults solution and providing the necessary feedback to evaluate the platform relies on the execution of the demonstrators in a coordinated and unified manner, forming a representative sample audience of users to evaluate each version.

Deliverable D6.1 “Project and Pilots Evaluation and Impact Measurement Plan” and the present deliverable D6.2 are compiled in parallel. D6.1 provides an inclusive demonstrators’ evaluation framework as well as a general guideline document to be used to monitor and align the demonstrators’ phases.

The present deliverable D6.2 lays down the planning and coordination of the demonstrators, detailing the plans to be executed in tasks 6.3-6.7 of WP6. The appropriate guidelines are authored to support demonstration responsible partners in their implementation of solutions with the help of the DataVaults platform.

---

## 1.1 RELATION TO OTHER TASKS AND WPs

The input to “Pilot Scenarios and Implementation Plan” consists of the requirements definition in WP1 and the technical implementation of the platform in WP3, WP4 and WP5. Deliverable D1.3 “DataVaults MVP and Usage Scenarios” describes the functionality offered by the DataVaults platform, while the implementation & integration Work Packages, i.e. WP3, WP4 and WP5, build incremental releases of the platform, namely Alpha on M20, Beta on M24 and V0.5 on M30, with V1.0 expected to be the final product on M36 (i.e. at the end) of the project.

Deliverable D6.1 covers the Evaluation Framework to be employed in DataVaults, while Annex A of D6.1 sorts out the details related to it and will become a living document, to be updated and added to throughout the rest of the project

The present deliverable shall be used as input to the next tasks set in operation once the operation of the demonstrators starts.

Extensive data collection, regarding the experience of the demonstrator partners with the DataVaults platform, will be conducted. It will be documented in deliverables D6.3 “Pilots Evaluation of Alpha Platform Version” and D6.4 “Demonstrators’ Evaluation of Beta Platform Version”.

Based on the data collected, an overall assessment and evaluation of DataVaults will be carried out. It will be reported in deliverable D6.5 “Final Evaluation and Impact Assessment Report”.

---

## 1.2 DOCUMENT STRUCTURE

The structure of this document is as follows:

**Section 1** provides the introduction and the relation of this deliverable to other deliverables, providing its positioning within the project.

Going forward, **Sections 3,4,5 and 6** provide the description, target audience, scenarios, evolution plan and expected impacts of each DataVaults demonstration, namely

Demonstrator #1 which is about Sports and Activity Personal Data, Demonstrator #2 which focused on Strengthening Entrepreneurship and Mobility, Demonstrator #3 that is working on Secure Healthcare Data Retention and Sharing, Demonstrator #4 that has its interest on Smart Home Personal Energy Data and Demonstrator #5 which is exploiting Personal Data For Municipal Services and The Tourism Industry.

**Section 8** draws the conclusions of the present deliverable.

## 2 DATAVAULTS DEMONSTRATION BASELINE ACTIVITIES

### 2.1 DEMONSTRATION EVOLUTION PLAN

From development & integration Work Packages, i.e. WP3, WP4 and WP5, we expect three releases, namely Alpha on M20, Beta on M24 and V0.5 on M30, with V1.0 expected to be the final product on M36 (i.e. at the end) of the project.

At each release, a different set of features/functionalities shall be made available.

The Alpha version shall offer basic functionalities (as described in the final MVP (deliverable D1.4 and in deliverables D3.1 and D4.1), for example for the Individual: fetch simple data sources (e.g. structured data from known API), define access control and share personal data under the access control schema, write the Individuals' sharing configuration on the ledger; or for the Data Seeker: simple search on datasets, acquisition of dataset (no compensation), write the Data Seekers retrieval of dataset(s) to the ledger, basic analytics.

The Beta version shall offer full functionality, with non-critical features (e.g. SSE) being at an early experimental level of maturity.

Version 0.5 shall fix all major defects identified after the evaluation of Alpha and Beta versions and shall offer full functionality plus mature, experimental-to-functional level of non-critical features.

At the same time, the audience who will evaluate each version of the DataVaults platform shall also evolve. The Alpha version shall be evaluated mainly by a small, closed group of users, usually employees of the demonstrator who have also taken part in setting up DataVaults platform at the premises of the demonstrator, i.e. with good knowledge of the goals of the project and both the technical details (e.g. connection of the data sources) and the business details (e.g. the goals that each demonstrator aspires to achieve via the adoption of the DataVaults platform). The Beta version shall be evaluated by a larger scale closed group, or a limited open group, depending on the nature of each demonstrator. At this stage, the users should have limited to no knowledge of the technical and business details behind the demonstrator's adoption of the DataVaults platform. Finally, V0.5 shall be evaluated by a fully open group of users, who could range from invitees of the demonstrator to groups of clients of the demonstrator or even the open public.

The test scenarios devised for all five demonstrators in sections 3, 4, 5, 6 and 7 of the present deliverable are split into 3 phases, which are designed in order to gradually evolve at full scale, following the Alpha, Beta and V0.5 releases of the DataVaults platform. In addition, the audience expected to evaluate each version of the DataVaults platform is described in the abovementioned demonstration evolution plans, according to the particularities of each demonstrator.

It is stressed out that the evaluation of the system and of the scenarios will be done within WP6 and the methods to be employed are the ones specified in Deliverable D6.1 which is providing the evaluation plan. The current deliverable provides the different routes that will

be taken to unfold the platform’s features in the different demonstrator settings, running the necessary scenarios which will provide the input to be used for the evaluation within WP6.

## 2.2 DEMONSTRATION HIGH LEVEL EXECUTION PLAN

The high-level execution plan for the whole demonstration effort (WP6) in Year 2 and Year 3 of the project and its relations with input coming from development & integration Work Packages, i.e. WP3, WP4 and WP5, is laid down in Figure 1 below.

From the development & integration Work Packages, we expect three releases, namely Alpha on M20, Beta on M24 and V0.5 on M30, with V1.0 expected to be the final product on M36 (i.e. at the end) of the project.

The testing effort in WP6 has already started with pilot planning and preparation activities, well before the Alpha version is available, and shall conclude with the current deliverable, D6.2 “Pilot Scenarios and Implementation Plan”. The development of the connectors and necessary customisations at each demonstrator follows this phase, so that every demonstrator can get ready to connect to the Alpha version of the platform around M20. The evaluation of the Alpha version by a small, closed group of users lasts from M20 to M24 and concludes with deliverable D6.3 “Pilots Evaluation of Alpha Platform Version”.

Once the Beta version of the DataVaults platform is available on M24, the evaluation of the Beta version by a larger scale closed group, or a limited open group of users starts and lasts until M30. It concludes with deliverable D6.4 “Demonstrators' Evaluation of Beta Platform Version”.

The availability of V0.5 on M30 marks the start of the final testing round within WP6 and the project as a whole. It is evaluated by a larger, open group of users and the results are reported in deliverables D6.5 “Final Evaluation and Impact Assessment Report” and D6.6 “DataVaults Scaleup Roadmap and Key Takeaways”.

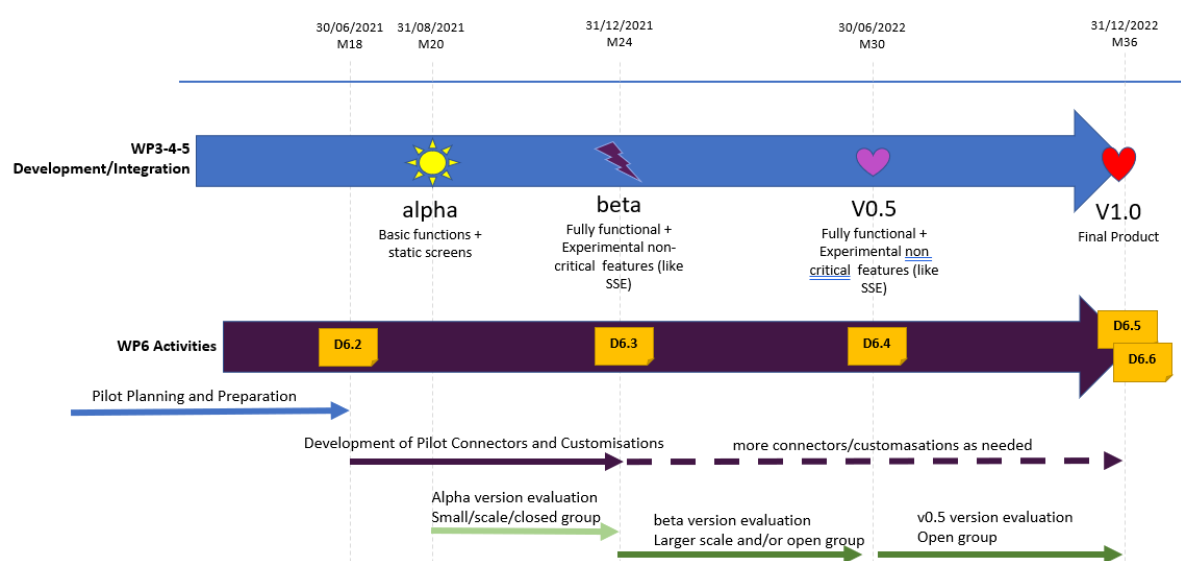


Figure 1: Demonstration High Level Execution Plan

### 3 DEMONSTRATOR #1 - SPORTS AND ACTIVITY PERSONAL DATA

#### 3.1 BUSINESS DESCRIPTION AND THE NEED FOR DATAVAULTS

Olympiacos has a large base of 130,000 members and 60,000 fans who register or renew their subscription on an annual basis. This base and the satisfaction of the members & fans are very important issues for the club as they are the most important source of revenues. Moreover, as the club has eighteen different sport departments at the competitive level and the academies, it needs to manage a large base of 2,700 professional and young athletes, including important contact, activity and medical details.

A crucial issue for the club is the management of the data included in these two bases:

- On the one hand, the data included in the database of members & fans is compulsory to organize and coordinate procedures such as the participation of the members in the club's General Assembly, the right to vote in the Administrative Elections and take part in the decision-making process. Moreover, this database is also beneficial for the planning and the implementation of segmented and targeted marketing campaigns, the improvement of fans' engagement and finding sponsors. This database concerns personal information such as demographics, contact details, personal governmental data, social media accounts, etc. For the club, it would be extremely beneficial to collect information such as user preferences, social media activity and mobility data in the future to focus on the real preferences of the members & fans and improve the services offered to them and the fans' engagement in order to ensure their loyalty.
- On the other hand, the athletes' base requires better management of the results of the ergometric and medical examinations, as well as the statistical reports regarding their performances in the training and the matches. This database also includes personal information, demographics, personal governmental data, and activity data (sport or exercise). Better management is very crucial for the club to adapt and plan the training sessions and the team tactics, make player transfers with specific physical skills, cover the athletes' expectations by offering the appropriate medical and sport equipment.

Regarding the data collection of the members' and fans' base, the club uses the 2016 Microsoft Dynamics CRM system. Moreover, it offers a possibility to register and renew their subscription via the official club website ([olympiacossp.gr](http://olympiacossp.gr)). In this case, an important issue is that a number of fans are hesitant to provide all the requested data because they are concerned about data protection. So, they prefer to leave some data such as demographics and social media accounts incomplete or avoid updating them on a regular basis. So, the available CRM tools seem to export incomplete reports making it difficult to implement a successful strategic and marketing plan. As far as the athletes' data collection is concerned, this is a manual procedure facilitated by Microsoft Office platforms (e.g., Excel, Word). However, the club has also started to use the above-mentioned CRM tool for the data management of the athletes. A critical issue is that the version of the CRM that the club uses, doesn't allow the entry of more complicated data such as the results of the ergometric and

medical examinations. So, there is a need for an appropriate platform that facilitates the entry of more complicated information.

---

## 3.2 TARGET AUDIENCE DESCRIPTION – DATA OWNERS AND DATA SEEKER

---

### 3.2.1 Data Owners

**Fans and members** of the club, who own and continuously generate personal data, social activity data and location data, usually without understanding what it comprises, or having any means to manage and share such data.

The principal type of data envisaged to be collected from fans and members in the context of the current demonstrator are:

- Facebook profile activity
- Instagram profile activity
- Location History
- How many times each fan/member was at the stadium and when/which game they followed, as the online ticketing system is linked to the fan/member card number

**Athletes** who belong to the club across eighteen different sport departments at the competitive level and the Academies. They possess personal data and continuously generate athletic activity data and ergometric and medical examination data, in parallel with their activities as athletes of the club.

The principal type of data envisaged to be collected from athletes in the context of the current demonstrator are:

- Athletic activity data (e.g., from wearable sensors during practice)
- Ergometric and medical examination data (e.g., from club measurements recorded at the beginning of the season)

---

### 3.2.2 Data Seekers

Various data seekers are envisaged to be interested in the data offered by the demonstrator. This could be club sponsors, for example a telecommunications/mobile telephony company affiliated with the club, who could be interested to search through the anonymised personal and online activity data of members & fans of the club, in order to offer better customized and targeted products and services towards, for example, the average season fan or the high-income fan who attends the most prestigious matches.

At the level of athletes, Data Seekers could be an athletic equipment provider who are interested in searching through the athletic activity data of athletes of the club, in order to offer better customized and targeted products and services (e.g. depending on what sport each athlete does).

The envisaged Data Seekers shall be on-boarded by the club (e.g. by direct invitation) either in the final phase of the scenarios or at the latest in the post-project phase.

---

### 3.2.3 Means to Reach other Data Seekers

Reaching other Data Seekers will be done in the first phases of the demonstrator by direct invitation by the club.

---

## 3.3 DEMONSTRATOR SCENARIOS

---

### 3.3.1 Scenario A - Club Fans and Members Personal Data Marketplace

In this scenario, fans and members of the club are encouraged to collect their personal data, understand what it comprises, manage and share them through the DataVaults platform. For the fans and members, awareness on how the type and quantity of personal data relates to compensation offered from the club shall encourage the former to push forward sharing and keeping it up these data. For the club, this will help build a stronger relationship/interaction with the fans and members, understand their needs and offer better services, as well as incentives to share more of their personal data, respond to the requests that may come from different organizations (e.g., Sponsors/NGOs/Sports Federations/Local authorities who want to run a campaign or host an event for the club members/fans and athletes) and find new sponsors/partnerships.

The main goal of this scenario is to collect information related to social media activity and the preferences/likes of individuals who are already fans and members of the club. A secondary goal would be to collect location data related to the position of the individual, for example if s/he is present at the stadium during the time the club's team is playing. Collecting and analysing such data could assist the club to combine personal data already existing in its systems with social media and location activity in order to engage and interact better with the fans and members. Evidently, such information could help reorganize the marketing plan (new market segmentation, marketing campaigns for specific target groups, finding specific sponsors, etc.), attract new sponsors based on personas (collective profiles) extracted from DataVaults, improve the services offered to the fans and members and finally, increase the revenues and online presence of the club. All this, in a transparent exchange between the individual and the club, respecting and enhancing the privacy of individuals and directly compensating those who wish to be part of this incentive.

#### 3.3.1.1 Challenges Faced

**Data availability issues:** At this point, the mandatory personal data that each fan or member needs to provide to the club include name, surname, e-mail address and either the username (for fans) or the club registration number (for members). All other personal data are optional, with the club's employees trying to keep and update as much as possible, e.g. telephone number, postal address, social security number, etc. Additionally, the club uses an online system to offer members the opportunity to register and annually renew their subscription on their own via the official club website ([olympiacossp.gr](http://olympiacossp.gr)). A number of fans are hesitant to give all their data also in this case, either because they are concerned about data protection or simply because the information is not mandatory.

The club does not collect social media profiles at all but has an overview of demographic data as shared by social media sites regarding the people that interact with the profile pages of the club.

The club and the football team both have official branded apps<sup>1</sup> but these do not use geolocation-based functionality.

**Availability of data collected from third parties:** The club uses an external contractor/third party company to manage the online ticket issuing process and the automatic admission to the stadium through electronically operated gates. The contractor collects data related to how many times each fan/member was at the stadium and when/which game they followed, as the online ticketing system is linked to the fan/member card number. However, such data belongs to the external contractor and is not owned by the club.

**Analytics capacity:** The internal CRM used by the club to collect the personal data of fans and members does not offer any analytics-relevant tools for exporting complete reports, making the implementation of successful strategic and marketing plan difficult.

### 3.3.1.2 Main Objectives of Demonstrator's Scenario

The Main Objectives of this scenario are:

- Connect internal CRM of the club to DataVaults cloud platform.
- Share data of fans and members with DataVaults cloud platform.
- Brand the DataVaults personal app, for example as an affiliate of Olympiacos.
- Promote DataVaults personal app to existing fans and members (e.g. via e-mail).
- Advertise and implement compensation (e.g. perks) for connecting the DataVaults app with social media data sources (e.g. Facebook profile, Instagram profile, etc.) based on the features that will be offered by the platform.
- Push questionnaires to fans to extract information in the absence of other data sources.

### 3.3.1.3 Scenario A Evolution Plan

This scenario is split into 3 phases, which are designed in order to gradually reach the full scale, following the alpha (August 2021), beta (December 2021) and v0.50 (June 2022) release of the DataVaults platform. Those are the following:

#### **I. Scenario A - Alpha Phase: Data Sources Connection and Closed Group Data**

**Collection:** This phase covers the initial interaction between the club's internal systems and the alpha version of the DataVaults infrastructure. Therefore, during the alpha phase: the development/set-up of the connector between the internal CRM of the club and the DataVaults infrastructure takes place; the sharing of the existing data of fans and members with DataVaults is initiated; and a closed group evaluation from a handful of members (e.g. employees of the club who are also active fans/member of the club) is performed, for example a handful of club employees could connect their social media accounts to DataVaults and start feeding it with data similar to those we intend to collect.

- Connect internal CRM of the club to DataVaults so that data owners can easily retrieve their data in the Personal App.
- Share data of fans and members with DataVaults.

---

<sup>1</sup> <https://play.google.com/store/apps/details?id=gr.olympiacosfp&hl=en&gl=US> ;  
[https://play.google.com/store/apps/details?id=com.olympiacosapp&hl=en\\_SG](https://play.google.com/store/apps/details?id=com.olympiacosapp&hl=en_SG)



- Stakeholders from the club inspect the collected data at the DataVaults and propose cleansing/transformation rules.
- Collection of data from a handful of members who are already in close contact with the club and could be recruited as very early evaluators, even outside the formal evaluation schedule of the project

## **II. Scenario A - Beta Phase: Extension of Data Owners base and operational readiness testing:**

This phase covers the early operation of the DataVaults solution, both the cloud platform and the personal app. During the beta phase the personal app opens up to greater groups of users, for example all the members, who are by position closer to the proceedings of the club. Since the platform will be functional and a certain amount of personal data have been collected and cleansed during the Alpha phase, during this phase analytics on collected data will be made available for the data seekers. Moreover, a customized (e.g. branded) version of the Personal App will be considered and advertised from the club to the fans and members.

- Open up the use of DataVaults Personal App to greater groups of users.
- Allow Data Seekers to acquire analytics from the DataVaults cloud platform.
- Pilot a branded version of the DataVaults personal app and advertise it to the fans and members as an affiliate of the club.

## **III. Scenario A - Final Phase – Branding and Operational Deployment of the Marketplace:**

This phase covers the full-fledged pilot operation of DataVaults platform related to the activities of Olympiacos (fans and members). During the final phase, there is enough amount of data in the DataVaults cloud platform to elaborate the personas (collective profiles) of, for example, the average season fan or the high-income fan who attends the most prestigious matches. Then, during this phase the club enables the compensation mechanisms (e.g. in the format of perks) for the members & fans who take part in the initiative. Finally, during this phase other Data Seekers will be invited to get on-board, a prominent example of this could be club sponsors who are interested to search through the personal and online activity data of members & fans of the club, in order to offer better customized and targeted products and services.

- Make the branded version of the DataVaults personal app available to a larger group (e.g. upload it to the prominent online app stores, or directly invite individuals).
- Acquire advanced analytics from the DataVaults cloud platform.
- Stakeholders from the club inspect the proposed personas from the DataVaults cloud platform and verify their accuracy using their experience.
- Activate the sharing compensation mechanisms and link it with perks for sharing data from the DataVaults app during the period of the pilot execution.
- Invite club sponsors to use the DataVaults cloud platform as a Data Seeker.

### ***3.3.1.4 Data Availability and Data Needs***

Data Available in the CRM of Olympiacos for each fan and member (mandatory fields marked with a \*):

- Username\*, e-mail\*, active/deactivated.
- Name\*, Surname\*, Father's Name, Sex.
- Birth Date, Birthplace, Nationality, Citizenship, Social Security ID, Job Title.
- Postal Address, Postal Code, City, Region/Prefecture, Country.
- Registration Start Date (renews yearly), Registration End Date (renews yearly), Registration Type (e.g. "Member Yearly"), Member Since (history), Member Number or Fan Number (linked to the fan/member card number).
- Phone Number, Mobile Phone Number.

Data Needs from each fan and member:

- Facebook profile activity.
- Instagram profile activity.
- Location History.
- How many times each fan/member was at the stadium and when/which game they followed, as the online ticketing system is linked to the fan/member card number.
- Updates in personal and contact data (if/once the club's records are outdated).
- Replies in satisfaction surveys.

### 3.3.2 Scenario B - Athletes Sports and Activity Data Sharing

The club has eighteen different sport departments at the competitive level and the Academies which comprise a large base of 2,000 professional and young athletes including important contact and medical details. In this scenario, athletes of the club are encouraged to share their personal data, athletic activity data and ergometric and medical examination data through the DataVaults platform.

For the athletes, awareness on the type and quantity of personal data, athletic activity data and ergometric and medical examination data belonging to one individual combined with the offering of the appropriate medical and sport equipment by the club and coupled with incentives offered from Data Seekers to share such data with the platform shall encourage the athletes to share not only already collected but also future athletic activity and medical data.

For the club, this will help with the better management of the sport departments, the planning of the training sections and team tactics, the opportunity to know what specific physical skills are needed when looking into players' transfers, and finally to cover the athletes' expectations by offering the appropriate medical and sport equipment.

The main goal of this scenario is to collect especially athletic activity data and ergometric and medical examination data of athletes who already belong to the club. Collecting and analysing such data could assist the club since the current portal on the one hand includes name, surname, address, telephone number, e-mail, date joined the club, sports (e.g. tennis, sailing, etc.), but on the other hand it doesn't allow the entry of more complicated data such as the results of the ergometric and medical examinations. The athlete's base should be also complemented by statistical reports regarding the performance of the athletes in training and during the matches. In those areas DataVaults could cover the club's need for accessing more

complicated information and offering better analytics. All this, in a transparent exchange between the athlete and the club, respecting and enhancing the privacy of individuals.

#### 3.3.2.1 Challenges Faced

**Data availability issues:** The folders of the club's athletes are kept in spreadsheets and more recently in an internal portal. Such structures don't allow the entry of more complicated data such as the results of the ergometric and medical examinations.

**Availability of data collected from third party:** The club doesn't keep in the internal systems the athletic activity data coming from external/proprietary sensors (in cases where these are used, for example in the football teams), because the data belong to the manufacturer of the device/sensor.

**Analytics Availability:** The data already possessed don't allow the generation of analytics and statistical reports regarding the performance of the athletes in training and during the matches.

#### 3.3.2.2 Main Objectives of the Scenario

- Connect internal portal of the club to DataVaults to fetch information upon individuals' command.
- Share personal data and athletic activity data of athletes with DataVaults cloud platform.
- Promote DataVaults Personal App to athletes, trying to include a handful of professional athletes.
- Run analytics on the performance of the athletes based on the collected data.
- Use DataVaults platform as the main channel to collect athletic activity and ergometric and medical examination data from the athletes who take part in the initiative.

#### 3.3.2.3 Scenario B Evolution Plan

This scenario is split into 3 phases, which are designed in order to gradually reach the full scale. As this scenario is considered to be of a lower scale than the previous one, the alpha phase will be realized once the beta release of the platform is ready (December 2021), and the next two phases will be executed after the release v0.50 of the platform (June 2022). Those are the following:

##### **I. Scenario B - Alpha Phase: Implementation of Data Migration Connectors and Sharing**

**Testing:** This phase covers the initial interaction between the club's internal systems and DataVaults. Therefore, during the alpha phase the focus is on gathering personal, athletic activity and ergometric and medical examination data from the athletes who already belong to the club to the DataVaults cloud platform. Connecting the two could mean the need to implement custom scripting for one-off migration of data kept in spreadsheets or folders, the transfer of which cannot be automated and then create the connectors to allow data owners to retrieve them and share them through the Personal App.

- Transform static isolated data into data sources.
- Connect data sources to DataVaults.
- Testing the features for share personal data and athletic activity data of athletes with DataVaults cloud platform.

- Stakeholders from the club inspect the collected data at the DataVaults cloud platform and propose cleansing/transformation rules.

**II. Scenario B - Beta Phase: Onboard a Small Group of Athletes:** During the beta phase the personal app is available to the athletes. The DataVaults cloud platform is now functional and a certain amount of personal, athletic activity and ergometric and medical examination data from the athletes have been already collected and cleansed during the alpha phase. The focus is on on-boarding some (professional) athletes and showcase the creation of basic analytics on collected data.

- Showcase basic analytics extracted with the help of the DataVaults cloud platform (for those athletes who were already on-boarded during the alpha phase).
- Recruit (professional) athletes to enrol in the platform and share athletic activity data and ergometric data with it (perhaps data already collected using equipment that belongs to the club).

**III. Scenario B - Final Phase: Introduction of the App to the club's athlete's community:**

This phase covers the full-fledged pilot operation of DataVaults platform related to the activities of Olympiacos (athletes). During the final phase, a good proportion of the club's athletes starts using the DataVaults personal app. Then, during this phase the club directly uses the DataVaults cloud platform as a solution to collect athletic activity and ergometric and medical examination data from the athletes who take part in the initiative. Finally, during this phase other Data Seekers are on-boarded, a prominent example of this could be athletic equipment companies who are interested to search through the athletic activity data of athletes of the club, in order to offer better customized and targeted products and services (e.g. depending on what sport each athlete does).

- Make the branded version of the DataVaults personal app available to everyone.
- Acquire analytics from the DataVaults cloud platform.
- Use DataVaults cloud platform as the main channel to collect athletic activity and ergometric and medical examination data from the athletes who take part in the initiative.
- Invite club sponsors to use the DataVaults cloud platform as a Data Seeker.

#### **3.3.2.4 *Data Availability and Data Needs***

Data Available for each athlete:

- Name, Surname
- Address
- Telephone number, e-mail
- Date joined the club
- Sports that the athlete does (e.g. tennis, sailing, etc.)
- Other data in spreadsheets

Data Needs from each athlete:

- Athletic activity data
- Ergometric and medical examination data
- Updates in personal and contact data (in case the club's records are outdated)
- Replies in satisfaction surveys

---

### 3.4 DEMONSTRATOR TECHNICAL DEVELOPMENT NEEDS

---

In order to utilise DataVaults for the needs of the Olympiacos demonstrator and in the context of the two different scenarios presented above, the following technical needs arise.

From the development team(s) of DataVaults:

- (for Scenarios A+B) Deploy DataVaults cloud platform instance.
- (for Scenario A) Create a connector between the internal CRM of the club and DataVaults cloud platform instance.
- (for Scenario B) Transform static isolated data into data sources (e.g. custom migration script from the club's records of athletes in MS Excel to machine readable format, e.g. JSON files) and connect those data sources to DataVaults cloud platform instance.
- (for Scenarios A+B) Help with technical issues during the transfer of personal data to the DataVaults cloud platform instance.
- (for Scenarios A+B) Implement cleansing/transformation rules on collected data (both from members & fans and from athletes), as proposed by club's employees.
- (for Scenarios A+B) Brand the DataVaults personal app, according to the needs of Olympiacos demonstrator.

---

### 3.5 COMPENSATION METHODS CONSIDERED

---

The list below provides a first view on the compensation methods that are considered by the club. These will be concretized after the Alpha phase.

Types of compensation considered:

- bonuses (discount coupons) on new official jerseys of our multisport club,
- benefits from our sponsor area partners, and
- discount coupons from our club and from other partners.

---

### 3.6 DEMONSTRATOR'S ETHICS CONSIDERATIONS

---

Participation of members, fans and athletes is completely on a voluntary basis, with clear rights to privacy and safety (informed consent).

Confidentiality and Anonymity guarantees for the members, fans and athletes who wish to part in the initiative will be provided, according to the functions of the platform.

Communicating Results on the use of DataVaults platform to the members, fans and athletes who took part in the initiative, as well as to the management of the club will be also performed, so that it is transparent on how data are used.

The DataVaults Ethical Policy outlined in D9.2 will be followed, including the envisaged Ethics and Data Protection Impact Assessment Methodology, as well as the indications of the project's Ethics Advisory Board.

Further information on the ethics dimension can be retrieved in D10.1 and D10.2: the former describes the informed consent procedure that will be adopted during the pilot actions and includes the templates of the informed consent and information sheet, the latter describes the procedures for data collection, storage, protection, retention, and destruction and the ethics risk evaluation

### 3.7 IMPACTS EXPECTED

In this section we describe the impact expected for the demonstrator based on the two scenarios identified above. The measurement of the actual impact will be provided in deliverable D6.5.

Scenario	KPI	Calculation Method	AS-IS Value	TO-BE Value	Verification Means
<b>A</b>	Effective management of members & fans data	# of members & fans who are "satisfied" and "very satisfied" with the club's methods of handling personal data / # of members & fans who completed the satisfaction survey	50%	75%	Satisfaction surveys completed by members & fans who participate in DataVaults initiative
<b>A</b>	Number of registered club members	Number of members who have renewed their membership with the club for the current season	130,000	135,000	Members database count
<b>A</b>	Number of active club members in the club's activities	Number of members who have participated in the club's advertised activities	20,000	22,000	In-house analytics (e.g. records of members who took part in the annual general assembly of the club)
<b>A</b>	Details about active fans in the club's activities	Number of fans who have followed the club's games in more than one sports department	60,000	70,000	In-house analytics
<b>B</b>	Effective management of athletes'	# of athletes who are "satisfied" and "very satisfied"	60%	75%	Satisfaction surveys completed by athletes who

Scenario	KPI	Calculation Method	AS-IS Value	TO-BE Value	Verification Means
	sports and activity data	with the club's methods of handling personal data / # of athletes who completed the satisfaction survey			participate in DataVaults initiative
<b>A+B</b>	Stakeholder Trust	# of club's stakeholders who are "satisfied" and "very satisfied" with the club's participation in the DataVaults initiative / # of club's stakeholders who completed the satisfaction survey	65%	80%	Satisfaction surveys completed by club's stakeholders

Table 1: Impact Target KPIs for Demonstrator #1

### 3.7.1.1 Impact for the Organisation

The expected inputs for this demonstrator are to be seen in the following areas:

Effective management of members & fans data: Olympiacos has a large base of members and fans which requires a better data management including collection, mining processing, security, and cryptography. It's very important for the club to secure the personal data protection in the best way to ensure the trust of the individuals. Moreover, as the club develops a large range of innovative and high-quality marketing campaigns, it should emphasize on the collection of valuable insights without violating privacy principles. This database is also beneficial for the planning and the implementation of segmented and targeted marketing campaigns, the improvement of fan engagement and finding sponsors. For the club, it would be extremely beneficial to collect information such as user preferences, activity, and mobility data in the future and focus on the real preferences of the members and fans and improve the services and the fan engagement to ensure their loyalty.

Number of registered club members: Olympiacos SFP has a large base of 130,000 registered members and 60,000 registered fans. This base is growing rapidly which makes the club the largest sport one in Greece and one of the most popular worldwide. Using the CRM system, the club has built a strong relationship with the member-fan base which seems to be an important issue for the sport successes.

Number of active club members in the club's activities: Number of members who have participated in the club's advertised activities. The data included in the base of members are compulsory to organize and coordinate procedures such as the organization and the

participation of the members in the club General Assembly, the right to vote in the Administrative Elections and take part in the decision-making.

Number of active fans in the club's activities: Number of fans who have followed the club's games in more than one sports department. This base and the satisfaction of fans are very important issues for the club as they are the most important source of revenues.

Effective management of athletes' sports and activity data: Olympiacos has a large base of athletes which requires a better data management including collection, mining processing, security, and cryptography. The sport departments and the Academies have more than 2,000 athletes on their rosters. It proves that Olympiacos is the most appropriate instance for the implementation and the usage of the platform as it will encourage the athletes to upload their personal data on the platform. The athletes' base requires better management of the results of the ergometric and medical examinations and statistical reports regarding their performances in the training and the matches. This base also includes personal information, demographics, personal governmental data, and sport exercise/activity one. Better management is very crucial for the club to adapt and plan the training sections and the team tactics, make players' transfers with specific physical skills, cover the athletes' expectations offering the appropriate medical and sport equipment.

Revenue from sponsorships: The club organizes meetings with the sponsors on a regular basis to discuss the progress of the sponsorship deals and control the sponsor satisfaction. These meetings could include a presentation of the platform and its utility to motivate the sponsors to join it.

Stakeholder Trust: The club will encourage the members, fans, and athletes to capture, witness and manage their personal data through this platform and undertake the decision to share them responding to the requests that may come from different organizations. The club will make clear that the platform ensures the data protection following modern encryption techniques to guarantee privacy and total information security. The club will inform the stakeholders about the possibilities that the platform offers such as full or partial data completeness, encrypted or unencrypted data security, data privacy etc.

#### *3.7.1.2 Impact for the Data Owners*

Effective management of members and fans data: In this case, an important issue is that a number of fans are hesitant to give all their data because they concern about data protection.

Effective management of athletes' sports and activity data: The data collection of the athletes is concerned, this is a manual procedure via Microsoft Office platforms (e.g., Excel, Word). However, the club has also started to use the above CRM tool for the data management of the athletes. A critical issue is that the version of the CRM that the club uses, doesn't allow the entry of more complicated data such as the results of the ergometric and medical examinations. So, there is a need for an appropriate platform that facilitates the entry of more complicated information.

#### *3.7.1.3 Impact for Collaborators (Data Seekers)*

Stakeholder Trust: Olympiacos aims to adopt specific services that the project includes such as:



- Holistic personal data management services, including collection, mining processing, normalization, appropriate formatting, and availability at individuals' personal devices level.
- Privacy preserving and data security retention mechanisms, to accommodate the generation of anonymized "digital twins" of individuals, as well as specimen clusters ("persona groups") powering group analytics that contain valuable insights without violating privacy principles.
- Smart balancing of analytics methods to accommodate Edge analytics as well as centralized operations depending on the degree of data volume, velocity, and variety, always in conjunction with the security and privacy modalities allowed by the individual for each kind of analysis.
- Provision of intuitive analytics, reports, smart dashboards, and visualizations tailored to the needs of each stakeholder of the domain, including the individual, as well as generic ones for wider use by any interested organization and by the public.

It needs to adopt the platform usage and promote it to the key stakeholders (members, fans, athletes, sponsors) to develop a stronger relationship with them and encourage them to care about the sharing, availability and updating of their data (especially those are related to social media activity and the preferences) securing their privacy. This would assist the club to re-use the current data and use new ones to engage and interact with the key stakeholders better and reorganize its strategic plan (new market segmentation, marketing campaigns for specific target groups, finding specific sponsors etc.) aiming to the improvement of the services and the increase of the revenues.

### 3.8 DEMONSTRATOR'S ACTIVITIES TIMELINE

The following table presents the main activities to be performed within the demonstrator towards realising the two scenarios described above.

Demonstrator 1 OLYMPIACOS	M19	M20	M21	M22	M23	M24	M25	M26	M27	M28	M29	M30	M31	M32	M33	M34	M35	M36
<b>Scenario A - Club Fans and Members Personal Data Marketplace</b>																		
<b>Alpha Phase</b>																		
Connect internal CRM to DataVaults																		
Share data of fans and members																		
Collect data from a few of early adopters																		
Club stakeholders inspect data																		
<b>Beta Phase</b>																		

Demonstrator 1 OLYMPIACOS	M19	M20	M21	M22	M23	M24	M25	M26	M27	M28	M29	M30	M31	M32	M33	M34	M35	M36
Open up the use of DataVaults personal app																		
Acquire analytics from the DataVaults cloud platform																		
Pilot a branded version of the DataVaults personal app																		
<b>Final Phase</b>																		
Branded DataVaults app available to everyone																		
Activate the sharing compensation mechanisms																		
Acquire analytics from the DataVaults cloud platform																		
Club stakeholders verify personas																		
Invite club sponsors																		
<b>Scenario B - Athletes Sports and Activity Data Sharing</b>																		
<b>Alpha Phase</b>																		
Transform static isolated data																		
Connect data sources to DataVaults																		
Test sharing features																		
Club stakeholders inspect data																		
<b>Beta Phase</b>																		
Showcase basic analytics extracted with DataVaults																		
Recruit (professional) athletes																		
<b>Final phase</b>																		
Branded DataVaults app																		

<b>Demonstrator 1 OLYMPIACOS</b>	<b>M19</b>	<b>M20</b>	<b>M21</b>	<b>M22</b>	<b>M23</b>	<b>M24</b>	<b>M25</b>	<b>M26</b>	<b>M27</b>	<b>M28</b>	<b>M29</b>	<b>M30</b>	<b>M31</b>	<b>M32</b>	<b>M33</b>	<b>M34</b>	<b>M35</b>	<b>M36</b>
available to everyone																		
Acquire analytics from the DataVaults cloud platform																		
Showcase acquisition of athletic activity and ergometric and medical examination data																		
Invite club sponsors																		

Table 2: Execution Timeline for Demonstrator #1

## 4 DEMONSTRATOR #2 - STRENGTHENING ENTREPRENEURSHIP AND MOBILITY

### 4.1 BUSINESS DESCRIPTION AND THE NEED FOR DATAVAULTS

Piraeus is the largest & busiest port in Greece, among the biggest European ports and the main hub connecting Europe, Asia and Africa. Piraeus is the maritime and former industrial centre of the Athens metropolitan area, and one of the most densely populated cities in Europe (over 180.000 inhabitants, 15.065 citizens/km<sup>2</sup>) based on the 2011 population report. Apart from commercial trading, the traffic of cruise passengers is very remarkable (over 12 million travellers in 2018), causing equally heavy traffic in the road network of the Piraeus centre (40.350 daily moves).

In Piraeus there are important monuments, such as a part of the Long Walls of the 5th century BC, remains of ancient temples and buildings. There are also the neoclassical Municipal Theatre, industrial buildings of the early 20th century and impressive modern buildings such as the Peace & Friendship Stadium and the G. Karaiskakis Stadium. Moreover, there are several museums with the most important the Archaeological and the Hellenic Maritime Museum. Furthermore, the city boasts the Greek multisport clubs, Olympiacos CFP which is also a demonstrator partner in DataVaults.

Aside from its importance as a tourist destination and a starting point for hoping onto cruises, Piraeus has one of the biggest commercial areas in Greece, having a large market that is based both on local citizens, as well as on tourists.

However, the recent financial crisis has also hit this market, resulting in a decrease of new business ventures of around 81,6% compared to the growth witnessed 10 years ago. In order to counterbalance this deficit, and to restart the local economy, the city of Piraeus got engaged in the Digital Strategy actions go the Greek State, while an Integrated Territorial Investment program of the amount of 6M€ for Information Technology actions has been issued, to make the city more attractive both to citizens and visitors, offering amongst others novel ICT services. One of those is the smart phone application PireApp<sup>2</sup>, aiming to boost citizen engagement. Through the use of this app, citizens and visitors are able to get in direct communication with the local authorities and report problems associated to urban the environment, while at the same time a database of citizens' profiles is generated, helping the city to better identify its citizens and their needs.

However, the information recorded through the app is still limited and does not allow the municipality to perform complex analyses that could provide insights to help the city understand and support local entrepreneurship and mobility.

Through DataVaults, PIRAEUS aspires to renovate the way it collects, manages and analyses data, moving from simple information provision and limited feedback solutions - such as PireApp - towards embracing novel citizens and visitors' engagement practices. The DataVaults app will allow the targeted stakeholders to provide more and richer data to the

---

<sup>2</sup> <https://pireapp.novoville.com/>

city, entailing more personalized services, always in a trusted and secure manner, respecting their privacy.

---

## 4.2 TARGET AUDIENCE DESCRIPTION – DATA OWNERS AND DATA SEEKER

---

### 4.2.1 Data Owners

**Citizens and tourists**, who own and continuously generate personal data, social activity data and location data, usually without understanding what it comprises, or having any means to manage and share such data.

The principal type of data envisaged to be collected from citizens and tourists in the context of the current demonstrator are:

- Facebook profile activity
- Instagram profile activity
- Location History
- Interests and hobbies
- Purchases and commercial trends

---

### 4.2.2 Data Seekers

The envisaged Data Seekers, besides the Municipality of Piraeus, shall be contacted by the Municipality either in the final phase or at the post-project phase. Possible Data Seekers include the Piraeus Chamber of Commerce & Industry and the local Destination Management Office (an organization under the Municipality of Piraeus) and other local companies operating in the area of Piraeus.

---

### 4.2.3 Means to Reach other Data Seekers

Reaching other Data Seekers will be mainly done by direct invitation by PIRAEUS

---

## 4.3 DEMONSTRATOR SCENARIOS

---

### 4.3.1 Scenario A - Smart Mobility Services for Individuals

Smart Mobility Services for Individuals. This scenario will engage both OLYMPIACOS and PIRAEUS and will use the data shared by the interested citizens, as well as by the members taking part in the OLYMPIACOS demonstrator, to better schedule the mobility strategy and the relevant services within the city. The specific area of interest during the course of the project will be the surroundings of the OLYMPIACOS sport venues.

The main goal in this scenario is to collect all relevant information from citizens and fans, especially routes of motion around the sport venues and information of attendance to specific sport events. The next step is to establish through data analytics the best road traffic diversions in order to avoid traffic congestion around the sport venues and the entrance roads towards the city of Piraeus.

#### 4.3.1.1 Challenges Faced

**Data availability:** All routing and location information must be collected from Olympiacos fans and citizens of the city of Piraeus. Moreover, attendance information is to be provided also from the fans. This information is not currently available and could be collected through the use of a DataVaults app.

**Analytics capacity:** The Municipality of Piraeus does not currently analyse data for traffic management nor possesses any specialized tool to perform such task. However, the Municipality is currently in the process to formulate a sustainable urban mobility plan and during this process both IT tools and experience are to be obtained.

#### 4.3.1.2 Main Objectives of Demonstrator's Scenario

- Setup the DataVaults personal app to contain the necessary for this scenario information.
- Extract data of citizens from existing apps (e.g., PireApp) to be used for promotional purposes, always following the agreement/instruction of the user to retrieve their data.
- Brand the DataVaults personal app, to be able to promote it through the Municipality of Piraeus channels.
- Promote DataVaults personal app to citizens (e.g., via e-mail or through the website).
- Advertise and implement compensation (e.g., perks) for connecting the DataVaults app with social media data sources (e.g., Facebook profile, Instagram profile, etc.) based on the features that will be offered by the platform.
- Push questionnaires to citizens and fans in order to extract information in the absence of other data sources.

#### 4.3.1.3 Scenario A Evolution Plan

This scenario is split into 3 phases, which are designed in order to gradually reach the full scale, following the alpha (August 2021), beta (December 2021) and v0.50 (June 2022) release of the DataVaults platform. Those are the following:

**I. Scenario A - Alpha Phase: Creation of closed group of people for evaluation:** This phase covers the initial interaction between the Municipality of Piraeus and the citizens. Therefore, during the alpha phase: an initial list of citizens is to be formulated, using contact information from the Piraeus application PireApp; a closed group evaluation from a handful of people (e.g., employees of the Municipality, a small random group of citizens, a small group of Olympiacos employees) is performed.

Actions to be taken:

- Collect citizen's data from PireApp
- Form initial group of users in collaboration with Olympiacos (Municipality users, Olympiacos employees, small group of citizens from PireApp)
- Collection of data from the above group outside the formal evaluation schedule of the project
- Examine and analyse this initial data to evaluate the possibility of taking action on them

**II. Scenario A - Beta Phase: Extension of user base** This phase is a preparatory phase towards the full pilot operation of DataVaults platform related to the activities of for Smart Mobility in the Municipality of Piraeus. During this phase the DataVaults personal app will be opened up to a greater number of users.

Actions to be taken:

- Open up the use of DataVaults personal app to greater groups of users.
- Acquire analytics from the DataVaults cloud platform.

**III. Scenario A - Final Phase: Full pilot operation:** This phase covers the full pilot operation of DataVaults platform related to the activities of for Smart Mobility in the Municipality of Piraeus. During the final phase, the DataVaults app should have collected enough data in the DataVaults platform to elaborate on possible actions to be taken regarding rerouting traffic during sport events. During this phase the Municipality will enable the compensation mechanisms e.g., in the form of discounts for Municipal services (theatre, training grounds, swimming pool) for the citizens taking part in the initiative.

Actions to be taken:

- Make the branded version of the DataVaults personal app available to everyone (e.g., upload it to the prominent online app stores).
- Acquire advanced analytics from the DataVaults cloud platform and propose actions for smart mobility.
- Activate the sharing compensation mechanisms and link it with perks for sharing data from the DataVaults app during the period of the pilot execution.

#### 4.3.1.4 *Data Availability and Data Needs*

Data Available from Olympiacos and PireApp (mandatory fields marked with a \*):

- Username\*, e-mail\*, active/deactivated
- Name, Surname, Father's Name, Sex
- Postal Address, Postal Code, City, Region/Prefecture, Country
- Mobile Phone Number

Data Needs from citizens and Olympiacos members:

- Location History (routes taken)
- Ticket purchase or season ticket holder
- Updates in personal and contact data (if/once the club's records are outdated)
- Replies in satisfaction surveys

---

#### 4.3.2 Scenario B - Empowering local entrepreneurship

In this scenario, the data to be provided by the DataVaults users will be used to better understand consumer behaviours and preferences, with the aim to strengthen the local economy through activities that can be brought forward by the municipality. Moreover,

PIRAEUS will invite local entrepreneurship associations (i.e., the Piraeus Traders Association) and other interested stakeholders to either join the platform or act as 2<sup>nd</sup> tier data seekers, to test the aspects of the project that have to do with value generation and sharing with entities not directly using personal data but that access the derivatives of the latter.

The main goal in this scenario is to collect all relevant information from citizens and tourists, especially commercial interests and preferences. The next step is to establish through data analytics personalized promotional actions to enhance the local commerce and entrepreneurship in the city of Piraeus.

#### *4.3.2.1 Challenges Faced*

**Data availability:** All personal information has to be collected from tourists and citizens of the city of Piraeus. This information is not currently available and could be collected through the use of the DataVaults app.

The data collection will be initiated by the citizen/tourist themselves, by turning on the relevant connectors to DataVaults, providing to the latter the commands to retrieve these data. However, the challenge here is to persuade the individuals to do so, and to build the connectors with other systems, that might not be open or might not offer APIs for user's data retrieval.

**Analytics capacity:** The Municipality of Piraeus does not currently analyse data for commercial purposes or possesses any specialized tool to perform such task. However, the Municipality is in collaboration with the Piraeus Traders Association, an organization capable of performing such an analysis.

#### *4.3.2.2 Main Objectives of the Scenario*

- Setup the DataVaults personal app to contain the necessary for this scenario information.
- Extract lists of citizens from existing apps (e.g., PireApp or an upcoming touristic application) to be used for promotional purposes.
- Brand the DataVaults personal app, to be able to promote it through the Municipality of Piraeus channels.
- Promote DataVaults personal app to citizens (e.g., via e-mail or through the website) and tourists (e.g., via posters and leaflets at the Cruise Terminal).
- Advertise and implement compensation (e.g., perks) for connecting the DataVaults app with social media data sources (e.g., Facebook profile, Instagram profile, etc.) based on the features that will be offered by the platform.
- Push questionnaires to citizens and tourists in order to extract information in the absence of other data sources.

#### *4.3.2.3 Scenario B Evolution Plan*

This scenario is split into 3 phases, which are designed in order to gradually reach the full scale, following the alpha (August 2021), beta (December 2021) and v0.50 (June 2022) release of the DataVaults platform. Those are the following:



**I. Scenario B - Alpha Phase: Creation of closed group of people for evaluation** This phase covers the initial interaction between the Municipality of Piraeus and the citizens/tourists. Therefore, during the alpha phase: an initial list of citizens is to be formulated, using contact information from the Piraeus application PireApp; a closed group evaluation from a handful of people (e.g., employees of the Municipality, a small random group of citizens) is performed. Also, in this phase the initial engagement of the Piraeus Trade Association is foreseen.

Actions to be taken:

- Collect citizen's data from PireApp.
- Form initial group of users (Municipality users, small group of citizens from PireApp).
- Collection of data from the above group of outside the formal evaluation schedule of the project.
- Piraeus Trade Association initial engagement.
- Examine and analyse this initial data to evaluate possibility of forming a promotional strategy.

**II. Scenario B - Beta Phase: Extension of user base** This phase is a preparatory phase towards the full pilot operation of DataVaults platform related to the activities of empowering local entrepreneurship in the Municipality of Piraeus. During this phase the DataVaults personal app will be opened up to a greater number of users.

Actions to be taken:

- Open up the use of DataVaults personal app to greater groups of users
- Acquire analytics from the DataVaults cloud platform

**III. Scenario B - Final Phase: Full pilot operation** This phase covers the full pilot operation of DataVaults platform related to the activities of empowering local entrepreneurship in the Municipality of Piraeus. During the final phase, the DataVaults app should collect enough data in the DataVaults platform to elaborate on possible actions to empower local entrepreneurship. During this phase the Municipality in cooperation with Piraeus Trade Association will enable the compensation mechanisms e.g., in the form of discounts for Municipal services (theatre, training grounds, swimming pool) or other perks from local traders (discounts, gifts) for the citizens taking part in the initiative.

Actions to be taken:

- Make the branded version of the DataVaults personal app available to everyone (e.g., upload it to the prominent online app stores).
- Acquire advanced analytics from the DataVaults cloud platform and propose actions to empower local entrepreneurship.
- Activate the sharing compensation mechanisms and link it with perks for sharing data from the DataVaults app during the period of the pilot execution.

#### 4.3.2.4 Data Availability and Data Needs

Data Available from PireApp (mandatory fields marked with a \*):

- Username\*, e-mail\*, active/deactivated
- Mobile Phone Number

Data Needs from citizens and tourists:

- Facebook profile activity
- Instagram profile activity
- Online purchases
- Commercial preferences and interests
- Updates in personal and contact data (if/once the club's records are outdated)
- Replies in satisfaction surveys

---

#### 4.3.3 Scenario C - Services for Personalized cultural and touristic experiences

Services for Personalized cultural and touristic experiences. This scenario will build on data analysed from the profiles and preferences of the DataVaults app users, in order to create services that target tourists and citizens visiting the city of Piraeus. During this scenario, the data to be analysed will generate reports that will assist the departments of the municipality to better design their strategies regarding the services offered to meet the touristic and cultural event demand. This scenario is both aligned and complementary to the Digital Strategy 2020 of PIRAEUS in terms of implementing an integrated Destination Management System, engaging citizens and visitors in the interactive definition of the cultural content of interest through the analysis of public (i.e., museums & touristic organizations) and private (i.e. travel agencies, cruise operators, booking organizations, etc.) data sources.

##### 4.3.3.1 Challenges Faced

**Data availability:** All personal information has to be collected from tourists and citizens of the city of Piraeus. This information is not currently available and could probably be collected through the use of the DataVaults app. A tourist app is currently under procurement for the city of Piraeus. If in time, this app could also contribute with tourist information and contact details.

**Analytics capacity:** The Municipality of Piraeus does not currently analyse data for touristic purposes or possesses any specialized tool to perform such task. However, such tools and specialized personnel hiring is under procurement and will possibly be in use in the second phase of this scenario.

##### 4.3.3.2 Main Objectives of the Scenario

- Setup the DataVaults personal app to contain the necessary for this scenario information.
- Extract lists of citizens from existing apps (e.g., PireApp or the upcoming touristic application) to be used for personalized cultural and touristic experiences.
- Brand the DataVaults personal app, to be able to promote it through the Municipality of Piraeus channels.
- Promote DataVaults personal app to citizens (e.g., via e-mail or through the website) and tourists (e.g., via posters and leaflets at the Cruise Terminal).

- Advertise and implement compensation (e.g., perks) for connecting the DataVaults app with social media data sources (e.g., Facebook profile, Instagram profile, etc.) based on the features that will be offered by the platform.
- Push questionnaires to citizens and tourists in order to extract information in the absence of other data sources.

#### 4.3.3.3 *Scenario C Evolution Plan*

This scenario is split into 3 phases, which are designed in order to gradually reach the full scale, following the alpha (August 2021), beta (December 2021) and v0.50 (June 2022) release of the DataVaults platform. Those are the following:

**I. Scenario C - Alpha Phase: Creation of closed group of people for evaluation** This phase covers the initial interaction between the Municipality of Piraeus and the citizens/tourists. Therefore, during the alpha phase: an initial list of citizens is to be formulated, using contact information from the Piraeus application PireApp; a closed group evaluation from a handful of people (e.g., employees of the Municipality, a small random group of citizens) is performed.

Actions to be taken:

- Collect citizen's data from PireApp.
- Form initial group of users (Municipality users, small group of citizens from PireApp).
- Collection of data from the above group of outside the formal evaluation schedule of the project.
- Examine and analyse this initial data to evaluate possibility of forming personalized cultural and touristic experiences proposals.

**II. Scenario C - Beta Phase: Extension of user base** This phase is a preparatory phase towards the full pilot operation of DataVaults platform related to the activities of personalized cultural and touristic experiences in the Municipality of Piraeus. During this phase the DataVaults personal app will be opened up to a greater number of users.

Actions to be taken:

- Open up the use of DataVaults personal app to greater groups of users.
- Acquire analytics from the DataVaults cloud platform.

**III. Scenario C - Final Phase: Full pilot operation** This phase covers the full pilot operation of DataVaults platform related to the activities of personalized cultural and touristic experiences in the Municipality of Piraeus. During the final phase, the DataVaults app should collect enough data in the DataVaults platform to elaborate on possible personalized cultural and touristic experiences proposals. During this phase the Municipality will enable the compensation mechanisms e.g. in the form of discounts for Municipal services (theatre, training grounds, swimming pool).

Actions to be taken:

- Make the branded version of the DataVaults personal app available to everyone (e.g., upload it to the prominent online app stores).

- Acquire advanced analytics from the DataVaults cloud platform and propose personalized cultural and touristic experiences.
- Activate the sharing compensation mechanisms and link it with perks for sharing data from the DataVaults app during the period of the pilot execution.

#### 4.3.3.4 Data Availability and Data Needs

Data Available from PireApp (mandatory fields marked with a \*):

- Username\*, e-mail\*, active/deactivated
- Mobile Phone Number

Data Needs from citizens and tourists:

- Facebook profile activity.
- Instagram profile activity.
- Touristic and cultural preferences and interests.
- Updates in personal and contact data (if/once the club's records are outdated).
- Replies in satisfaction surveys.

---

## 4.4 DEMONSTRATOR TECHNICAL DEVELOPMENT NEEDS

The need for the pilot scenarios is the development of a data repository and API in order to feed information into the DataVaults app. Personal data coming from different data sources, like excel sheets or the PireApp application can then be stored in a structured way and be fed directly to the DataVaults platform.

---

## 4.5 COMPENSATION METHODS CONSIDERED

The list below provides a first view on the compensation methods that are considered by the Municipality of Piraeus. These will be concretized after the Alpha phase.

Types of compensation considered:

- discount on Municipal sport grounds (track, swimming pool).
- discount on Municipal Theatre tickets.
- discount coupons from local traders (through the Piraeus Trade Association).

---

## 4.6 DEMONSTRATOR'S ETHICS CONSIDERATIONS

Participation of citizens and tourists is completely on a voluntary basis, with clear rights to privacy and safety (informed consent).

Confidentiality and Anonymity guarantees for the citizens and tourists who wish to part in the initiative will be provided, according to the functions of the platform. It is feasible for the purposes of all three pilot actions.

Communicating Results on the use of DataVaults platform to citizens and tourists who took part in the initiative, as well as to the management of the club will be also shared, so that it is transparent on how data are used.

The DataVaults Ethical Policy outlined in D9.2 will be followed, including the envisaged Ethics and Data Protection Impact Assessment Methodology, as well as the indications of the project's Ethics Advisory Board.

Further information on the ethics dimension can be retrieved in D10.1 and D10.2: the former describes the informed consent procedure that will be adopted during the pilot actions and includes the templates of the informed consent and information sheet, the latter describes the procedures for data collection, storage, protection, retention, and destruction and the ethics risk evaluation

#### 4.7 IMPACTS EXPECTED

in this section the impact expected is described. The impact assessment will be measured in deliverable D6.5.

Scenario	KPI	Calculation Method	AS-IS Value	TO-BE Value	Verification Means
<b>A</b>	Decrease in time required to reach the sports venue	Decrease in time required to reach the sports venue as percentage	To be measured before the Alpha phase	10% decrease	Satisfaction surveys completed by citizens and Olympiacos members
<b>A</b>	Decrease in time to park around the sports venue	Decrease in time to park around the sports venue as percentage	To be measured before the Alpha phase	10% decrease	Satisfaction surveys completed by citizens and Olympiacos members
<b>B</b>	Number of entrepreneurs involved	To be extracted from the platform	0	20	Platform statistics
<b>B</b>	Number of shared datasets	To be extracted from the platform	0	50	Platform statistics
<b>B</b>	Increase in the number of customers entering the local stores	Increase in the number of customers entering the local stores as percentage	To be measured before the Alpha phase	10% increase	Satisfaction surveys completed by local shop owners
<b>B</b>	Increase in revenues of the local shops participating in the pilot	Increase in revenues of the local shops participating in the pilot as percentage	To be measured before the Alpha phase	10% increase	Satisfaction surveys completed by local shop owners
<b>C</b>	Number of tourists and citizens participating	To be extracted from the platform	0	200	Platform statistics
<b>C</b>	Number of data analysis procedures	Number of data analysis procedures run by the local Destination Management Organization	0	3	DMO report
<b>C</b>	Number of actions taken by the local Destination	Number of actions taken by the local Destination Management	0	5	DMO report

	Management Organization based on DataVaults data	Organization based on DataVaults data			
--	--	---------------------------------------	--	--	--

Table 3: Impact Target KPIs for Demonstrator #2

#### 4.7.1.1 Impact for the Organisation

Scenario A: The impact for the Municipality of Piraeus for scenario A, if the above KPIs are met, is the overall satisfaction of the citizens and the improvement of their quality of life.

Scenario B: The impact for the Municipality of Piraeus for scenario B, if the above KPIs are met, is the overall satisfaction of the local entrepreneurs. Another effect is the increase of the Municipalities revenues as a percentage (tax) of the increases shops revenue.

Scenario C: The impact for the Municipality of Piraeus for scenario is the overall satisfaction of tourists and an increase in the local attractions visitors.

#### 4.7.1.2 Impact for the Data Owners

The main impact for the Data Owners will be given by the possibility of managing all personal data in the same framework in the full respect of privacy issues and control of data sharing procedures, together with the possibility of getting a compensation for data sharing.

#### 4.7.1.3 Impact for Collaborators (Data Seekers)

The main impact on collaborators is the ability provided by the DataVaults tools to the search for personal data in an easier manner and without the need to pay large amounts of money to private big data players. This will result also in an improvement in data analytics procedures for more accurate and effective planning.

## 4.8 DEMONSTRATOR'S ACTIVITIES TIMELINE

The following table presents the main activities to be performed within the demonstrator towards realising the scenarios described above.

Demonstrator 2 PIRAEUS	M19	M20	M21	M22	M23	M24	M25	M26	M27	M28	M29	M30	M31	M32	M33	M34	M35	M36
<b>Scenario A - - Smart Mobility Services for Individuals</b>																		
<b>Alpha Phase</b>																		
Collect citizen's data from PireApp																		
Form initial group of users																		
Collection of data																		
Examine and analyse data																		
<b>Beta Phase</b>																		
Open up the use of DataVaults																		

[illegible]

[illegible]

Table 4: Execution Timeline for Demonstrator #2



## 5 DEMONSTRATOR #3 - SECURE HEALTHCARE DATA RETENTION AND SHARING

### 5.1 BUSINESS DESCRIPTION AND THE NEED FOR DATAVAULTS

A7 Software is a Belgian eHealth company, founded in 2014 by Vincent Keunen. It was created to significantly improve communication between doctors and patients and empower them to work effectively together. Our team of software engineers and medical specialists is focused on improving healthcare for everyone. The realization of this common object led to the development of Andaman7.

Andaman7 is a patient centric mobile tool for collaborative management of medical records between health professionals, patients, authorized relatives of such patients (parents or spouse, for example) and connected devices. Exchange and synchronization of information are Peer to Peer, based on original concepts maximizing security and confidentiality.

With the Andaman7 app, patients can

- feed their medical records by importing health data: The platform captures data from EMRs, sensors, wearable and QoL questionnaire data
- access their data in a structured, secure and standardized way and consult them anytime, anywhere;
- share their file with authorized third parties: general practitioners, specialists, or family members.

But Andaman7 is also much more than “a simple mobile app”. It is advanced health IT software combining mobile applications, server processing, smart peers and a full HIP - Health Interoperability Platform - for the easy and efficient distribution of data, while preserving the privacy of patients and the security of the overall system.

Our solution touches on a universal problem: human health and our Business Model is based on a wide distribution of the application in "Freemium". As such, revenues come from paid advertising and add-ons (including Clinical Studies apps).

Our application is currently well differentiated and complementary to existing solutions. We do not want to be profiled as a competing solution to EHR.

There are more than 22.000 users registered on the Andaman7 platform coming from 35 countries. The platform is available in 22 languages, and 80% of the users are from Europe. Andaman7 is also connected to some hospitals, laboratories, pharmaceutical companies to either send personal data to the patient or gather information from them (as in clinical trials).

---

## 5.2 TARGET AUDIENCE DESCRIPTION – DATA OWNERS AND DATA SEEKER

---

### 5.2.1 Data Owners

Data owners are identified as the final users of the Andaman7 application. They can be either patients or health specialists. For now, the app is able to collect health data from various sources - manually entered by the owner, collected from care facilities (hospitals, labs, etc.), gathered by connected devices (Apple Health, Google fit, etc.). This creates the Andaman7 Electronic Health Record (EHR), that can be shared to any other user of the Andaman7 platform via a “peer to peer” exchange mechanism (data is not stored in the cloud). The content of this EHR (or a part of it depending on what the user is willing to share) is the data that is envisaged to be collected.

Data may also come from hospitals, laboratories, etc. (actually any health company that holds data about a patient). But, from a legal perspective, the patient is the real owner of the data and are the only ones that should be able to control the access to data seekers.

Collection of data will be mainly done through secured APIs. eHealth is currently evolving a lot and many systems now have their own APIs or are planning to have one soon. In Andaman7, APIs such as Google fit, or other connected device brands can provide public and/or private APIs that could be called by the DataVaults platform to get data. Most US hospitals also have implemented FHIR and are now able to provide their data easily. Europe is also currently adopting this standard.

Also, the other way is also possible: The DataVaults platform has its own API that can be called by any actors of the (e-)health industry. This solution is more attractive from DataVaults perspective but requires extra work from companies that want to upload data to the platform, and some may be difficult to convince.

---

### 5.2.2 Data Seekers

We can identify three main types of data seekers:

1. Private companies that hold and maintain electronic health records.

They are generally companies that provide services to the data owners (health applications, EHR data store, etc.). They usually are interested in all data of specific(s) type(s) concerning one user. The precise purpose really depends on the company and the services they provide to the user. Most of the time, the ultimate goal is to provide better and/or more services to the user and so make their service more attractive, having more users and increasing revenues. Depending on the provided services they can also be interested in collecting statistics from the community to provide even more information to their users.

2. Health institutions that hold and maintain electronic health records.

Generally, hospitals and national EHR organizations. They are interested in all medical data about a specific patient/user to keep and provide the most complete EHR possible. The goal is to make it available to the patient but also to health professionals that need as

much information to provide better care to the user. They are also interested in follow up and feedback data from patients before and after a care (care path, home care, etc.).

3. Health companies that collect anonymized/pseudonymized data to improve global health.

These can be laboratories, pharmaceutical companies, hospitals, etc. They are interested in more specific data coming from groups of people corresponding to specific criteria. One of their goals is to identify if there are any people corresponding to searched criteria in an anonymized way and get enough information about them to be able to reach them (for example through their hospital). For example, that could ease recruitment for clinical trials. Another goal will be to retrieve data specific to clinical trials. That could be specific questionnaires, health data collected by connected devices, events, follow-up, etc. They are usually interested in having statistics.

---

### 5.2.3 Means to Reach other Data Seekers

Regarding the first group mentioned above (Private companies that hold and maintain electronic health records) this is probably the easiest to reach. Bigger companies can be contacted directly through their business representatives. They should be convinced by the product and especially by the user base and quantity of data they could get/process from the platform. Smaller and newer companies should be easier to convince but also more difficult to find and contact directly. A good communication through the website, social networks, ads, etc. is the best way to bring them to the projects. Bringing smaller companies can also create a lot of visibility and attract bigger one.

The second group (Health Institutions) is probably the hardest one to reach because they are generally not looking to connect to private systems, and they are too many to contact them one by one. Besides that, it generally takes a lot of time and energy to convince them to collaborate (many approvals, security audits, ethics, compliance to some criteria). A better solution would be to have a partnership with software editors of such companies that could integrate the connection to the DataVaults platform directly into their software solution that is already in place in many hospitals.

For the third group (Health companies), the main argument in favour of the platform is the ability to find pertinent candidates to participate (or directly get data from citizens corresponding to some criteria) in the clinical study they want to run as it can be difficult for them to discover and reach them anonymously. They are too many to contact them directly and they don't always run such clinical trials so timing can be a key. A good communication on our key points for this specific purpose can bring some of them when they are looking for such solutions. Also like for the second group, partnership with software solutions they use can be a better approach.

---

## 5.3 DEMONSTRATOR SCENARIOS

---

### 5.3.1 Scenario A - Get data from Andaman7 and provide to health sector

Current users of Andaman7 will be able to connect to DataVaults to store all or part of their health data. This storage can be used as a backup to retrieve data when lost. This can also be used by third parties in the health sector (e.g.: clinical trial, research).

#### 5.3.1.1 Challenges Faced

Main challenge here relies on security, privacy of data and communication to the user. As we are dealing with sensitive data this is very important to have a platform that is able to ensure this security in every aspect and have a great level of anonymization/pseudonymization to avoid any leak of identity. Even though we achieve this challenge, we also have to inform the user correctly about that so its confidence towards the platform allows him to share his data.

A second challenge will be to provide enough data and added value for third parties of the health sector and convince them to use it. The medical field is a very difficult one and even if they are slowly evolving to the digital world, some are still very reticent.

#### 5.3.1.2 Main Objectives of Demonstrator's Scenario

The main objective for Andaman7 is to improve the general attractiveness of the application for both users and partners. To measure if this goal is reached, we can divide it in small measurable goals:

- Increase the number of registered users from expected 36k (based on a linear evolution of registered users) to 40k at the end of the project. An increase in the average daily/monthly registration should also be noted.
- Increase the number of active users from 2k at the beginning of the project to 2.8k at the end.
- Increase the number of provided services by the integration of new features and creating new partnerships.

The first two should show an increase of interest from the final users. The third one should help convince more partners. To achieve those goals we will in this scenario:

- Develop the backup of Andaman7 content through the DataVaults platform. Backup feature is asked by a lot of users to increase confidence towards the application and especially the storage of their data.
- Develop and run a fictitious clinical study to show the potential of the DataVaults platform associated with Andaman7 and communicate to potential partners about the results.

This will be measured through analytics tools integrated in the Andaman7 platform.

#### 5.3.1.3 Scenario A Evolution Plan

**I. Scenario A – Alpha Phase: Development of Connectors:** This first phase will mainly concern the development of connectors and the recruitment of data owners from Andaman7 users. A customized version of the Andaman7 application should be ready by the end of this phase.

- Development of the connector to connect to DataVaults Personal App and upload data from Andaman7.
- Development of the connector to retrieve Andaman7 data stored on the DataVaults Personal App
- Development of the restoration of data downloaded from the previous connector.
- Recruitment of Andaman7 users that want to test the DataVaults platform.

**II. Scenario A – Beta Phase: Patient recruitment and onboarding:** This second phase will onboard recruited users. Of course, this will allow us to test the feature and fix/improve it if needed. In parallel a fictitious clinical study will be developed with a special recruitment for this.

- Creation of new Andaman7 service and resources to connect Andaman7 to the DataVaults Personal App (privacy policy, informed consent, login, configuration of data to be uploaded, etc.)
- Onboarding of users.
- Collection of data according to patient preferences.
- Integration of a demonstration clinical study into Andaman7.
- Recruitment of patients to this fictitious clinical study from the existing group (based on criteria found in uploaded data) or from other groups.

**III. Scenario A – Final Phase: Clinical trial usage evaluation:** This final phase will be based on the retrieval of such data and what we can get from the platform to help running real clinical trials and/or improve health in general. The idea is to exploit as much as possible all features provided by the DataVaults platform to provide meaningful data to clinicals, document them, analyse the results and present it to potential data seekers.

- Retrieval of clinical study data from a Data Seeker perspective with compensations mechanism with analytics and statistics data retrieved.
- Documentation of results obtained from this demonstration study.
- Presentation of the solution and obtained results to Andaman7 partners for onboarding.

#### **5.3.1.4 Data Availability and Data Needs**

The first phase collects content of the electronic health record stored in the Andaman7 application. Those are raw data (Andaman7 proprietary format), documents and standardized data that are already collected from many sources (some hospitals, connected devices, other A7 applications, etc.). Patient recruitment will try to maximize the availability of those different types of data and increase the variety that will be uploaded on the DataVaults platform.

The second phase collects clinical trials data that are not available for now. Andaman7 is already running some clinical trials, but we are legally not allowed to communicate such data to third parties. The proposed solution is to create a fictitious clinical study based on our knowledge in this field and ask participants to complete this trial and so generate the needed data.

### 5.3.2 Scenario B - Data collection and Andaman7 improvements

Current users of Andaman7 will be able to connect to DataVaults to collect their data (coming from various sources) and store them in Andaman7 on their smartphone. This will make the data available to patients for reviewing, learning, using in other set ups (e.g., share additional data with their doctors, hospital, etc.). Data will mostly be raw personal data but also aggregated data (e.g., result of a clinical trial, comparison to a specific group, etc.).

#### 5.3.2.1 Challenges Faced

The first challenge in this scenario will be to find new sources of medical/health data in this “short” time lapse. As mentioned in the first scenario, the medical community is very hard and slow to convince, even more when the platform is new/under development.

A second challenge is about the quality of computed data. We have to be really careful on the aspects that will be used to identify similarities in patient profiles with the minimum storage and no profiling. We also have to be careful on the computation and how we can present it to the final user, so it is useful and has a positive impact. Presenting such data is usually a challenge as it shouldn't be considered as medical advice nor should it bring frustration to the patient.

#### 5.3.2.2 Main Objectives of the Scenario

This scenario's objectives are the same already described in the scenario A, but two additional objectives can be added:

- Increase the number of supported data types by 20 %.
- Increase the quantity of data available in each category and/or the number of categories.

The interest of the final user is here the main goal as getting more meaningful data from various sources to complete their record should attract more users. To achieve those goals, we will in this scenario:

- Complete the health record of patients with health data available on the DataVaults platform.
- Compute some statistics based on data of the patient or other users of the DataVaults and integrate them in the health record of patients as indication to bring some intelligence to the storage of data.

This will be measured through analytics tools integrated in Andaman7.

#### 5.3.2.3 Scenario B Evolution Plan

**I. Scenario B – Alpha Phase: Development of connectors:** In this first phase, we will mainly develop the connector to retrieve data from other sources and recruit patients to test this new functionality. The integration into the Andaman7 app will be done but will remain at a basic level.

- Development of the connector to connect to download medical data coming from other sources.
- Retrieval, conversion and insertion of data into Andaman7 smartphone app (basic).
- Recruitment of Andaman7 users that want to test the DataVaults platform.

**II. Scenario B – Beta Phase: Patient onboarding and addition of data sources:** This second phase will bring a better version of this connector that will add analytics to analyse what and how data are retrieved and used. This will bring information to improve and finalize the feature and especially display them in the Andaman7 app in a meaningful and helpful manner. Bringing new sources of data is a key point to improve this feature. Aggregation of data can also be a way to improve the content distributed to the user.

- Creation of new service and resources to connect Andaman7 to the DataVaults Personal App (privacy policy, informed consent, login, configuration of data to be uploaded, etc).
- Onboarding of users.
- Add analytics on retrieval, types of data, sources, etc.
- Retrieval, conversion and insertion of data into Andaman7 smartphone app (improvements)
- Search and involve new sources of data

**III. Scenario B – Final Phase: Statistics from other patients:** This final phase will bring more value to the user by trying to match him with people with similarities in their medical profile, retrieve statistics/analytics data from this community and display it to the user so he can compare his results to others. The identification of data that will be retrieved and displayed can be sensitive as it shouldn't be considered as medical advice.

- Identify criteria to find people that have similarities in their medical records and corresponding statistics/analytics data that could be interesting.
- Development of connectors to retrieve statistics/analytics data from people that have similarities.
- Retrieval, conversion, and insertion of data into Andaman7 smartphone app.

#### 5.3.2.4 Data Availability and Data Needs

Data stored in Andaman7 can be used for some parts such as analytics and statistics data. This will be uploaded to the platform with the first scenario. To successfully fulfil this scenario other sources of health data should be found and integrated into the platform. That can be Health applications, hospitals, pharmaceuticals, clinics, data manually uploaded to the platform, etc.

---

## 5.4 DEMONSTRATOR TECHNICAL DEVELOPMENT NEEDS

---

We need an open authenticated API to build connectors to the DataVaults platform. This API is very important to automate every step of the scenario and ease any interaction with the platform. Of course, this API should be documented so it is easy to understand and implement its usage. This API should allow to:

- Upload data to the DataVaults Personal App
- Download data from my DataVaults Personal App. Results should be filterable on some criteria (original source, type of data, etc.)

- Search and retrieve pseudonymized users or data based on criteria over the DataVaults Cloud Platform
- Query data on some criteria over the DataVaults Cloud Platform

To multiply sources of data, it is also important that we are able to get data directly from the source and link them to the right patient. This could be development of connectors to external systems or ways for sources to upload data to the personal cloud of a specific user through an API call or via a web application.

---

## 5.5 COMPENSATION METHODS CONSIDERED

---

From the Andaman7 perspective the DataVaults platform allows to add new services and features to the application. Those are supposed to help the user and so is already part of the compensation. In addition, premium features can also be enabled.

If the reason for collection is not the improvement of the user's experience or if another data seeker is implied to be the main reason behind this data collection, this strategy can evolve. In this scenario, monetary vouchers or services rewards can be used depending on the partner we are working with. For example, hospitals usually reward participants of a clinical trial with money. But we can also work with pharmaceuticals that could be interested in rewarding users with vouchers for cosmetics or parapharmacy products.

---

## 5.6 DEMONSTRATOR'S ETHICS CONSIDERATIONS

---

As we are dealing with medical data, ethics is very important. But actually, such usage of data doesn't bring too many constraints from the ethics perspective. Of course, the platform should ensure security and provide enough trust to the user to store such sensitive data. Also, the user should always give his consent before any process of data for specified purposes.

In some use cases, an approval by an ethics committee may be required:

- When very sensitive data is processed. Such data contains data revealing racial or ethnic origin, genetic data, biometric data for the purpose of uniquely identifying a natural person, etc.
- Processing is done for clinical and medical research i.e., medical analysis of the results with the aim of publishing scientific medical results from collected data.

A Data Protection Impact Assessment may also be required and will be conducted as we are dealing with sensible data even though it doesn't seem to be needed for our use cases.

The DataVaults Ethical Policy outlined in D9.2 will be followed, including the envisaged Ethics and Data Protection Impact Assessment Methodology, as well as the indications of the project's Ethics Advisory Board.

Further information on the ethics dimension can be retrieved in D10.1 and D10.2: the former describes the informed consent procedure that will be adopted during the pilot actions and includes the templates of the informed consent and information sheet, the latter describes the procedures for data collection, storage, protection, retention, and destruction and the ethics risk evaluation



## 5.7 IMPACTS EXPECTED

Scenario	KPI	Calculation Method	AS-IS Value	TO-BE Value	Verification Means
<b>A, B</b>	No of active users	Analytics from Andaman7 database + Firebase analytics tools	22k registered, 2k monthly active	40k registered, 2.8k monthly active	Goals reached + increase of new daily registration
<b>B</b>	No of data types supported	Implemented data type in Andaman7 app + Analytics (Firebase) on usage of each data	100	120	Collect the number of new types of data that could be retrieved from DV.  Usage of data type will be measured through analytics tools.
<b>B</b>	Increase of Volume of data per/ category	Analytics (Firebase) of volume of data per user for each data category	To be split by categories calculated (in average 326 pieces of data per user for all categories)	15%	Increase of data available for each category
<b>A, B</b>	New services Implemented	Implemented services in Andaman7 app + Analytics (Firebase) on usage of service	2	4	At least 2 new services implemented services from described scenarios.  Usage of services will be measured through analytics tools

**Table 5: Impact Target KPIs for Demonstrator #3**

### 5.7.1.1 Impact for the Organisation

The main impact will be to make the Andaman7 app more attractive for end-users but also for potential partners.

The addition of services and data sources, bringing more data of different types should increase the attractiveness of the application. If the app is more attractive, more users will be convinced to install it and use it on a regular basis.

All this should bring more visibility to the app and, in addition with the increase of users and usage, it should also convince more companies to trust our platform and start a partnership to bring or to collect data.

Analytics tools already present in the application will be able to measure easily the number of active users, usage of data types and categories, usage of specific services with analytics which will confirm the impact.

### 5.7.1.2 Impact for the Data Owners

Data Owners will get more out of the Andaman7 application thanks to the addition of services, automatic retrieval data coming from various sources and additionally, calculation/analytics on their

data which can lead to more features and intelligence to the health record. This should increase the global satisfaction of the application.

As data owners are users of the Andaman7 app, we can measure the number of active users (increase means a positive impact on users) and average rating of the app (measured on app stores or via NPS). Usage of data type and categories can also be an indicator. This can also be easily done thanks to analytics tools.

#### 1.1.1.1 Impact for Collaborators (Data Seekers)

In the medical field, Data Seekers place great emphasis on the amount of collected data, the quality of data and the persona of the data owner. Through the DataVaults platform they will be able to find and gather data more easily data from a specific target group while reducing R&D costs. As all pieces of data are stored in the same place, they can get a huge amount of data (raw or computed) easily. Quality can also be easily checked because the owner of the data is known and can be verified by the platform.

## 5.8 DEMONSTRATOR'S ACTIVITIES TIMELINE

This is a theoretical timeline of activities for the next month as it may evolve depending on the development and the deployment of the DataVaults platform itself. Especially for the developments of connectors that are mostly at the beginning of this timeline. Of course, as the development of the platform will occur during this whole time, connectors and integration may be adapted/improved according to those evolutions.

Demonstrator 3 ANDAMAN7	M19	M20	M21	M22	M23	M24	M25	M26	M27	M28	M29	M30	M31	M32	M33	M34	M35	M36
<b>Scenario A - Get data from Andaman7 and provide to health sector</b>																		
<b>Alpha Phase</b>																		
Dev. upload connector																		
Dev. download connector																		
Dev. backup																		
Patient recruitment																		
<b>Beta Phase</b>																		
Service creation																		
Users onboarding																		
Data collection																		
Fictitious clinical study																		
Study recruitment																		
<b>Final Phase</b>																		
Study data collection																		
Result documentation																		

[illegible]

Table 6: Execution Timeline for Demonstrator #3

## 6 DEMONSTRATOR #4 – SMART HOME PERSONAL ENERGY DATA

### 6.1 BUSINESS DESCRIPTION AND THE NEED FOR DATAVAULTS

MIWenergia is an SME electricity retailer founded in Murcia, Spain. It operates in Spain nationwide with a varied range of customers: residential, industries, small and medium enterprises, and large tertiary buildings.

Providing service to more than 3.000 customers with an annual energy transit of 100 GWh, MIWenergia offers clients with additional services related to energy efficiency such as energy consulting and audits, design and installation of photovoltaic systems for self-consumption, electric vehicles charging stations, and energy monitoring.

The first step of all methodologies used to provide the previously mentioned services is collecting the needed data to be processed. Some data are public, others are technical ones and are indeed in the hand of the service provider, but the most important data are only known and owned by the user, like the history of electricity consumption, number of inhabitants in the house, or the building orientation.

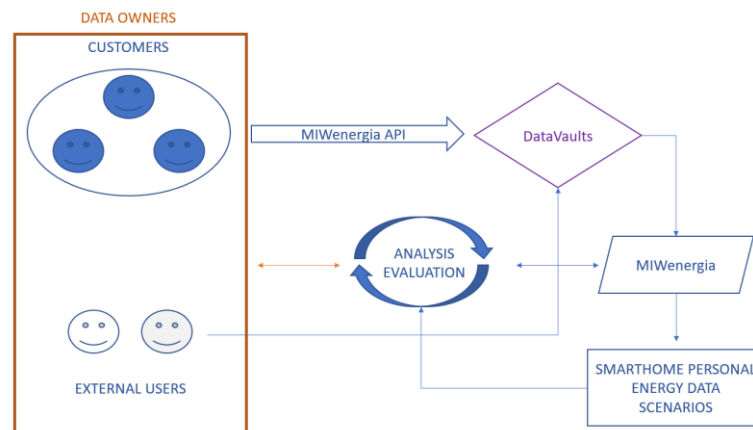
There are currently payment platforms for third-party access to electricity consumption data. These platforms provide monthly aggregated data per supply point, so the information is not very detailed, being almost useless for the design of a photovoltaic system or the energy prediction. A platform like DataVaults, where all the user generated data is provided to the platform by the data owner, can give access to all the companies interested in those data, connecting the data processors, distribution system operators, or electricity retailers, and letting other energy entities use that detailed information about electricity consumption with the user's permission. MIWenergia considers a platform for accessing data like DataVaults needed due the low quality of the platforms and data which are available now. Furthermore, the company only has access to the data of their customers, DataVaults gives the opportunity of having the data of all the platform users. MIWenergia will utilise DataVaults, to try and acquire such data directly from customers that are willing to share those, towards providing better services to them.

### 6.2 TARGET AUDIENCE DESCRIPTION – DATA OWNERS AND DATA SEEKER

#### 6.2.1 Data Owners

To analyse the usability of having access to user's energy data in addition to other personal data, there is a targeted audience, who will use the platform for its analysis. Not only MIWenergia's customers are planned to be engaged, but additional data owners are expected to be enrolled to evaluate the potential of DataVaults. The differences between having the hourly energy consumption data through DataVaults or an annual prediction will be demonstrated with the implementation of the scenarios of the pilot. At least 15 to 20 residential customers with various profiling will be actively participating, to test the phases of the platform and check the functionalities of it. Also, to verify the proper integration of MIWenergia's data to the platform. Supplementary users, 5 to 10, without business relations

with the company are foreseen to participate, to mitigate the possible subjectivity from customers due to the relationship with the company.



**Figure 2: Data owners data integration in DataVaults**

The supplementary users will be crucial to point out which aspects of the platform provide an additional value to the actual energy market, and the reasons to promote DataVaults to other electricity retailers or data processors. The distinction of the two groups can define the basis to promote to third parties the integration with DataVaults, acting as a bridge between the data owner and the data storage entity.

This connection will be established through an API for the hourly electricity consumption data, additional information should be manually uploaded by the users.

For full security and privacy compliance, it is planned that in the first phase of the platform no real user data will be transmitted. The use of fabricated data will allow API communication using the same associations and restrictions as real users.

### 6.2.2 Data Seekers

As a data brokerage platform, DataVaults needs to be fed with data sources coming from individual users (customers of energy suppliers) and that are not automatically collected by the energy companies, as they are produced mostly inside the premises of each customer. By having such data, the DataVaults platform will be highly attractive to data seekers, who are in need to acquire such data in order to improve their services. Energy services are based on the history of electricity consumption, which sets the electricity pattern/behaviour of the customer, without the past energy performance it is impossible to predict future energy consumption or to define actions to reduce it. Energy services companies haven't direct access to the data needed, they normally request it to the data owners, who lack the knowledge to collect and send it to the service provider.

There are many national and international entities in the energy sector. From regulated ones like the transport system operators and the distribution system operators, to liberalized companies like electricity retailers and engineering offices that provide energy efficiency services. The last group has a limited access to data due all the actual restrictions from the first group. A complete list of all active electricity retailers operating nationwide has been compiled and can be used to build synergies. Those, like MIWenergia, seek the data of other users in order to offer them their services.

From commercial campaigns, to improve the performance of the own company, hourly energy consumption data are highly demanded to set and achieve goals, like reducing the deviation of market electricity purchase, reducing the timing and increasing the effectiveness in dimensioning photovoltaic systems and offering energy efficiency personalized offers to the customers.

Using national standards for the data format let third parties to participate during the project and after the finalization. With 468 electricity retailers in Spain and 6175 in Europe<sup>3</sup> DataVaults “SmartHome Personal Energy Data” scenario has a high replicability and scalability.

DataVaults is a game changing platform in the energy sector, aiming to be disruptive and fair. Data Seekers can be reduced their operation costs at the same time that data Owners are benefited. In other hand, energy services will have a higher impact due the deep knowledge about the data owners (e.g., energy consumers), improving the quality of the solutions. The brokering platform set the data seekers in a privileged position where all parties obtain a profit satisfying the customer’s needs.

MIWenergia will participate not only as the bridge to the data owners, but also as a data seeker, additionally, other private and public institutions such as universities or research centres can be considered data seekers for research and statistical purposes. Within this framework, it is proposed to involve at least two data seekers to provide the point of view of both the private and the public sector. The participation of external data seekers is not foreseen until the final phase of the platform. In the event that the planned developments are advanced, this participation can be transferred to when the platform contains the necessary implementations.

---

### 6.2.3 Means to Reach Data Seekers

The energy sector is a populated market where all the members are interconnected, with new tools and regulations emerging constantly. Even though, the changes are slowly implemented and accepted.

The first way to reach data seekers is to ignite the spark of curiosity in them. The dissemination and communication actions from WP8 must turn the reactivity into proactivity. Once the platform is familiar and the functionalities are running, MIWenergia will contact collaborators such engineering office and other electricity retailers to promote the use of DataVaults. Other entities such as public universities with research relationships to MIWenergia will also be invited to use the platform.

The contact will be done by the usual methods with each entity, normally by email, introducing the advantages and potential of the platform. A manual or tutorial will be developed to enhance and simplify the start of use.

---

<sup>3</sup> Eurostat electricity market indicators 2019 – Available at: <https://ec.europa.eu/eurostat/statistics-explained/pdfscache/9144.pdf>

The first phase of the platform will be used as an introduction to the platform, as a demonstration. Access to external data seekers is not expected to be provided until the minimum functionalities have been implemented and tested by the pilots.

During all phases of the platform, the feedback of the data seekers will be collected, and it will determine the level of acceptance of the data seeker's tools. Suggestions and comments are expected to be analysed to improve the user experience.

## 6.3 DEMONSTRATOR SCENARIOS

### 6.3.1 Scenario A - PV installation design for self-consumption

In order to design an adequate PV installation, its needed not only the electricity consumption of the buildings but also some extra information. The data providers can introduce the required manual information for the data seekers to use it and well design a photovoltaic system.

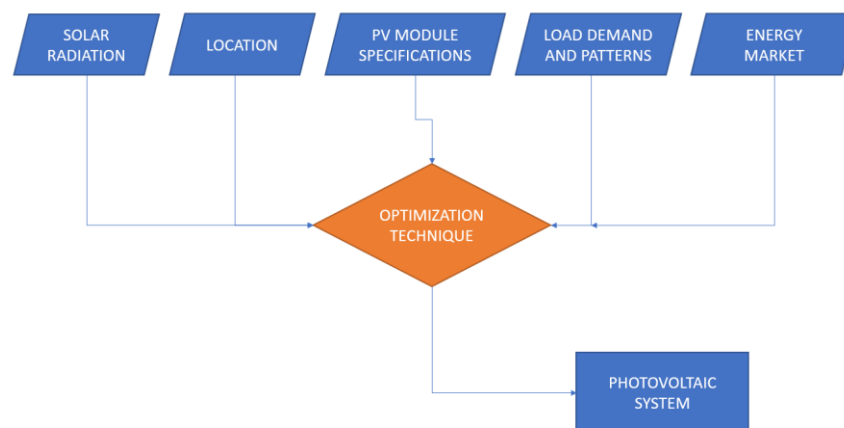


Figure 3: Design of a photovoltaic system

Additional data to be “manually” submitted include the following:

Data	Details
<b>Location</b>	At least city, the more detailed the better
<b>Available outside space</b>	Square meters
<b>Kind of roof</b>	Flat or slope
<b>Available outside space orientation</b>	North, South, East, West

Table 7: Description of data needs for Demonstrator 4 - Scenario A

With the supplementary data and the hourly electricity consumption, there is a complete package with the minimum inputs to calculate properly by the data seekers the specifics for a photovoltaic installation to reduce the electricity consumption by the data owners.

Even without the additional information on the electricity consumption, there are some basic analyses the platform should perform over the electricity consumption. Those analytics can be used as initial indicators of the viability of using solar energy and are the following:

- Annual electricity consumption. Total annual aggregation of electricity consumption in kWh, calculated as the sum of the last 8.760 (8.784 if a leap year) registered hourly consumption.

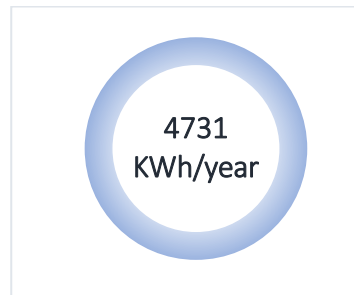


Figure 4: Annual electricity consumption analysis

- Distribution of the daily consumption per hour in one year. Total hourly aggregation of electricity consumption in percentage, calculated as the sum of the 365 (366 if leap year) registered daily consumption for each hour divided by the annual total consumption and multiply by 100.

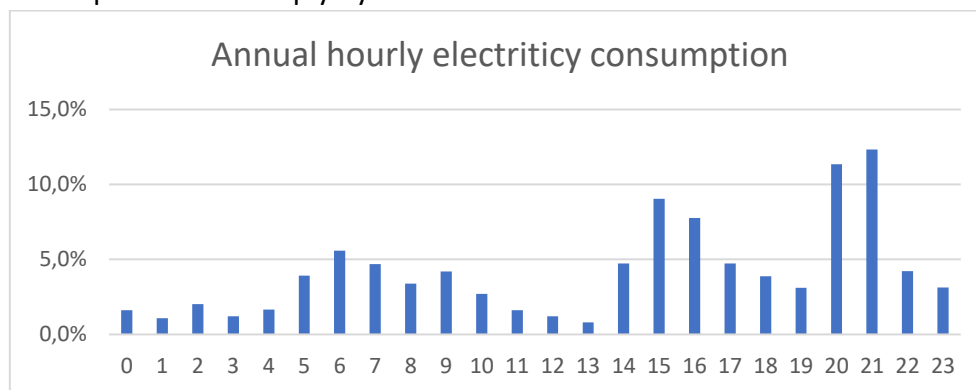


Figure 5: Hourly energy consumption profile analysis

- Distribution of the monthly consumption per year. Total monthly aggregation of electricity consumption in percentage, calculated as the sum of all the hourly consumption in each month registered divided by the annual total consumption and multiply by 100.

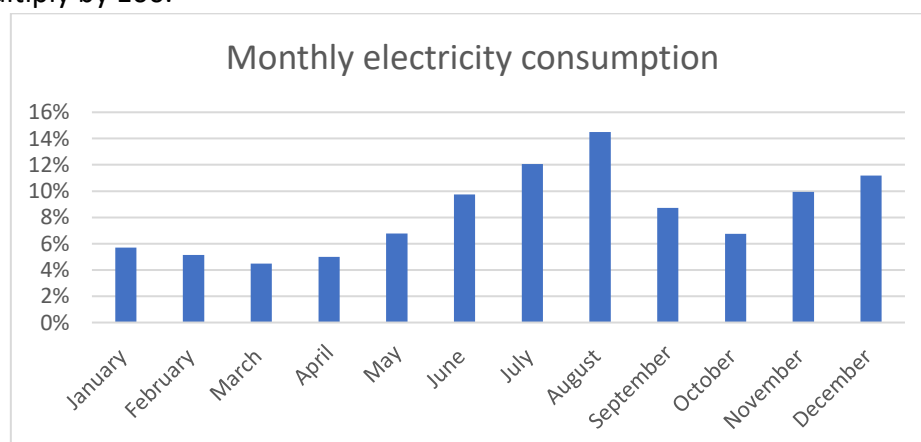


Figure 6: Monthly energy consumption profile analysis



For the correct sizing of a photovoltaic installation, consumption and production must be matched. The production is defined by the earth's movement around the sun, so the electricity consumption should take place during the day, more specifically between 08:00 and 16:00, in order to obtain a high yield.

Besides, solar radiation is more intense in the summer months, so the distribution of consumption should be concentrated in this period (May to September).

The sizing criteria of the photovoltaic installation are variable depending on the person in charge of its development, a variable widely used is the amount of energy that is self-consumed. Using this criterion, a rate of 40% can be a minimum target to achieve.

The annual consumption, together with the distribution of this previously described, will be the one that will finally define the power to be installed.

#### *6.3.1.1 Challenges Faced*

The current problem is the access to the electricity consumption data for which the photovoltaic system is to be designed.

Due to the current electrical system, the electricity distribution company is in charge of measurement. But it is the electricity retailer that bills the customer. The electricity distributor provides the metering data to the supplier but is not always obliged to do so, and the frequency of the measurements may not be sufficiently detailed to be useful (hourly frequency). Even if the data chain is complete, and the measurement frequency is valid, the customer only perceives the electricity consumption as a cost on his bill, a paper, or pdf invoice.

Although there is currently an attempt to digitize the sector, and many electricity distributors and retailers have online platforms to view and download electricity consumption. This download must be done by the customer after registering manually on the platform, logging in, and going through an arduous manual process on the web interface. This generates a file that has to be sent to the company that wants to use it, often being manually modified by the technicians to adapt it to the process they use for the matching of electricity consumed and demanded.

This process is complex for customers, who are usually not willing to invest their time and effort in it.

In addition to the difficulty of accessing the data, the lack of knowledge in the energy sector, such as the units of measurement (kW, kWh, V, A), their meaning and implication, make conversations with customers interested in photovoltaic systems not very effective, and there are usually errors in the information provided by the customer.

Due to all the issues commented on, many times the photovoltaic designers use approximate data instead of detailed one, reducing the reliability of the proposals. Generating dissatisfaction among customers, lack of credibility and distrust in the sector.

### 6.3.1.2 Main Objectives of Demonstrator's Scenario

As an electricity trading company, MIWenergia uses to exchange files with electricity market participants based on standard formats and structures.

Based on the knowledge of the current state of difficult access to electricity consumption data, and the high interest in renewable energy production systems such as photovoltaics. MIWenergia wants to facilitate the communication of this information for its correct integration in DataVaults.

The integration of additional information in the platform, which is currently requested by phone or email, facilitates the accuracy of the information. The fulfilment of the company's objectives, such as the installation of photovoltaic power, can be achieved more easily thanks to the consumption analysis of the platform, being able to easily segment the optimal customers to recommend photovoltaic self-consumption.

Furthermore, the aim is to expose the differences between a photovoltaic design based on monthly data and hourly data. The first one is commonly extended as a method of design, but with an accuracy unreliable. As such, the objective in this case is to have:

- A functional API that returns customers' hourly consumption.
- Annual consumption analysis.
- Analysis of annual hourly consumption.
- Annual monthly consumption analysis.
- Collection of the data described in 1.3.1 through DataVaults.
- Analyse reliability of photovoltaic proposals based on approximations.
- Improve image in the energy market.

By achieving these objectives, the barrier of accessing to electricity consumption data is eliminated, at the same time that simple but useful analysis are providing knowledge about the energy patterns of the users.

### 6.3.1.3 Scenario A Evolution Plan

For each phase of the platform, there will be a different scope of this scenario, to adapt the reach to the functionalities and tools of the software.

**I. Scenario A - Alpha Phase: Implementation of data sharing connectors and user recruitment:** The focus on this scenario is on establishing the connection and the data dump and making the necessary adaptations and the definition of the changes or developments required for the next phase.

- Testing of the MIWenergia API.
- API adaptation.
- Additional mock-ups definition.
- Verification of shared data.
- Participant Recruitment.

**II. Scenario A - Beta Phase: Energy analyses for photovoltaic design:** Once the platform has ensured the correct functioning of the data collection platform, ensuring the security and

privacy of the data, we will proceed to the use of the actual participants. Performing the proposed analyses in the playground

- Add functionalities to the API, data reliability.
- Participation of real users.
- Energy basic analysis at the playground.
- Data extraction using DataVaults.
- Implementation of virtual wallet.
- Implementation of compensation methods.
- Development of photovoltaic system calculation method.

**III. Scenario A - Final Phase: Photovoltaic design models:** In the final phase of the platform, it is intended to be able to integrate electricity consumption data from non-customer users. Likewise, the use of the playground for the development of more complex analyses.

- Development of API bridge data for non-customers.
- Use of compensation methods.
- Use of photovoltaic system calculation method.
- Comparison of photovoltaic design models.

#### 6.3.1.4 Data Availability and Data Needs

Electricity consumption information with the hourly frequency of customers is already stored in the MIWenergia database. This information is provided by the electricity distributor through file transfer protocols. MIWenergia is responsible for collecting this data and storing it in a structured way for its customers. The information is not collected or transmitted in real-time, so the data available for consultation is delayed by three to four days.

In addition to the non-availability of real-time information, there are two types of data regarding electricity consumption: metered data and billed data. The electricity distribution company periodically sends the metered electricity consumption data, and once these data are validated in the billing period and the billing period is closed, it sends the invoiced electricity consumption data.

Source	Type of data	Description of the data	Volume	Origin	Source Update
<b>Energy consumption data</b>	Structured data	Consumption given by the DSO. Hourly frequency. Only our clients.	1,27MB/user/year	MIWenergia Database (API)	Although the initial source is the electricity distribution company, the connection to DataVaults will be through the MIWenergia API to its database.
<b>Personal data</b>	Structured data	Full name, bank account, address, email, phone number. Only our clients.	Unknown	The user by signing the contract.	

**Table 8: Data Availability for Demonstrator 4 - Scenario A**

Normally there are no major modifications between the two measurements sent. Because the billed measurement data are the legally valid ones, they overwrite the measurement data. Therefore, the result of a query before closing the customer's billing may not give the same result as a later query. Currently the origin of each data is traced, so it is possible to verify whether the data is metered or billed, the latter being more valid.

Source	Type of data	Description of the data
<b>Dwelling</b>	Structured data	<ul style="list-style-type: none"> <li>• Location</li> <li>• Available outside space.</li> <li>• kind of roof (slope or flat)</li> <li>• Available outside space orientation</li> </ul>
<b>Energy consumption data</b>	Structured data	<ul style="list-style-type: none"> <li>• Electricity consumption from non-clients</li> </ul>

**Table 9: Data Needs for Demonstrator 4 - Scenario A**

### 6.3.2 Scenario B - Improve profiling of the clients to enhance energy efficiency

At the moment, all European Union countries are under a framework of common regulatory recommendations. In the field of energy, energy efficiency criteria are established in national plans. These plans determine the maximum energy consumption ratios of a home or business. They also establish levels for energy certification. This energy certification is currently mandatory for housing operations such as renting or selling a dwelling.

The electricity and energy sector are currently labelled as uncertain for the vast majority of customers. This is due to the lack of information and training in the fields of knowledge, together with the malpractice of many of the participants in the sector. For these reasons, customers are generally unable to know whether they are consuming too much energy, whether these are normal values or whether there may be a technical problem. They simply know that they are paying for it.

Reducing energy consumption and polluting emissions are sustainable development goals set by the United Nations. To improve this knowledge in the sector, as well as to facilitate energy certification and energy efficiency analysis. This scenario seeks the collection of additional information to energy consumption for its evaluation.

In order to carry out measures with a greater impact, we will try to intervene first in the buildings with the worst ratings. This requires the classification of the supply points. In order to perform this classification, private data will be needed, which will be collected manually through DataVaults.

Additional data to be “manually” submitted include:

Data	Details
<b>Year of construction</b>	Year of the building construction
<b>Size of the dwelling</b>	Area in square meters
<b>Location</b>	Province
<b>Number of inhabitants</b>	Number of people living together

**Table 10: Description of data needs for Demonstrator 4 - Scenario B**

Based on the location of the house there are ranges of energy consumption, as average values for Spain can be used as follows.

A	e<44.6 kWh/m <sup>2</sup> year
B	e<72.3 kWh/m <sup>2</sup> year
C	e<112.1 kWh/m <sup>2</sup> year
D	e<172.3 kWh/m <sup>2</sup> year
E	e<303.7 kWh/m <sup>2</sup> year
F	e<382.6 kWh/m <sup>2</sup> year
G	e>382.6 kWh/m <sup>2</sup> year

Figure 7: Energy efficiency class approximation, Spain

It is proposed to create a new energy efficiency criterion that takes into consideration the number of inhabitants of a house. This criterion will be evaluated and compared with the existing official one to analyse its usefulness both for users and for its official use. Since it is considered that the current one used does not take into consideration relevant factors.

#### 6.3.2.1 Challenges Faced

To collaborate in the reduction of energy consumption in homes through energy efficiency actions, it is necessary, on the one hand, to know where to focus efforts and, on the other hand, to previously inform users about their current status so that they are more receptive to the measures. The lack of implementation of correct indicators, or their follow-up, monitoring and control are the main sources of error when it comes to meeting energy efficiency objectives. In addition, the difficulty in accessing energy consumption or personal data generates uncertainty in the processing of information. Ignorance of the subject keeps the user from showing interest in taking measures that could reduce electricity consumption and therefore reduce the cost of the electricity bill. If the service provider had access to the necessary data to compare users, it could show comparisons with respect to similar users. Often, the benchmarking method is effective in encouraging users with poorer evaluations to try to improve, as well as rewarding those with good results to continue to do so. The cost-effectiveness of energy efficiency measures, such as the replacement of luminaires with energy-saving technologies like LED bulbs, is increasingly high. There are proposals where there is no investment by stakeholders, who pay for these energy efficiency measures with the savings they achieve.

#### 6.3.2.2 Main Objectives of Demonstrator's Scenario

The main objective for this scenario is to encourage the implementation of energy efficiency measures, reducing customers' electricity consumption. While increasing knowledge of the energy sector among participants. In order to plan and to evaluate energy efficiency measures first of all the actual energy performance is needed, and so the main developments of this scenario are:

- To have a functional API that returns customers' hourly consumption.
- Annual consumption analysis per square meter of building.
- Analysis of annual hourly consumption per inhabitant.
- Collection of the additional data through DataVaults.
- Promote energy efficiency measures.
- Reduce participants' electricity consumption.

### 6.3.2.3 Scenario B Evolution Plan

For each phase of the platform, there will be a different scope of the scenario, to adapt the reach to the functionalities and tools of the software.

**I. Scenario B - Alpha Phase: Data retrieving and energy efficiency:** The main actions to be carried out in the first phase of the platform are those necessary for the subsequent testing of the scenario. Establishing the connection and the data dump and making the necessary adaptations. As well as the definition of the changes or developments required for the next phase.

- Testing of the MIWenergia API.
- API adaptation.
- Energy efficiency mock-ups definition.
- Verification of shared data.
- Participant Recruitment.

**II. Scenario B - Beta Phase: Energy efficiency analyses:** Once the platform has ensured the correct functioning of the data collection platform, ensuring the security and privacy of the data, we will proceed to the use of the actual participants. Performing the proposed analyses in the playground.

- Add functionalities to the API, data reliability.
- Use of real data.
- Clustering based on energy efficiency.
- Data extraction using DataVaults.
- Implementation of virtual wallet.
- Implementation of compensation methods.
- Development of energy efficiency factors.

**III. Scenario B - Final Phase: Energy efficiency recommendations:** In the final phase of the platform, it is intended to be able to integrate electricity consumption data from non-customer users. Likewise, the use of the playground for the development of more complex analyses.

- Development of API bridge data for non-customers.
- Use of compensation methods.
- Recommendation of energy efficiency measures.

### 6.3.2.4 Data Availability and Data Needs

On the basis of the above, only the hourly electricity consumption is available. The intention is to obtain it through the platform to check the correct integration of the MIWenergia API with DataVaults.

Source	Type of data	Description of the data	Volume	Origin	Source Update
<b>Energy consumption data</b>	Structured data	Consumption given by the DSO.	1,27MB/user/year	MIWenergia Database (API)	Although the initial source is the electricity distribution company, the

		Hourly frequency. Only our clients.			connection to DataVaults is through the MIWenergia API to its database.
<b>Personal data</b>	Structured data	Full name, bank account, address, email, phone number. Only our clients.	Unknown	The user by signing the contract.	

**Table 11: Data Availability for Demonstrator 4 - Scenario B**

With the personal information that data owners must enter through the platform, it will be possible to make consumption ratios and assess the potential improvement in electricity consumption.

Source	Type of data	Description of the data
<b>User data</b>	Structured data	<ul style="list-style-type: none"> <li>• Number of inhabitants per building</li> <li>• Worktime</li> <li>• Preferences</li> <li>• Hobbies</li> <li>• Tastes</li> </ul>
<b>Dwelling</b>	Structured data	<ul style="list-style-type: none"> <li>• Location</li> <li>• Year of construction</li> <li>• Size of the dwelling</li> </ul>
<b>Energy consumption data</b>	Structured data	<ul style="list-style-type: none"> <li>• Electricity consumption from non-clients</li> </ul>

**Table 12: Data Needs for Demonstrator 4 - Scenario B**

### 6.3.3 Scenario C - Energy consumption patterns with personal preferences

One of the biggest problems and concerns of service providers is to maintain and improve their customer portfolio. Currently, MIWenergia usage of the electricity consumption records is limited to billing and demand forecasting.

In order to provide personalized services or offers, information such as preferences, tastes or hobbies must be known in order to provide more targeted and accurate offerings, that would capture the attention of consumers. As this information is not available at present, it is not possible to analyse it and make decisions based on it. If customers would provide this information, discounts or customized promotions could be offered. Such as access to on-demand content platforms, discounts on take-away food apps, subscriptions to cultural and sporting events, etc.

The analysis of energy consumption together with this information and other information of interest, such as working hours, can provide relevant information on patterns or habits. For example, it could be intuited that the user eats out on weekends, does not cook at home on Friday nights, or spends more time at home. This information can encourage the acceptance of offers, while improving customer satisfaction. On the other hand, one of the costs that is difficult for an electricity trader to foresee are the deviations in the purchase of energy in the electricity auction. MIWenergia forecasts the consumption of its customers hour by hour using algorithms based on historical data and past experience. But one-off events or personal patterns are difficult to predict.

With additional information given by the user, other methods of forecasting electricity demand with better fits can be tested.

Additional data to be “manually” submitted for the execution of this scenario include:

Data	Details
<b>Worktime</b>	Select schedule
<b>Preferences</b>	Communication channel, Compensation methods
<b>Hobbies</b>	List (Art, Games, Sport, Music, Tech...)
<b>Tastes</b>	Food list, film/book/games Genres
<b>Daily habits</b>	Calendar of plans (Holidays, events ...)

**Table 13: Description of data needs for Demonstrator 4 – Scenario C**

#### 6.3.3.1 Challenges Faced

Due to the current state of development of the company's technologies, such as the website it is not possible to obtain customer information through cookies or other methodologies. This makes it difficult for the marketing department to design promotional campaigns, or to launch offers for customers. The absence of innovation and the lack of data keeps the company from standing out, while the effectiveness of campaigns is reduced. Obtaining this data through DataVaults provides the opportunity to analyse and use it to improve commercial campaigns and customer satisfaction. In the same way, obtaining additional information such as work schedules or plans, together with historical electricity consumption and other variables, allows the implementation of new demand forecasting models and to evaluate them against those currently used by the company.

#### 6.3.3.2 Main Objectives of Demonstrator's Scenario

The main objective of this scenario is to improve the availability of personal information for usage in marketing and promotional campaigns and at the same time planning energy bids in the market, by understanding how customers behave. MIWenergia participates in the hourly energy auctions, buying the energy predicted for its clients. In order to develop an energy forecasting model a prediction model based on neural networks is planned to be developed in DataVaults' playground.

- To have a functional API that returns customers' hourly consumption.
- Collection of the additional data through DataVaults
- Implementation of commercial campaigns based on customers' tastes and hobbies.
- Realization of consumption patterns and demand prediction through the playground.

#### 6.3.3.3 Scenario C Evolution Plan

For each phase of the platform, there will be a different scope of the scenario, to adapt the reach to the functionalities and tools of the software.

**I. Scenario C - Alpha Phase: Hobbies and interests data sharing:** The main actions to be carried out in the first phase of the platform are those necessary for the subsequent testing of the scenario. Establishing the connection and the data dump and making the necessary adaptations. As well as the definition of the changes or developments required for the next phase.

- Testing MIWenergia API.
- API adaptation.



- Hobbies and interests mock-ups definition
- Verification of shared data
- Participant Recruitment

**II. Scenario C - Beta Phase: External data loading to playground:** Once the platform has ensured the correct functioning of the data collection platform, ensuring the security and privacy of the data, we will proceed to the use of the actual participants. Performing the proposed analyses in the playground, as well as the testing of external data loading.

- Add functionalities to the API, data reliability.
- Use of real data.
- Implementation of virtual wallet.
- Profiles and basic analysis at the playground.
- Data extraction using DataVaults.
- Implementation of compensation methods.
- Definition of commercial campaigns.
- External data loading test on the playground.

**III. Scenario C - Final Phase: Demand forecasting using neural network:** In the final phase of the platform, it is intended to be able to integrate electricity consumption data from non-customer users. Likewise, the use of the playground for the development of more complex analyses. In this case, new technologies such as machine learning will be used.

- Development of API bridge data for non-customers
- Use of compensation methods
- Selection of commercial campaigns by customer
- Use of neural networks for consumption prediction in playground

#### 6.3.3.4 Data Availability and Data Needs

Based on the above, only the hourly electricity consumption is available. The intention during this scenario is to obtain it through the platform to check the correct integration of the MIWenergia API with DataVaults.

Source	Type of data	Description of the data	Volume	Origin	Source Update
<b>Energy consumption data</b>	Structured data	Consumption given by the DSO. Hourly frequency. Only our clients.	1,27MB/user/year	MIWenergia Database (API)	Although the initial source is the electricity distribution company, the connection to DataVaults is through the MIWenergia API to its database.
<b>Personal data</b>	Structured data	Full name, bank account, address, email, phone number. Only our clients.	Unknown	The user by signing the contract.	

**Table 14: Data Availability for Demonstrator 4 - Scenario C**

The following table shows also other data which are relevant to this scenario, and which will be requested by the user in order to be able to execute the scenario

Source	Type of data	Description of the data
User data	Structured data	<ul style="list-style-type: none"> <li>• Number of inhabitants per building</li> <li>• Worktime.</li> <li>• Daily habits</li> </ul>
Energy consumption data	Structured data	<ul style="list-style-type: none"> <li>• Electricity consumption from non-clients</li> </ul>

**Table 15: Data Needs for Demonstrator 4 - Scenario C**

## 6.4 DEMONSTRATOR TECHNICAL DEVELOPMENT NEEDS

Although the development and deployment of the MIWenergia API to provide access to participants' electricity consumption is being finalized, it is expected that future adaptations and modifications will be necessary. In addition, MIWenergia plans to implement access to electricity consumption data of non-customer participants. Acting as a gateway to other platforms.

Those could be:

- Electricity consumption API development
- API customization
- Development of photovoltaic system calculation method
- Development of demand prediction model
- Virtual wallet integration
- Development of API bridge data for non-customers.

## 6.5 COMPENSATION METHODS CONSIDERED

The methods of compensation for the data will be those established in the DataVaults platform. The virtual currency used in the platform will have a real equivalent to the user, so that participants will benefit from their collaboration. However, feedback from users is necessary in order to fine-tune compensation methods and adapt them to their needs and criteria.

Some of the proposed methods are:

- Electricity bill discounts .
- Promotions on smart devices.
- Gift cards for online platforms.
- Gift cards for local business.

## 6.6 DEMONSTRATOR'S ETHICS CONSIDERATIONS

User participation is completely voluntary. The selection of candidates and end users will ensure equal opportunity. Participants will be duly informed about the purpose of the project,

the data and their use, as well as the rights they have and the means to exercise them. Participation will be expressly consented to by filling out the project consent form.

All the information collected, stored and processed will be committed in full compliance with the Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and the free movement of such data (GDPR) and the Spanish Organic Law 3/18 of 5 December on the Protection of Personal Data and guarantee of digital rights.

The DataVaults Ethical Policy outlined in D9.2 will be followed, including the envisaged Ethics and Data Protection Impact Assessment Methodology, as well as the indications of the project's Ethics Advisory Board.

Further information on the ethics dimension can be retrieved in D10.1 and D10.2: the former describes the informed consent procedure that will be adopted during the pilot actions and includes the templates of the informed consent and information sheet, the latter describes the procedures for data collection, storage, protection, retention, and destruction and the ethics risk evaluation

## 6.7 IMPACTS EXPECTED

An overall positive impact is expected, which starts with improved customer satisfaction. This is expected to lead to an increase in the number of customers. Encouraging the participation of other data seekers is expected to improve the number of MIWenergia collaborators, as well as satisfaction and trust.

In addition, scenarios such as scenario 1 should allow for the development of other business avenues that will bring in new paid services.

In addition, scenarios such as scenario 1 should allow for the development of other business avenues that will bring in new paid services.

Scenario	KPI	Calculation Method	AS-IS Value	TO-BE Value	Verification Means
A, B	EMC Effective Management of customers	Survey (Likert scale 1-5)	60	80	Annual Surveys
All	RPS Revenues through offering personalized services	Income Statement	0	30000	Annual Economic Reports
A, B	RSA Revenues through sales agreements	Income Statement	0	20000	Annual Economic Reports
All	NOP Number of partners	Number of Registrations	10	20	Annual Economic Reports
All	CST Clients' satisfaction and trust	Survey (Likert scale 1-5)	50	70	Annual Surveys
All	PST Partner's satisfaction and trust	Survey (Likert scale 1-5)	80	95	Annual Surveys

**Table 16: Impact Target KPIs for Demonstrator #4**

#### **6.7.1.1 *Impact for the Organisation***

It is estimated that the greatest impacts will be on the organization due to involvement. The work with data owners and partners involves developments on the part of MIWenergia, which, in addition to improving the relationship with them, plans to increase its services.

In order to evaluate the proposed objectives, additional indicators are proposed to assess the improvement between the real and expected status. This difference is referred to the proposed increase in order to evaluate the achievement of objectives per unit. This is described in D6.1 Appendix A.

The main expected impacts on the company are improved customer management, increased revenue from personalized offers, and increased revenue from sales agreements.

Due to the offer of special services such as photovoltaic systems, or energy efficiency. New lines of business are expected. In addition, access to personal customer information can have an impact through offers and promotions for better customer management. The information to be evaluated will come mainly from the company's annual reports. It is planned to monitor the variables of interest in the shorter term, so shorter-term reports will be requested from the company.

#### **6.7.1.2 *Impact for the Data Owners***

The main data owners to be dealt with in the scenarios are MIWenergia's customers. Although participation in the project requires effort and dedication on the part of the participants. It is expected that the compensation mechanisms, together with MIWenergia's developments will be sufficient to improve their satisfaction and confidence.

#### **6.7.1.3 *Impact for Collaborators (Data Seekers)***

Access to innovative tools for partner companies allows them to be at the forefront of innovative measures.

The correct design of photovoltaic installations thanks to the access to electricity consumption data, or the selection of customers with greater margins for improving energy efficiency. It not only promotes the achievement of sustainable development objectives, but also the effectiveness of proposals to customers. This improvement is expected to increase the satisfaction and trust of current partners, as well as increasing the number of new collaborators.

---

## **6.8 DEMONSTRATOR'S ACTIVITIES TIMELINE**

The development of the different phases of the platform involves constant developments and improvements by the pilot. The following is the temporal approach of the actions planned in three scenarios. Some of the actions are common to the scenarios, so they are only proposed once.

Demonstrator 4 MIWENERGIA	M19	M20	M21	M22	M23	M24	M25	M26	M27	M28	M29	M30	M31	M32	M33	M34	M35	M36
<b>Scenario A - PV installation design for self-consumption</b>																		
<b>Alpha Phase</b>																		
Testing MIWenergia API																		
API adaptation																		
Dwelling mock-up definition																		
Data verification																		
Participant Recruitment																		
<b>Beta Phase</b>																		
API, data reliability																		
Real users																		
Energy basic analyses																		
Data extraction																		
Implementation virtual wallet																		
Implementation of compensation methods																		
PV calculation method																		
<b>Final Phase</b>																		
API non-customers																		
Use compensation methods																		
Use PV calculation method																		
<b>DataVaults</b> benefits analysis																		
<b>Scenario B - Improve profiling of the clients to enhance energy efficiency</b>																		
<b>Alpha Phase</b>																		
Energy efficiency mock-ups definition																		
<b>Beta Phase</b>																		
Clustering basic analyses																		
Data extraction																		
Energy efficiency factors																		
<b>Final phase</b>																		
Energy efficiency measures																		
<b>Scenario C - Energy consumption patterns with personal preferences</b>																		
<b>Alpha Phase</b>																		

Hobbies and interests mock-ups definition																			
<b>Beta Phase</b>																			
Profiles and basic analyses																			
Data extraction																			
Def. Commercial campaigns																			
Playground external data test																			
<b>Final phase</b>																			
Use commercial campaigns																			
Neural networks analysis																			

Table 17: Execution Timeline for Demonstrator #4

## 7 DEMONSTRATOR #5 - PERSONAL DATA FOR MUNICIPAL SERVICES AND THE TOURISM INDUSTRY

### 7.1 BUSINESS DESCRIPTION AND THE NEED FOR DATAVAULTS

As a public administration, Prato is often contacted by commercial subjects (e.g., Google, credit card companies, mobile operators) interested in selling their customers' data to the city. Such approach not only is expensive for the administration, but it's even questionable from the privacy point of view and sometimes the supplied data is limited to specific context (localisation, payments) and no third parties owning personal data can be included. This introduces limitations in the administration's capability of leveraging citizens' personal data to analyse and improve its service and the same problem can be registered for other subjects operating in the city area, like for example cultural institutions (around 20, including theatres and museums).

The DataVaults approach would allow the administration (and other potentially interested subjects) to safeguard the users' privacy, while also collecting data from third parties provided by the users themselves, in a safe and privacy-oriented approach, since data will always be in the individual's full control. In addition, the money currently spent to acquire data from operators like Google or similar, could be partially transferred to the citizen data owners through appropriate compensation mechanisms.

Eventually, the DataVaults ecosystem could be tested as a possible further service provided to the citizens by the administration as a trustable and efficient tool for the sharing of personal data.

### 7.2 TARGET AUDIENCE DESCRIPTION – DATA OWNERS AND DATA SEEKER

#### 7.2.1 Data Owners

In all the three proposed scenarios (see below for their description), Data Owners will be individual citizens interested in testing the functionalities offered by the DataVaults tools.

From the technical point of view, there is no specific pre-requirement in the selection of the data owners, provided they have a smartphone and a minimum technological background to interact with the DataVaults tools. On the other hand, since the pilot scenarios are strongly connected with the Prato area, it is requested that participants are resident (actually this will be necessary in the certificate scenario) or, at least, having work, personal life or culture interests in the city.

Personal data will be collected from the different sources that the DataVaults tools will be able to interface with, in particular:

- data coming from the population registry (civil certificates)
- position data
- mobility preferences
- cultural preferences

- subscription data (newsletters)
- personal information (name, telephone, email, etc.)
- social network profiles

The data will be collected directly by the data owner, through the DataVaults Personal App loaded on his/her smartphone and shared according to the own preferences, with no direct involvement of the pilot managers or other parties.

---

### 7.2.2 Data Seekers

In the pilot phase it is planned to have four different Data Seekers:

- Administration Mobility Office
- Textile Museum
- Museum of the Pretorio Palace
- CAAF – CGIL fiscal support centre

According to the pilot results, some others might be involved in the final demonstration phase, like for example other cultural institutions, but also other public/private service providers.

The Data Seekers' interests in terms of personal data might be large enough and will depend on both data availability and analysis capacity.

The Administration Mobility Office is mostly interested in contact data (email, address, etc.), mobility preferences, position, in order to select citizen samples for surveys and questionnaires to enrich and refine mobility analysis and planning in the city.

Cultural institutions are mostly interested in contact data (email associated with real names) and in cultural preferences, including participation and liking of cultural events that can be extracted both from social networks and information provided by the data owners themselves. Such data will allow to improve the analysis of cultural offers and build rational citizen samples to whom send surveys and questionnaires.

Fiscal support centres usually require some documents like civil certificates (residence, marriage, etc.) to elaborate users' procedures: citizens need to collect such documents from the administration prior going to the centre, while DataVaults would allow to directly recover them by fully respecting the user's privacy.

---

### 7.2.3 Means to Reach other Data Seekers

The pilot experience and results obtained by the four involved Data Seekers (Mobility Office, Textile Museum, Museum of the Pretorio Palace, CAAF – CGIL fiscal support centre) will be the pulling factor to involve other Data Seekers. In particular, demonstrative actions will be organised to explain the DataVaults approach and show the pilot results, not only by the administration, but also from the Textile Museum, Palazzo Pretorio Museum and the CAAF – CGIL in the role of testimonials.

In addition, some specific reports and information contents, also on the social networks, will be produced to spread the pilot results.



The more effective and interesting the pilot results will be, the easier it will be to involve other Data Seekers and testify the innovative approach offered by DataVaults to obtain and manage individuals' personal data.

---

### 7.3 DEMONSTRATOR SCENARIOS

---

#### 7.3.1 Scenario A - Access to personal data for the analysis of mobility solutions

In this scenario, the Mobility Office acts as a Data Seeker and access the DataVaults platform to look for citizens' personal data (location, itineraries, means of transportation) in order to accomplish different types of activities:

1. to plan and verify mobility solutions in the city, also as an integration to the mobility plan released by the administration on a ten-years basis,
2. to identify adequate samples of citizens for the sending of surveys and questionnaires on traffic and mobility, according to their mobility preference and itineraries, in order to enrich the office knowledge base.

##### 7.3.1.1 Challenges Faced

Presently, the administration carries out investigations on the city mobility just in occasion of the redrawing of the city mobility plan, through the positioning of traffic sensors only in the most relevant locations and by sending questionnaires to the citizenship, to learn about mobility uses. This approach is very heavy from the organizational and economic point of view, and it is impossible for the administration to carry out such work more than once every 10 years. Moreover, in such an approach the questionnaires are sent to the widest audience, without any prefiltering that are not available at the moment, since individuals' personal data are only available as a pay service from big data providers like Google and this is expensive for the administration. In addition, the use of such data presents strong concerns regarding the individuals' privacy, and this represents a limitation for the administration's action.

The possibility offered by DataVaults to get personal data (including answers to questionnaires and surveys) tuned on mobility behaviours will improve the analysis of local mobility also on the basis of short-term evolutions. In addition, this will also improve the quality of surveys and questionnaires, due to the possibility of building citizens' samples that are more coherent with planning objectives. Eventually, the DataVaults approach allows to lighten the administrative burden in managing the users' privacy, since the data remains always in the full possess of individuals who are free to decide if and how to share it. On the other hand, the implementation of compensation schemas will facilitate the building of a trust and valuable relationship with data owners that will encourage them to share their personal data and participate in questionnaires and surveys pushed through the DataVaults platform.

##### 7.3.1.2 Main Objectives of Demonstrator's Scenario

The main objectives of this scenario are summarised below:

- To involve a reasonable number of data owners to provide their personal data including mobility behaviours and preferences through the DataVaults app installed on their smartphone. Personal data will include also answers and feedbacks to

questionnaires and surveys pushed through the DataVaults platform by the data seeker.

- To test the platform as a mobility operator (Data Seeker) and check the functionalities of data search and purchase, including the available compensation schemas to reward data owners.
- To evaluate the platform tools for data analytics to improve and integrate the current procedures adopted for mobility planning.
- To build citizens' samples as a mobility operator, on the basis of specific profiling specifications, and to push them specific questions to enrich the existing knowledge base.
- To build specific questionnaires and surveys as a mobility operator and send them to the selected citizens' samples.
- To get back sample's answers as a shared dataset according to the data sharing procedures implemented in the platform, including compensation schemas, and analyse them with the platform tools.
- To provide feedbacks and comments for the technical improvement of the DataVaults app and platform.

#### 7.3.1.3 Scenario A Evolution Plan

**I. Scenario A – Alpha Phase: Testing the tools with a small group:** At this phase, the focus will be on the test and verification of the basic functionalities available from the platform and app, namely fetching simple data sources and sharing with access control for the data owner and making simple search and data acquisition for the data seeker. This will imply the setup of a small group of users (data owners) with the aim of testing and providing feedbacks.

The main actions at this phase will be:

- Development of an external repository for collecting contact information distributed in different data sources and the API to connect it to the DataVaults app
- Set up of a small user group (3-4 data owners and the Mobility Office) to test the main user functionalities and the UIs: data fetching and sharing for data owners and data query and acquisition for the data seeker
- Collection of feedbacks in cooperation with the technical partners to improve the platform functionalities

**II. Scenario A – Beta Phase: Expanding the pilot action:** At this phase it is planned to have in place a more consolidated version of both the app and the platform, with the full set of functionalities in place, including the connections (API) with different data sources. The pilot action will therefore be expanded to a wider number of data owners (e.g., 10-20), to test the available functionalities and customisations, collect and share personal data from the app, verify the available compensation schemas and wallet. The larger amount of available data will allow the Mobility Office (data seeker) to improve data analysis, test the possibility of making user categorisation to obtain citizens' samples for surveys and questionnaires.

The main actions at this phase will be:

- Setup of a larger user group (data owners) to test the current version of DataVaults tools and procedures.

- Test of the connection to the available data sources and retrieve of available personal data on data owners' smartphone, in a realistic environment.
- Test of the available data query procedures as a mobility officer, with the adoption of available filters.
- Test of the available procedures of personal data sharing (data owners) and purchasing (data seeker) on the basis of the smart contracts built by the platform, including the compensation mechanisms in place and the wallet usage, if already in place.
- Analysis of the personal data obtained from the platform (data seeker) and evaluation of possible adoption of such data for mobility planning and building of citizens' samples based on user profiling. This procedure will help in identifying the most suitable data analyses for mobility requirements.
- Preparation of questionnaires/surveys on mobility issues with the platform tool, to be sent to selected citizens' samples and collection of related results.

**III. Scenario A – Final Phase: Consolidating the pilot action in a real context:** At this phase, the plan is to enlarge the data owner group to the city (up to 100 citizens) and to consolidate the adoption of all the available functionalities of both the DataVaults app and platform, including the building of persona profiles by the involvement of the Data Analyst role and the full deploy of the compensation mechanisms among all parties.

The availability of more data will allow the Mobility Office to experiment different types of data analysis with the platform tools and to integrate those with other type of data owned by the office. The Data Analyst role will be also crucial to provide further investigation opportunities based on personal data provided by the platform.

The outcomes of the pilot scenario will also pave the way to the possibility of including other data seekers, at least through the dissemination of the obtained results with the aim of planning possible exploitation of the opportunities provided by the DataVaults approach.

The main actions at this phase will be:

- Enlargement of the data owner group to the city, to make a sufficient quantity of personal data available for the test needs of the Mobility Office.
- Setup of the full data chain procedures foreseen in the DataVaults approach in a real context, including all the functionalities of both the app and the platform.
- Enrichment of the data analysis by the Mobility Office, also through the building of user personas and profiles, the preparation and sending of questionnaires/surveys to collect users' mobility preferences and the involvement of the Data Analyst role.
- Evaluation of the pilot results under different aspects: technological reliability, methodological effectiveness, data sharing approach acceptance for both data owners and data seekers. This will provide inputs for both the refinement of the technical approach and the definition of possible future business exploitation of the project rationale within the context of the public administration.

#### **7.3.1.4 Data Availability and Data Needs**

The available data for this scenario are personal data included in the CRM municipality system, the municipality newsletter system (Antherica) and other contact data in specific lists owned by the Mobility Office.

The data needs include individual's social network profiles (Facebook, Twitter, etc.), individual's GPS location, browser navigation history with reference to mobility interests, current mobility preferences, updates on personal contact info (if available), replies to surveys on mobility attitudes pushed through the DataVaults platforms.

---

### 7.3.2 Scenario B - Access to personal data for the improvement of cultural offer in the city

In this scenario, a Cultural Institution in the city acts as a Data Seeker and can access the DataVaults platform to look for citizens'/visitors' personal data to accomplish different types of activities:

1. to carry out data analysis for the improvement of its cultural offer,
2. to define adequate samples of citizens/visitors for the sending of surveys/market campaigns.

#### 7.3.2.1 *Challenges Faced*

In the current data business model approach, it is not easy for cultural institutions to collect a wide range of personal data, since this would require purchasing such data from external big players with high costs and a doubtful respect of privacy issues. In addition, data available from such players are not really focussed on cultural assets and specific analyses can be also very expensive. For the time being, the most common procedure for cultural institutions is to collect some information from those already attending their events through newsletter subscription or other feedback forms, but this makes obviously impossible to "capture" potential consumers and engage them.

The DataVaults platform can be used by cultural institutions for marketing surveys and loyalty campaigns for users of cultural services, identifying samples of individuals to whom they can send information material and ask for feedbacks, also including the provision of benefits (remuneration), such as example discounts, free tickets., etc. In addition, cultural institutions can use the DataVaults platform to collect information on people's cultural consumption and build more effective questionnaires and surveys to be pushed through the platform: Individuals will be able to share answers and feedbacks according to the data sharing and compensation procedures implemented by DataVaults and this will hopefully increase the number of replies obtained by the cultural institution. Eventually, the DataVaults approach will allow the lightening of the administrative burden to manage privacy issues in the relationships with data owners, since data always remains in full possession of the individuals.

#### 7.3.2.2 *Main Objectives of the Scenario*

The scenario has the following main objectives:

- To involve a reasonable number of data owners to provide personal data including cultural preferences, attendance and liking of cultural events.
- To test the platform as a cultural institution (data seeker) to check the functionalities of data purchasing including available the compensation schemas.
- To evaluate the platform tools for data analytics to improve the current procedures for the planning of cultural offer.

- To build citizens' samples on the basis of specific profiling items and to push them specific questions to enrich available information.
- To get back sample's answers as a shared dataset according to the data sharing procedures implemented in the platform, including compensation schemas, and analyse them with the platform tools.
- To get back sample's answers and analyse them with the platform tools.

#### 7.3.2.3 Scenario B Evolution Plan

**I. Scenario B – Alpha Phase: Testing the tools with a small group:** At this phase, the focus will be on the test and verification of the basic functionalities available from the platform and app, namely fetching simple data sources and sharing with access control for the data owner and making simple search and data acquisition for the data seeker. This will imply the setup of a small group of users (data owners) with the aim of testing and providing feedbacks.

The main actions at this phase will be:

- Development of an external repository for collecting contact information distributed in different data sources owned by the participating cultural institutions and the API to connect it to the DataVaults app.
- Set up of a small user group (3-4 data owners, Textile Museum and Pretorio Palace Museum) to test the main user functionalities and the UIs: data fetching and sharing for data owners and data query and acquisition for the data seeker.
- Collection of feedbacks in cooperation with the technical partners to improve the platform functionalities.

**II. Scenario B – Beta Phase: Expanding the pilot action:** At this phase it is planned to have in place a more consolidate version of both the app and the platform, with the full set of functionalities in place, including the connections (API) with different data sources. The pilot action will therefore be expanded to a larger number of data owners (e.g., 10-20), to test the available functionalities and customisations, collect and share personal data from the app, verify the available compensation schemas and wallet in a real environment. The larger amount of available data will allow the cultural institutions (data seeker) to improve data analysis and test the possibility of making user categorisation to obtain citizens' samples for surveys and questionnaires.

The main actions at this phase will be:

- Setup of a larger user group (data owners) to test the current version of DataVaults tools and procedures.
- Test of the connection to the available data sources and retrieve of available personal data on data owners' smartphone.
- Test of the available data query procedures as a cultural institution, with the adoption of available filters.
- Test of the available procedures of personal data sharing (data owners) and purchasing (data seeker) on the basis of the smart contracts built by the platform, including the compensation mechanisms and the wallet usage, if already in place.

- Analysis of the personal data obtained from the platform (data seeker) and evaluation of possible adoption of such data for the planning of culture offer and building of citizens' samples on the basis of user profiling.
- Preparation of questionnaires/surveys on cultural issues with the platform tool, to be sent to selected citizens' samples and collection of related results.

**III. Scenario B – Final Phase: Consolidating the pilot action in a real context:** At this phase, the plan is to enlarge the data owner group up to the city (up to 50 persons) and to consolidate the adoption of all the available functionalities of both the DataVaults app and platform, including the building of persona profiles by the involvement of the Data Analyst role and the full deploy of the compensation mechanisms among all parties.

The availability of more data will allow both the Textile Museum and the Pretorio Palace Museum to experiment different types of data analysis with the platform tools and to integrate those with other type of data owned by the two institutions to plan their cultural offer. The Data Analyst role will be also crucial to provide further investigation opportunities based on personal data provided by the platform.

The outcomes of the pilot scenario will also pave the way to the possibility of including other cultural institutions as data seekers, at least through the dissemination of the obtained results with the aim of planning possible exploitation of the opportunities provided by the DataVaults approach.

The main actions at this phase will be:

- Enlargement of the data owner group, in order to make a sufficient quantity of personal data available for the test needs of the participating cultural institutions.
- Setup of the full data chain procedures foreseen in the DataVaults approach in a "real" context, including all the functionalities of both the app and the platform.
- Enrichment of the data analysis by the two museums, also through the building of user personas and profiles with the involvement of the Data Analyst role, the preparation and sending of questionnaires/surveys to collect users' cultural interests and liking, to support the planning of the cultural events.
- Evaluation of the pilot results under different aspects: technological reliability, methodological effectiveness, data sharing approach acceptance for both data owners and data seekers. This will provide inputs for both the refinement of the technical approach and the definition of possible future business exploitation of the project rationale for the needs of cultural institutions.

#### *7.3.2.4 Data Availability and Data Needs*

The available data for this scenario are personal data included in the CRM municipality system, the municipality newsletter system (Antherica) and other contact data in specific lists owned by the participating cultural institutions.

The data need to include individual's social network profile (Facebook, Twitter, etc.), browser navigation history with reference to cultural interests and possible online ticket purchase,

users' GPS location, updates on personal contact info (if available), replies to satisfaction surveys on cultural events pushed through the DataVaults platform.

---

### 7.3.3 Scenario C - Access to personal data for the delivery of personal certificates

In this scenario, the CGIL-CAAF fiscal support centre accesses the DataVaults platform as a data seeker, with the objective of acquiring personal civil certificates from a user requiring some fiscal services.

#### 7.3.3.1 Challenges Faced

In the current approach, whenever requiring a fiscal service from the CGIL-CAAF fiscal support centre, a citizen has to procure some civil certificate in advance, either by going to the City Registry Office in person or by getting it from the online service. In both cases, this represents a burden and, moreover, the city administration becomes aware of the reason why the certificate is requested since the citizen has to declare it (at least to a broad extent) and this is not exactly in line with the GDPR prescription.

The automatic, authorised and regular downloading of personal certificates by third parties, made possible by the DataVaults tools, will reduce burdens for citizens and interested data seekers. In addition, although the certificates are always produced by the administration, the document exchange is carried out in a totally transparent way without the administration being aware of it and this makes the whole process more compliant with the GDPR legislation.

#### 7.3.3.2 Main Objectives of the Scenario

- To verify the software connection (API) between the city population registry and the DataVaults app for the download of the civil certificates provided by the Prato administration and implement the inclusion of the civil certificates inside the project data model.
- To test the proposed innovative document exchange between citizens and the CGIL – CAAF fiscal support centre by involving a small group of users (2-3 citizens) that will download the DataVaults app and register on the platform as data owner, while the CGIL – CAAF will act as a data seeker.
- To evaluate the whole procedure to check strong and weak points and suggest possible technical improvements, from both data owner's and data seeker's point of view.

#### 7.3.3.3 Scenario C Evolution Plan

This scenario is somehow restricted with respect to the previous ones since its main objective is the verification of the procedure to access individual's civil certificates by the CGIL – CAAF fiscal support centre through the DataVaults tools. Therefore, two different phases set in accordance with the platform technical developments should be sufficient to test such procedure.

**I. Scenario C – Alpha Phase: Testing the procedure:** At this phase, the focus will be on the test and verification of the basic functionalities available from the platform and app, in particular the connection to the municipality population registry to fetch the civil certificates

and share them within the data owner's full control. This will imply the setup of a small group of users (2-3 data owners) with the aim of testing and providing feedbacks.

The main actions at this phase will be:

- Development of APIs to connect the DataVaults app to the city population registry.
- Definition of certificate metadata to support the search functionalities.
- Building of the certificate dataset to be included in the individual's profile.
- Set up of a small user group (2-3 data owners, CGIL-CAAF fiscal support centre) to test the main user functionalities and the UIs: data fetching and sharing for data owners and data query and acquisition for the data seeker.
- Collection of feedbacks in cooperation with the technical partners to improve the platform functionalities.

**II. Scenario C – Final Phase: Consolidating the pilot action in a real context:** At this phase, the plan is to consolidate the full procedure of sharing and retrieval of civil certificates, by involving a maximum of 10 citizens. The outcomes of the pilot scenario will also pave the way to the possibility of including other data seekers requiring civil certificates, at least through the dissemination of the obtained results with the aim of planning possible exploitation of the opportunities provided by the DataVaults approach.

The main actions at this phase will be:

- Setup of the full data chain procedures foreseen in the DataVaults approach in a real environment, including all the functionalities of both the app and the platform for the sharing and retrieval of civil certificates.
- Evaluation of the pilot results under different aspects: technological reliability, methodological effectiveness, data sharing approach acceptance for both data owners and data seekers. This will provide inputs for both the refinement of the technical approach and the definition of possible future business exploitation of the project rationale towards third parties requiring civil certificates and/or other official documents from individuals.

#### **7.3.3.4 *Data Availability and Data Needs***

Available data are given by the information included in the municipality population registry and the civil certificates provided by administration.

There is no specific need for other types of personal data, but the civil certificates must be included in the individual's personal profile according to the technical requirements defined in the project data model.



---

## 7.4 DEMONSTRATOR TECHNICAL DEVELOPMENT NEEDS

---

The extra needs for the development of scenario A and B will be the development of a data repository and the API to connect it to the DataVaults app, to store personal data coming from different data sources, like for example contact lists in informal excel sheets. This will simulate a structured data source like those DataVaults can connect to with an API.

The extra need for scenario C is given by the need of building an appropriate API to connect the municipality population registry to the DataVaults app for the retrieval of the civil certificates.

---

## 7.5 COMPENSATION METHODS CONSIDERED

---

In principle, the adoption of a general compensation schemas based on points or tokens can be adopted both for scenario A and B. For scenario A, such compensation method can simulate a money payment to the data owners from the platform. For scenario B, an extension of such approach could be the spending of the accumulated points to get free entrance tickets to cultural events.

As far as scenario C is concerned, there is no need of a direct compensation scheme for the CGIL-CAAF fiscal support centre, since the sharing of the civil certificates is a necessity for the individual to conclude the administrative procedure that he/she is interested in. In case, we might test the possibility of having a compensation schema driven by the Prato administration, that might remunerate citizens through the DataVaults platform as an incentive not to go to the counter, thus ensuring the savings of administrative costs and personnel.

---

## 7.6 DEMONSTRATOR'S ETHICS CONSIDERATIONS

---

Data owners' participation will be completely voluntary. The selection of candidates and end users will ensure equal opportunity as much as possible and participants will be duly informed about the purpose of the project, the data and their use, as well as the rights they have and the means to exercise them. Participation will be expressly consented to by filling out the project consent form that will be available when registering on the project platform.

All the information collected, stored and processed will be committed in full compliance with the Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and the free movement of such data (GDPR) and the Italian Data Protection Code 101/2018 of 10 August on the Protection of Personal Data and guarantee of digital rights.

The DataVaults Ethical Policy outlined in D9.2 will be followed, including the envisaged Ethics and Data Protection Impact Assessment Methodology, as well as the indications of the project's Ethics Advisory Board.

Further information on the ethics dimension can be retrieved in D10.1 and D10.2: the former describes the informed consent procedure that will be adopted during the pilot actions and includes the templates of the informed consent and information sheet, the latter describes the procedures for data collection, storage, protection, retention, and destruction and the ethics risk evaluation

## 7.7 IMPACTS EXPECTED

Scenario	KPI	Calculation Method	AS-IS Value	TO-BE Value	Verification Means
<b>A</b>	Number of data owners involved	Number of registrations on the app/platform	0	100	Check of registration file
<b>A</b>	Number of available data sources	To be extracted from app functionalities	0	5	Check of App APIs
<b>A</b>	Number of shared datasets	To be extracted from the platform data base	0	200	Check of platform data base
<b>A</b>	Number of activated smart contract	To be extracted from the platform distributed ledger	0	100	Check of the platform DL
<b>A</b>	Number of data analysis procedures	To be extracted from the data seeker workflow	0	3	Production of data analysis report
<b>A</b>	Number of questionnaire s/surveys	To be extracted from the data seeker's workflow on the platform	0	2	Data seeker's documents
<b>A</b>	Improvement in the planning capabilities as perceived by the Office	Survey (Likert scale 1-5)	N/A	4,5	Survey results
<b>A</b>	Savings in the installation of traffic sensors and data acquisition procedures	Specific budget from the Mobility Office's records	10.000 euro	-50%	Expenditure records
<b>B</b>	Number of data owners involved	Number of registrations on the app/platform	0	50	Check of registration file
<b>B</b>	Number of available data source	To be extracted from app functionalities	0	5	Check of App APIs
<b>B</b>	Number of shared datasets	To be extracted from the platform data base	0	100	Check of platform data base
<b>B</b>	Number of activated smart contract	To be extracted from the platform distributed ledger	0	50	Check of the platform DL
<b>B</b>	Number of data analysis procedures	To be extracted from the data seeker workflow	0	3	Production of data analysis report
<b>B</b>	Number of questionnaire s/surveys	To be extracted from the data seeker's workflow on the platform	0	2	Data seeker's documents
<b>B</b>	Improvement in the planning capabilities of	Survey (Likert scale 1-5)	N/A	4,5	Survey results

	the cultural institutions				
<b>B</b>	Savings in data acquisition and analysis procedures	Specific budget figure	5.000 euro	-50%	Expenditure records
<b>C</b>	Number of involved data owners	Number of registrations on the app/platform	0	10	Check of registration file
<b>C</b>	Number of shared certificates	To be extracted from the platform data base	0	10	Check of the data owners' activity
<b>C</b>	Data owners' satisfaction in using the DataVaults tools	Survey (Likert scale 1-5)	N/A	4,5	Survey results
<b>C</b>	Improvement in the organisation workflow	Survey (Likert scale 1-5)	N/A	4,5	Survey results
<b>C</b>	Prato administration's savings in terms of costs and personnel	Cost of one certificate release operation at the counter [15 min personnel cost + paper cost]	Euro 4,10	~ 0	n. data owner x current cost at the counter

Table 18: Impact Target KPIs for Demonstrator #5

#### 7.7.1.1 Impact for the Organisation

From the organisation's point of view (Municipality of Prato) the number of data owners is a relevant KPI, as it gives the idea of the potential appeal of the DataVaults tools and procedures as a service that might be provided by the administration to the citizens, in order to support their personal data sharing with different data seekers in the full respect of the privacy issues.

In addition, the procedure carried out in scenario C regarding civil certificates (particularly if a compensation is foreseen) will represent a good incentive for citizens to avoid going to the registry counter to get civil certificates with a positive impact on administrative costs and personnel, thus extending the advantages already provided by the existing online procedure for the delivery of certificates that is currently not as much adopted as it could be.

Further impacts for the administration as a data seeker are reported in section 7.7.1.3.

#### 7.7.1.2 Impact for the Data Owners

The main impact for the Data Owners will be the management all personal data in the same framework in the full respect of privacy issues and control of data sharing procedures, together with the possibility of getting a compensation for data sharing. Therefore, from the Data Owners' perspective, a relevant KPIs is the number of available data sources, as this is connected to the potential completeness of personal data management by the individual, who can therefore perceive the DataVaults environment as a "one-stop" link point to the whole set of his/her personal data. Another meaningful KPI is given by the number of activated smart contracts, that is demonstrative of the effectiveness of the DataVaults tools as a mean to

establish a trustful relationship with data seekers and get possible rewarding for personal data sharing. On the whole, the experience carried out with the DataVaults tools should unburden the data owners' concerns on the privacy issues when dealing with personal data.

### 7.7.1.3 Impact for Collaborators (Data Seekers)

The main impact in this case is represented by the opportunity provided by the DataVaults tools to unburden and enrich the search of personal data with savings of money with respect to the current approach that sees the need of paying private big data players. This will result also in an improvement in data analytics procedures for more accurate and effective planning results both in the mobility and cultural scenario.

In addition, the opportunity to push surveys and questionnaires directly through the platform will greatly improve the possibility of getting answers and feedbacks from the individuals, who will get a compensation for sharing them with the data seeker.

Consequently, from the Data Seekers' perspective the most relevant KPIs are the number of data owners and the number of shared datasets, as they are representative of the utility of the DataVaults approach as a tool to support personal data availability and search. In addition, the number of activated smart contract is also a relevant KPI, as it gives an index of the effectiveness of the DataVaults approach in the building of the data value chain based on a trust relationship between data owner and data seeker.

## 7.8 DEMONSTRATOR'S ACTIVITIES TIMELINE

The table below reports the Gantt diagram related to the three above presented scenarios, where the three time periods are based on the planned phases for the release of the DataVaults app and platform. As discussed in the previous section 7.3.3.3, the development of scenario C will cover up to M30, since this will be sufficient to analyse and evaluate the proposed procedure for the sharing of personal certificates. Should there be any need to extend it, due to either technical issues or special needs on behalf of the CGIL-CAAF support centre, the activity will be extended accordingly in the period M30-M36.

Demonstrator5 PRATO	M19	M20	M21	M22	M23	M24	M25	M26	M27	M28	M29	M30	M31	M32	M33	M34	M35	M36
<b>Scenario A - Access to personal data for the analysis of mobility solutions</b>																		
Testing the tools																		
Releasing of external repository and API																		
Set up of a small testing group																		
Test activity on the platform basic functionalities																		
Collection of feedbacks																		

[illegible]

[illegible]

### Table 19: Execution Timeline for Demonstrator #5

## 8 CONCLUSIONS AND NEXT STEPS

In this deliverable, the DataVaults demonstrator partners, alongside with the technical partners that are supporting them, laid down the detail plans for the operation of the demonstrator scenarios, which will go hand in hand with the different releases of the platform in order to test the technical innovations of the project, evaluate the platform from various perspectives and most importantly to validate the main concept of DataVaults, by bringing into the picture real Data Owners who would be in a position to provide the anticipated feedback for optimising the platform and its offerings.

At the same time, the scenarios that have been devised go beyond the validation of the value and of the user experience of platform by Data Owners, and also test the usefulness of the platform and the business benefits it can provide from the perspective of Data Seekers, as the demonstrator partners take up this role and try through the operation of DataVaults to improve their services, procedures and offerings to their main customer audience. It is noted at this point, that all the technical verification and validation of the whole platform (like privacy guarantees, system responsiveness, etc.) will be conducted under WP5 and thus does not fall within the scope of the presented scenarios, without excluding the option that some tests might run having such a scenario as a reference/simulation point.

As such, the next steps which follow the delivery of D6.2 are to put the plans described in the document at hand in motion; that is to start executing in the different demonstrators the first stages of the identified scenarios that have to do with the integration of the demonstrators' systems with the DataVaults platform, the detailed definition and the acquisition of data from different data sources that need to be provided by the Data Owners, and the gathering of an initial set of individuals which will take up the role of Data Owners to run small-scale validation experiments, which are expected to gradually scale as new releases of the platform will surface and more complex editions of the prescribed scenarios will be kickstarted.

All of the above, will be documented in the following deliverables of WP6, evaluating both the technical and the business benefits that steam from the project.