

**D6.4 Validation and Evaluation of the Promoting independence of specific vulnerable groups**

**Use Case v2**

|  |  |
| --- | --- |
| Deliverable ID: | D 6.4 |
| Deliverable Title: | Validation and Evaluation of Promoting independence of specific vulnerable groups Use Case v2 |
| Revision #: | 1.0 |
| Dissemination Level: | Public |
| Responsible beneficiary: | SIVECO, VTT |
| Contributing beneficiaries: | SIVECO, BEIA, VTT, CGI, UniqAir, ISEC, Turgen, Semantik Ar-Ge |
| Contractual date of delivery: | 31.8.2019 |
| Actual submission date: | 2.10.2019 |

# Table of Content

[Table of Content 2](#_Toc20897331)

[1 Introduction 3](#_Toc20897332)

[1.1 Objective of this document 3](#_Toc20897333)

[1.2 Structure of this document 3](#_Toc20897334)

[2 Applied methodology for evaluation 4](#_Toc20897335)

[3 Rehabilitation decision support 7](#_Toc20897336)

[3.1 User level test results - system parameters evaluation done by client 7](#_Toc20897337)

[3.2 Restrictions of the evaluation processes 19](#_Toc20897338)

[4 Indoor air quality improvement at school 20](#_Toc20897339)

[4.1 Parameters and methods 20](#_Toc20897340)

[4.2 Reporting and decision making 29](#_Toc20897341)

[4.3 Restrictions of the evaluation processes 30](#_Toc20897342)

[5 Tracking of athletes with wearable sensors 31](#_Toc20897343)

[5.1 Parameters and methods 31](#_Toc20897344)

[5.2 Reporting and decision making 32](#_Toc20897345)

[5.3 Restrictions of the evaluation processes 33](#_Toc20897346)

[6 Conclusion 34](#_Toc20897347)

[7 References 35](#_Toc20897348)

# Introduction

## Objective of this document

This Deliverable 6.4 - Validation and Evaluation of the *Promoting independence of specific vulnerable groups* Use Case v2 deals with the validation tests performed in the task 6.4 in different pilots that are included in this use case; Rehabilitation decision support, Indoor air quality improvement at school, and Tracking of athletes with wearable sensors. This is an updated version of the first version of the pilot validation and evaluation report published one year ago.

The *Rehabilitation decision support* pilot combines environmental sensor data with physiological and behavioral sensor data to empower patients in a rehabilitation clinic with decision support tools for behavioral choices and treatment options. The goals are to

monitor health parameters to constantly improve the health of the population through rehabilitation and spa care, specifically targeting the patient's functional aspect of integration in everyday life, environment and work,

develop a decision support system and services based on the outdoors environment parameters and indoor location, and

reduce operations costs and improve quality of the services provided.

The *Indoor air quality improvement at school* pilot studied the use of a variety of indoor sensors and wearables combined with users’ personal feedback and environmental sensing information to provide a healthier living environment for pupils, teachers, and other staff members at school. The goal was to learn from sensor data by means of combining different data streams and applying environmental models, machine learning, data mining and supported by big data ICT. This will enable optimization of power provision and self-adaptive HVAC control and air purification in the house. The users can get personalized recommendations based on measured data and feedback got from the users on the impact of air quality.

The main objective of the *Tracking of professional / non-professional athletes with wearable sensors* pilot was to define an IoT infrastructure that integrates different sensor technologies. This infrastructure includes environmental factors and body basic state parameters that is collected and analyzed with machine learning and text mining methods. According to the results, the services and the applications that are able to provide guidance and direction can be developed.

As for the relation to other work in ESTABLISH project we have to emphasize that the task 6.4 reflects the work of all previous work packages (pilots’ specification, development and realization) and as for the validation processes the document D6.1 was created which provided the evaluation methodology in order to assess if all requirements were accomplished. This document represents the end of supply chain and gives the proof of the correct system design.

## Structure of this document

This deliverable 6.4 is composed of six sections, similarly to the first version. This introduction is followed by a section describing the applied evaluation methodology and separate sections with evaluation results for the three different pilots, and combined conclusions for the use case 3 evaluation.

# Applied methodology for evaluation

The evaluation methodology elaborated in D6.1 was adopted. The D6.1 recommends a list of generic and use case specific parameters that can be used for evaluation. The generic parameters are:

* Meeting general project objectives
* Meeting use case specific objectives
* System performance
* Privacy and security issues
* Feedback system performance
* System satisfaction

The specific parameters are:

* Self-awareness and self-adaptability
* Sensor system performance
* Visualization framework
* Management platform
* Big data management /processing
* Event processing
* Completeness of the use case
* Non-redundancy
* Operability
* Data sources utilization
* Storage, recording of the data
* Scope of the monitoring
* Monitoring applications
* Implementation
* Validation and integration within aimed domains
* Privacy and security
* Stakeholders’ Feedback
* IoT Interoperability
* Reliability of simulations / predictions
* Portability

More detailed description of the evaluation parameters with examples can be found in Table 1 of the deliverable D6.1. As a **Step 1** the deliverable D6.1 recommends to select relevant evaluation parameters for the use case in question. The following table presents the selected parameters for each pilot within use case 3. Use case 3 includes three pilots: 1) Rehabilitation decision support, 2) Indoor air quality improvement at school, and 3) Tracking of athletes with wearable sensors.

Table 1 Table of the relevant parameters within use case 3.

|  |  |  |  |
| --- | --- | --- | --- |
| Parameter | In use case 3 – pilot 1 | In use case 3 – pilot 2 | In use case 3 – pilot 3 |
| Self-awareness and self-adaptability | 🗷 | 🗷 |  |
| Meeting objectives (project main objectives and use case-specific objectives) | 🗷 | 🗷 |  |
| Systems’ performance | 🗷 | 🗷 |  |
| Sensor system performance | 🗷 | 🗷 |  |
| Visualization framework | 🗷 | 🗷 |  |
| Management platform | 🗷 | 🗷 |  |
| Big data management | 🗷 | 🗷 |  |
| Event processing | 🗷 |  |  |
| Use case implementation (Completeness of use case) | 🗷 | 🗷 |  |
| Redundancy percentage (Non-redundancy) |  |  |  |
| Usability | 🗷 | 🗷 |  |
| Data sources utilization | 🗷 | 🗷 |  |
| Storage, recording of the data | 🗷 | 🗷 |  |
| Scope of the monitoring | 🗷 |  |  |
| Monitoring applications | 🗷 |  |  |
| Implementation | 🗷 |  |  |
| Validation and integration within aimed domains | 🗷 |  |  |
| Privacy and security | 🗷 | 🗷 |  |
| Stakeholders’ feedback | 🗷 | 🗷 |  |
| IoT Interoperability | 🗷 | 🗷 |  |
| Reliability of simulations / predictions | 🗷 | 🗷 |  |
| Portability | 🗷 | 🗷 |  |
| Feedback system performance | 🗷 |  |  |
| System satisfaction | 🗷 | 🗷 |  |

In **Step 2**, the evaluator defines the methods used for evaluation of the above parameters. According to the chapter 2.2 of the deliverable D6.1 the evaluator

* chooses the criteria for the evaluation based on the guidelines in Step 1,
* determines the evidence that will be collected for each chosen criterion,
* determines the sample that will be used, if appropriate,
* determines ways to collect the evidence and,
* sets up a plan to collect the evidence.

In line with the chapter 2.3 (**Step 3**) of the deliverable D6.1, the evaluators

* informed the client of the scales that will be used to determine quality,

then the client

* developed the decision-making process to use the evaluated performance and
* set the standards that will be used in decision-making.

After the client has set the standards, the evaluators

* collected the evidence and
* documented the findings.

In **Step 4**, the findings are reported and appropriate decisions/proposed actions are described. The following chapters report the evaluation Steps 2-4 of the three pilots within the use case 3.

# Rehabilitation decision support

The test results will be reported as recommended in D6.1. Collecting data within the validation steps, described in the methodology, were processed and structured by the evaluator and client in the tables. Following action points were recommended for the evaluation meetings:

* The evaluator reports the information about the parameters to the client.
* The client checks the quality against the standards set in step 3.
* The client makes and implements decisions based on the findings.
* The client and/or evaluator documents the results appropriately.
* Either the evaluator or client reports the findings to the evaluatee, if appropriate.

## User level test results - system parameters evaluation done by client

The table 2 gives the results of the parameters evaluation done by client or system end user. The scaling for the evaluation is used as recommended in D6.1, the oral evaluation is easier for client and usually gives the supplier valuable information about how the system works in terms of user expectations.

Table 2. System evaluation results (client level)

|  |  |
| --- | --- |
| Parameter | In use case 3 \*/ pilot 1 |
| How the system copes with Self-awareness and self-adaptability | 3 |
| Meeting objectives | 4 |
| Systems’ performance | 4  Sporadic outage of the sensors, no outage in case of the whole system, system provides all functionalities available.  No evaluation in terms of number of users logged in neither number of sensors due to limited scope of the system. |
| Visualization framework | 4 |
| Management platform | 4 |
| Big data management | X |
| Event processing | 4 |
| Use case implementation | 4  Implementation of the system components the average technical skills, except sensors’ installation, which is done automatically. |
| Redundancy percentage | 4.5 |
| Usability | 4  Very user friendly and intuitive system |
| Data sources utilization | 4 |
| Storage, recording of the data | 4 |
| Implementation | 4 |
| Validation and integration within aimed domains | X |
| Privacy and security | 4  Login process into the system uses standard access procedure. No privacy nor security issues identified. |
| Stakeholders’ feedback | 3 |
| IoT Interoperability | 2 |
| Reliability of simulations / predictions | 5 |
| Portability | 3 |

### How the system copes with self-awareness and self-adaptability

|  |
| --- |
| **Description** |
| How the system consider to adapt when conditions are changed. |
| **Methods** |
| **Researcher** evaluates the success of each goal by 100%, if goal was fully met and 0% if not. Then the average percentage of the successfully met objectives is calculated. |
| **Accomplishment Scale** |
| 1) Bad (0-20%)  2) Poor (21-40%)  3) Medium (41-60%)  4) Good (61-80%)  5) Excellent (81-100%) |
| **Results** |
| Adaptability when sensor data is changed 70%  Adaptability when patients are changed 40%  Average 55% 🡪 **3 (medium)** |

### Meeting objectives

|  |
| --- |
| **Description** |
| The pilot 1 will combine environmental sensor data with physiological and behavioural sensor data to empower patients in a rehabilitation clinic with decision support tools for behavioural choices and treatment options. |
| **Methods** |
| **Researcher** evaluates the success of each goal by 100%, if goal was fully met and 0% if not. Then the average percentage of the successfully met objectives is calculated. |
| **Accomplishment Scale** |
| 1) Bad (0-20%)  2) Poor (21-40%)  3) Medium (41-60%)  4) Good (61-80%)  5) Excellent (81-100%) |
| **Results** |
| 1. Sensor data was visualised for indoor and outdoor. Current values and a history of values. 🡪 80% 2. Patient data was displayed for patients for a selected activity (75%) 3. Requirements were successfully collected via questionnaires and also additional interviews 🡪80%   🡪 average 77% 🡪 **4 (good)** |

### Systems’ performance

|  |
| --- |
| **Description** |
| Evaluation of the systems’ overall performance in terms of completeness of sensor set, robustness, quality of produced data. How well different components work in real-world environment. |
| **Methods** |
| **Researcher** evaluates the performance of the system as a whole on a scale 1-5 (1=bad, 5=Excellent).  Also, the **participant** evaluated the system performance at the end. SUS scale is used and the output gives values from 0 to 100. |
| **Accomplishment Scale** |
| *Expert* evaluation  1) Bad  2) Poor  3) Medium  4) Good  5) Excellent  *Participant* evaluation  1) Bad (SUS average 0-20)  2) Poor (SUS average 21-40)  3) Medium (SUS average 41-60)  4) Good (SUS average 61-80)  5) Excellent (SUS average 81-100) |
| **Results** |
| **Researcher** evaluation for the system performance is **4**: the sensor set was complete (5), system was fairly robust (4), quality of the produced data was good (4), network was working mostly ok (5), data components worked ok in real-world environment (4).  **Participant** evaluation for the whole system on average was 63 (SUS scale 0-100) 🡪 **4 (good)** |

### Visualization framework

|  |
| --- |
| **Description** |
| Visualizations of the pilot data are implemented with Web application developed by Siveco, and Grafana. The pilot participants are provided with graphical, as well as, numerical values measured by the outdoor, and indoor air quality sensors and collected with IoT system. Also the evolution of patient biometric data is presented graphically and text based. A calendar is used to plan activities for both patients and caregivers.  Evaluation of how well the visualizations meet the end-users’ needs and how well it performs data search, interpretation, and comparing of air quality parameters. |
| **Methods** |
| Pilot **participants** evaluate how useful the visualisations are for them with scale 1 to 10, 1 being fully useless and 10 being extremely useful. Average answer is calculated. |
| **Accomplishment Scale** |
| 1) Bad (when 1-2)  2) Poor (when 3-4 )  3) Medium (when 4.1-6)  4) Good (when 6.1-8)  5) Excellent (when 8.1-10) |
| **Results** |
| **Caregivers (Kineto)** evaluation for the visualisation framework was on average 8.5 🡪 4 (excelent)  No evaluations were registered for patients. Patient usage was tested also by caregivers. The result was 6.2 🡪 **4 (good)** |

### Management platform

|  |
| --- |
| **Description** |
| Management of system’s entities – operators, caregivers, patients, organizing logs, warnings, remote access possibilities, entering of data, user-friendliness, etc. |
| **Methods** |
| **Researcher’s** rough estimation of the platform’s   1. costs / year / user vs target costs (660-1600€/ year/user[[1]](#footnote-1)) (AWS, Establish web app, network connection, sensors, air purifiers, installation, maintenance…), 2. the level of automatic warnings related to data values, data losses, lost data connections etc should be considered, 3. Remote access. |
| **Accomplishment Scale** |
| 1) Bad (average between 1-1.8)  2) Poor (average between 1.9-2.6)  3) Medium (average between 2.7-3.4)  4) Good (average between 3.5-4-2)  5) Excellent (average between 4.3-5) |
| **Results** |
| **Researcher** evaluation for the management platform was   * Costs (AWS=2000€/year, 60€/year, IoT) * Automatic warnings (abnormal data values, data losses, lost data connections): There alerts and notifications in the system for all parameters and also for complex conditions and rules. * Average **4 (good)** |

### Big data management

There is no BigData usage. Data captured is stored in RDBMS (Postgres). The usage of postgres was evaluated as excellent (5) by researcher.

### Event processing

|  |
| --- |
| **Description** |
| Events are considered that related to   * Outdoor data registration * Indoor data registration * Patient data registration |
| **Methods** |
| **Researcher** evaluates how events are processed 1-5 (1=bad, 5=Excellent).  Also, the **caregivers (Kineto)** evaluated the system performance at the end. SUS scale is used and the output gives values from 0 to 100. |
| **Accomplishment Scale** |
| *Expert* evaluation  1) Bad  2) Poor  3) Medium  4) Good  5) Excellent  *Participant* evaluation  1) Bad (SUS average 0-20)  2) Poor (SUS average 21-40)  3) Medium (SUS average 41-60)  4) Good (SUS average 61-80)  5) Excellent (SUS average 81-100) |
| **Results** |
| **Researcher** evaluation for the event processing is **5 (excellent)**: the sensor set was complete (5), system was fairly robust (4), quality of the produced data was excellent (5), network was working mostly ok (5), components worked ok in real-world environment (5).  **Caregiver** evaluation for the whole system on average was 63 (SUS scale 0-100) 🡪 **4** **(Good)** |

### Use case implementation

|  |
| --- |
| **Description** |
| Evaluate the progress of all the components-subsystems to be developed and the integration between them  Main components to be evaluated are:   * air quality sensors * physiological sensor * calendar * self-reporting application * recommendations/automatic * patient activities * visualizations/feedback module * data collection backend |
| **Methods** |
| **Researcher** evaluates implementation of each module by scale 1 to 5. 5 is given if the module is implemented as planned and 1 if not. Evaluation of the sensor systems as a whole is given as an average of different module evaluations.  **Caregiver (kineto)** evaluates implementation of each module by scale 1 to 5. 5 is given if the module is implemented as planned and 1 if not. Evaluation of the sensor systems as a whole is given as an average of different module evaluations. |
| **Accomplishment Scale** |
| 1) Bad (average between 1-1.8)  2) Poor (average between 1.9-2.6)  3) Medium (average between 2.7-3.4)  4) Good (average between 3.5-4-2)  5) Excellent (average between 4.3-5) |
| **Results** |
| **Researcher, Tester** evaluation for the use case implementation was **5:**   1. air quality sensors: installed and integrated to the system (5) 2. physiological sensor: installed and integrated to the system (5) 3. calendar (5) 4. self-reporting application: installed and integrated to the system (5) 5. recommendations automatic/(3) 6. visualizations/feedback module: installed and integrated to the system (4)  * average **4,5 (good)**   **Caregiver** evaluation for the use case implementation was **4:**   1. air quality sensors: installed and integrated to the system (5) 2. physiological sensor: installed and integrated to the system (4) 3. calendar (5) 4. self-reporting application: installed and integrated to the system (4) 5. recommendations automatic/(3) 6. visualizations/feedback module: installed and integrated to the system (3)  * average **4.1 (good)** |

This chapter deals with system parameters test performance and then evaluation summary done by tester specialist. In the first round the application prototype was tested by the involved QA (Quality Assurance) teams, in the next phase a representative of the customer or end user will be asked for test performance. Test results of the both test rounds will be compared in order to make evaluation as objective as possible.

Using test cases forms provided in the D3.3 - Test plan – the tester specialist made the detail validation of all set up parameters of the system. The table 3 provides the list of tests results and table 4 shows the results of the tests performance.

Table 3. List of tests for the system evaluation

|  |  |  |  |
| --- | --- | --- | --- |
| **Test Case ID** | **Test Case name** | **Components affected** | **Function addressed** |
| TS.GUEST.01-TC.01 | Enter as Guest | Web application | Login |
| TS.GUEST.02-TC.01 | View “Current Data” tab information | Web application | Current Data |
| TS.GUEST.02-TC.02 | View “Environmental Data” tab information | Web application | Environmental Data |
| TS.GUEST.03-TC.01 | Sign up as a Patient | Web application | Sign up |
| TS.GUEST.03-TC.02 | Sign up as a Kineto | Web application | Sign up |
| TS.USER.01-TC.01 | Login as a Patient | Web application | Login |
| TS.USER.01-TC.02 | Login as a Kineto | Web application | Login |
| TS.USER.02-TC.01 | View “Current Data” tab information as a patient | Web application | Current Data visualization |
| TS.USER.02-TC.02 | View “Current Data” tab information as a kineto | Web application | Current Data visualization |
| TS.USER.02-TC.03 | “Environmental Data” tab information management | Web application | Environmental Data |
| TS.USER.02-TC.04 | “Physiological Data” tab information management | Web application | Physiological Data |
| TS.USER.02-TC.05 | Dashboard alerts management for patient | Web application | Dashboard Alerts |
| TS.USER.02-TC.06 | Dashboard alerts management for kineto | Web application | Dashboard Alerts |
| TS.USER.03-TC.01 | “Alerts” management for patient | Web application | Management alerts |
| TS.USER.03-TC.01 | “Alerts” management for kineto | Web application | Management alerts |
| TS.USER.04-TC.01 | View assigned activities | Web application | Activities visualization |
| TS.USER.04-TC.02 | View assigned activities | Web application | Activities visualization |
| TS.USER.04-TC.03 | Managing activities assigned to a patient | Web application | Activities management |
| TS.USER.04-TC.04 | Create new activity | Web application | Activities management |
| TS.USER.05-TC.01 | View patient list | Web application | Patients list Visualization |
| TS.USER.05-TC.02 | Edit patient’s data | Web application | Patients management |
| TS.USER.05-TC.03 | Add new patient | Web application | Patients management |
| TS.USER.06-TC.01 | View current account settings | Web application | Settings |

Table 4. Results of the tests of system evaluation

|  |  |  |  |
| --- | --- | --- | --- |
| **Test Case ID** | **Test Case name** | **Result of test** | **Remark** |
| TS.GUEST.01-TC.01 | Enter as Guest | OK |  |
| TS.GUEST.02-TC.01 | View “Current Data” tab information | OK |  |
| TS.GUEST.02-TC.02 | View “Environmental Data” tab information | OK |  |
| TS.GUEST.03-TC.01 | Sign up as a Patient | OK |  |
| TS.GUEST.03-TC.02 | Sign up as a Kineto | OK |  |
| TS.USER.01-TC.01 | Login as a Patient | OK |  |
| TS.USER.01-TC.02 | Login as a Kineto | OK |  |
| TS.USER.02-TC.01 | View “Current Data” tab information as a patient | OK |  |
| TS.USER.02-TC.02 | View “Current Data” tab information as a kineto | OK |  |
| TS.USER.02-TC.03 | “Environmental Data” tab information management | OK |  |
| TS.USER.02-TC.04 | “Physiological Data” tab information management | OK |  |
| TS.USER.02-TC.05 | Dashboard alerts management for patient | OK |  |
| TS.USER.02-TC.06 | Dashboard alerts management for kineto | OK |  |
| TS.USER.03-TC.01 | “Alerts” management for patient | **IP** | In progress, will be done in v2 version of the deliverable, tests are not finalized in the day of the deliverable release |
| TS.USER.03-TC.01 | “Alerts” management for kineto | **IP** | In progress, will be done in v2 version of the deliverable, tests are not finalized in the day of the deliverable release |
| TS.USER.04-TC.01 | View assigned activities | OK |  |
| TS.USER.04-TC.02 | View assigned activities | OK |  |
| TS.USER.04-TC.03 | Managing activities assigned to a patient | OK |  |
| TS.USER.04-TC.04 | Create new activity | OK |  |
| TS.USER.05-TC.01 | View patient list | OK |  |
| TS.USER.05-TC.02 | Edit patient’s data | OK |  |
| TS.USER.05-TC.03 | Add new patient | OK |  |
| TS.USER.06-TC.01 | View current account settings | OK |  |

### Redundancy percentage

|  |
| --- |
| **Description** |
| Redundancy of data and redundancy of actions are measured. |
| **Methods** |
| **Researchers** assess how much data is redundant, and how much action are redundant. |
| **Accomplishment Scale** |
| % of the data or actions redundancy  1) Bad (81-100%)  2) Poor (61-80%)  3) Medium (41-60%)  4) Good (21-40%)  5) Excellent (0-20%) |
| **Results** |
| **Researcher** evaluation for the redundancy was **4.5 (good)**  The visualisation is made from the indoor air quality data and self-reports, and for the physiological data, 🡪 75% ->4  Redundant data sources 20% ->5  Redundant activities 30% ->4 |

### Usability

|  |
| --- |
| **Description** |
| Evaluate the operability of the pilot components considering how simple it is to operate the system.  Main components to be evaluated are:   * air quality sensors * physiological sensor * reporting application * visualisations * calendars |
| **Methods** |
| **Researcher** evaluates the operability of each of the above-mentioned component with scale 1 to 5, 1 being bad, 5 being excellent. Average is calculated. |
| **Accomplishment Scale** |
| 1) Bad (average between 1-1.8)  2) Poor (average between 1.9-2.6)  3) Medium (average between 2.7-3.4)  4) Good (average between 3.5-4-2)  5) Excellent (average between 4.3-5) |
| **Results** |
| **Researcher** evaluation for the operability was **4 (good)**   * air quality sensors - easy to operate (5) * physiological sensor - not very easy (3) * reporting application - easy to operate (4) * visualisations easy to operate (4) * calendars very easy to operate (5) |

### Data sources utilization

|  |
| --- |
| **Description** |
| Are data sources used in optimal way? Do we have overloaded data sources? Do we have space for system optimization in this respect? |
| **Methods** |
| **Researchers** assess how much of the collected data (%) is utilized in the current feedback system. |
| **Accomplishment Scale** |
| % of the data source utilization.  1) Bad (0-20%)  2) Poor (21-40%)  3) Medium (41-60%)  4) Good (61-80%)  5) Excellent (81-100%) |
| **Results** |
| **Researcher** evaluation for the data sources utilisation was **4.3 (good)**  The visualisation is made from the indoor air quality data and self-reports, and for the physiological data, 🡪 75% ->4  Overloaded data sources 95% ->5  Space for optimizations 30% ->4 |

### Storage, recording of the data

|  |
| --- |
| **Description** |
| What percentage of the stored data is useful and reliable? How much of the collected data can be used in further analysis. |
| **Methods** |
| **Researcher’s** rough estimation of the good quality data in scale from 1 to 5, 5 being ‘all data was useful and reliable’ and 1 ‘the data was useless/ impossible to be used’. |
| **Accomplishment Scale** |
| 1) Bad (average between 1-1.8)  2) Poor (average between 1.9-2.6)  3) Medium (average between 2.7-3.4)  4) Good (average between 3.5-4-2)  5) Excellent (average between 4.3-5) |
| **Results** |
| **Researcher** evaluation for the storage and recording of data was **4 (good)**   1. air quality sensor data (noise, humidity, pressure, temperature…): (5) 2. physiological sensor: data was 3 3. calendar 5  * average 4 |

### Implementation

|  |
| --- |
| **Description** |
| Evaluate the operability of the pilot components considering how simple it is to operate and maintain and whether there was problems.    Main components to be evaluated are:   * air quality sensors * physiological sensor * reporting application * visualisations * calendars |
| **Methods** |
| **Researcher** evaluates the operability of each of the above-mentioned component with scale 1 to 5, 1 being bad, 5 being excellent. Average is calculated. |
| **Accomplishment Scale** |
| 1) Bad (average between 1-1.8)  2) Poor (average between 1.9-2.6)  3) Medium (average between 2.7-3.4)  4) Good (average between 3.5-4-2)  5) Excellent (average between 4.3-5) |
| **Results** |
| **Researcher** evaluation for the operability was **4 (good)**   * air quality sensors - easy to operate (5) * physiological sensor - not very easy (3) * reporting application - easy to operate (4) * visualisations easy to operate (4) * calendars very easy to operate (5) |

### Validation and integration within aimed domains

There was no integration in other domain. The system is working standalone.

### Privacy and security

|  |
| --- |
| **Description** |
| How well the system privacy and security is ensured?   * network security: encryption * interface security: authentication/access control * data security: backup, encryption, data isolation * legal issues security: data storage location, data sanitization, data management after project lifecycle, etc. |
| **Methods** |
| **Researcher** evaluates the security categories/attributes described above with scale 1 to 5. 5 is given f the category security is fully covered and 1 if not. Evaluation of the privacy and security as a whole is given as an average of different category evaluations. |
| **Accomplishment Scale** |
| 1) Bad (average between 1-1.8)  2) Poor (average between 1.9-2.6)  3) Medium (average between 2.7-3.4)  4) Good (average between 3.5-4-2)  5) Excellent (average between 4.3-5) |
| **Results** |
| **Researcher** evaluation privacy and security issues is **4 (good)**:   * network security: encryption (5) * interface security: authentication/access control (4) * data security: backup, encryption, data isolation (data back-up and encryption existed provided by Azure and limited access by admin) (5) * legal issues security: data storage location, data sanitization, data management after project lifecycle, etc. (data storage locates in Ireland in secure place, data was pseudonymised when stored, the data is stored behind password) (4) |

### Stakeholders’ feedback

|  |
| --- |
| **Description** |
| How satisfied the participants were with the pilot as a whole |
| **Methods** |
| **Pilot participants’** satisfaction for the entire system as a whole is collected though NSP (net promoters score) on a scale 0-10 at the end of pilot. Score gives output based on all answers between -100 to 100. |
| **Accomplishment Scale** |
| 1) Bad (when NPS -100- 61)  2) Poor (when NPS -60 – -21 )  3) Medium (when NPS -20- +20)  4) Good (when NPS 21-60)  5) Excellent (when NPS 61-100) |
| **Results** |
| **Participant** evaluation of the entire system satisfaction was 0 on NPS scale 🡪 **3 (medium)** |

### IoT Interoperability

|  |
| --- |
| **Description** |
| Ease of using other types of sensors and communication protocols, or the ease of using several IoT platforms at the same time. |
| **Methods** |
| **Researcher’s** rough estimation (%) on how much of the current implementation allows integration of other types of sensors/communication protocols/IoT platforms. |
| **Accomplishment Scale** |
| 1) Bad (0-20%)  2) Poor (21-40%)  3) Medium (41-60%)  4) Good (61-80%)  5) Excellent (81-100%) |
| **Results** |
| **Researcher** evaluation for the IoT operability was 25 since new sensors can be adapted if they comply MQTT protocol (50%)  Wearable are difficult to adapt (1%) |

### Reliability of simulations / predictions

|  |
| --- |
| **Description** |
| Predict the value of a biological parameter based on outdoor conditions, indoor conditions, activities |
| **Methods** |
| Employing machine learning method (Regression based on Decision trees with Giny impurity and random forests) with person-specific training (*i.e.*, model of each individual was trained using only his/her data) to classify user (patient) evolution under some specific circumstances. |
| **Accomplishment Scale** |
| 1) Bad (0-20%)  2) Poor (21-40%)  3) Medium (41-60%)  4) Good (61-80%)  5) Excellent (81-100%) |
| **Results** |
| Experimental results with real life data, collected in "Centrul stelutelor" during 12 weeks, show that is possible to predict values with an accuracy of 85%  **Researcher** evaluation for employing machine learning classifier Decision Tree) to classify personal Data 🡪 **5 (excellent)** |

### Portability

|  |
| --- |
| **Description** |
| Understanding portability, how easy it is to adapt and apply the pilot to other groups of persons. How coupled the software systems is developed? |
| **Methods** |
| **Researcher’s** rough estimation (%) on how much of the current implementation can be directly applied to other cities/buildings. |
| **Accomplishment Scale** |
| 1) Bad (0-20%)  2) Poor (21-40%)  3) Medium (41-60%)  4) Good (61-80%)  5) Excellent (81-100%) |
| **Results 3** |
| **Researcher** evaluation for the portability was 3, system can be implemented in another place under some conditions, a.i  Installation of outdoor and indoor sensors  Usage of the same wearable (Garmin)  Installation on a sever which supports Java 1.8 |

## Restrictions of the evaluation processes

The evaluation processes did not cover long term evaluation e.g. user long term experience, training program, maintenance, because of the limited time of the use cases operation within work package 6. As we expected the use cases were in operation, on the component level, starting M21, thus we were able to provide this report in version 1. Full system operation we plan in the month 28 when we expect to start second phase of the evaluation resulting in the report version 2.

### Achieved results after the evaluation process

The Romanian Consortium within the ESTABLISH project, BEIA CONSULT INTERNATIONAL and SIVECO, organized the first workshop within the use case 2 in order to evaluate the functionalities of the rehabilitation ESTABLISH platform, but also to identify the potential for improving the services developed within it.

The event was organized on November 16th, 2018 at the premises of the NGO Association M.A.M.E., where the rehabilitation pilot is also held. The results obtained by questioning the people present at the workshop were centralized in a POWER BI dashboard[[2]](#footnote-2) for graphical interpretation of the answers obtained.  
The questionnaire contained 24 questions, 31 people of different ages, gender, and occupation were questioned. The centralized results in the PowerBi dashboard are shown in the figures below;   
and, the centralized answers from the first section show the following statistics:

* 70% of the persons surveyed are over 25 years old,
* 45% of those surveyed considered that the air quality in the area where they live is moderate,

according to 32% of those surveyed, the period of the day in which they feel the pollution phenomenon is in the evening.

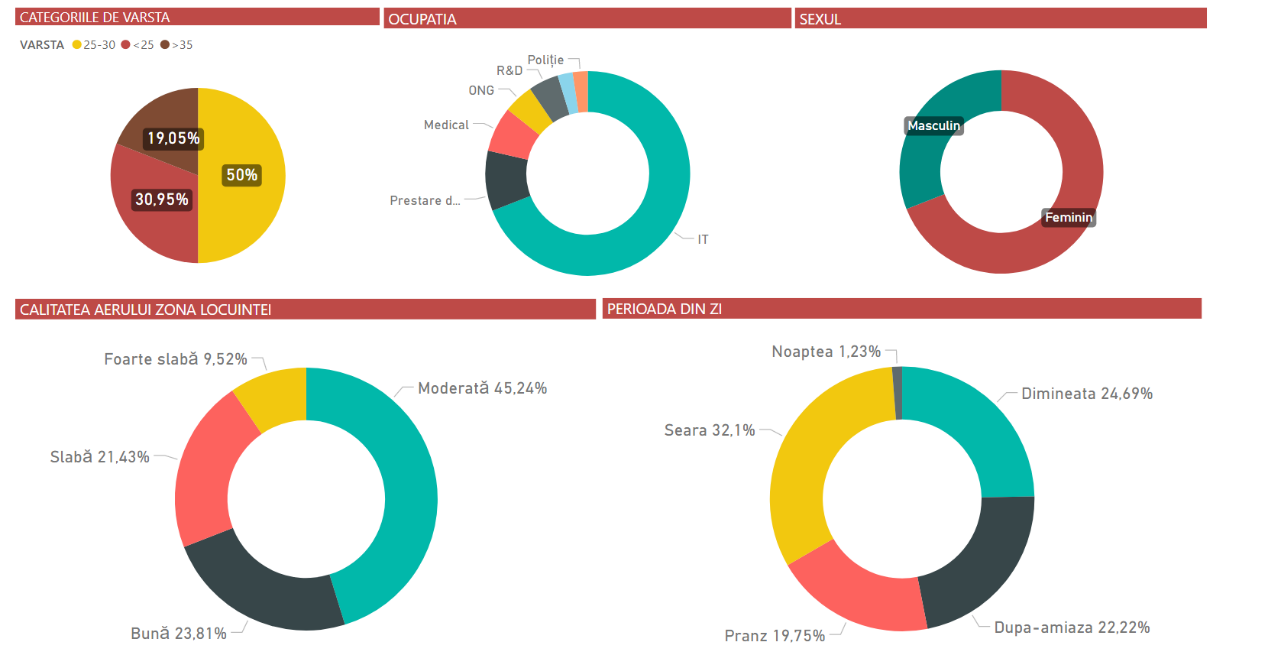


Figure 1. Graphic representation of the answers of the ESTABLISH Workshop questionnaire (section 1 of the questionnaire)

The centralized answers in the second section of the questionnaire present the following statistics:

* the risk categories, according to the persons questioned, are children and persons suffering from respiratory diseases,
* the main source of pollution is due to car traffic,
* the main polluting factors specific to the outdoor environment are considered to be the suspended particles (PMs), carbon dioxide and dust,
* the main polluting factor specific to the indoor environment is considered to be dust, and the phenomenon felt is realted to the sensation of lack of air (low amount of oxygen).

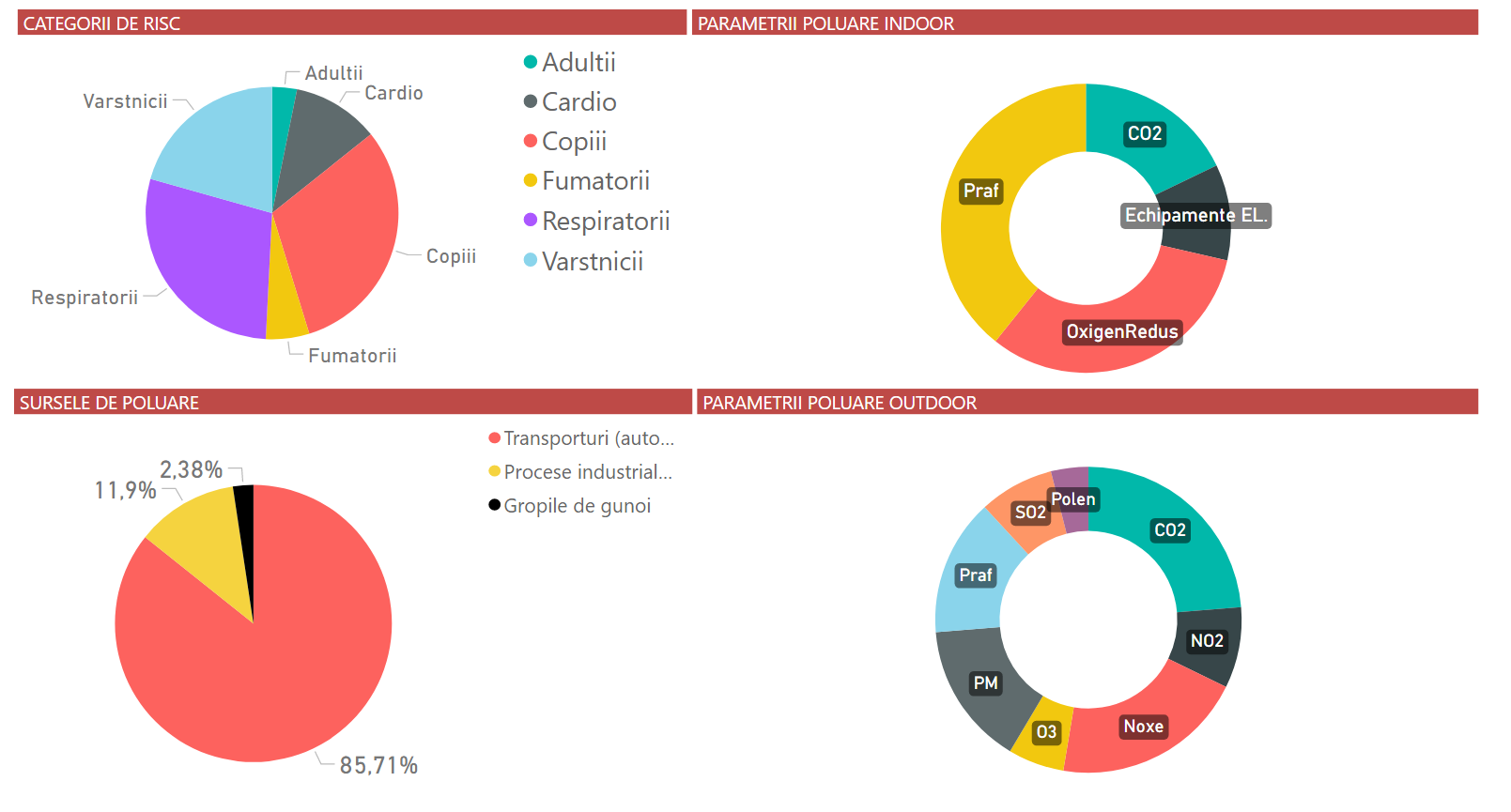


Figure 2. Graphic representation of the answers of the ESTABLISH Workshop questionnaire (section 2 of the questionnaire).

The centralized answers in the third section of the questionnaire present the following statistics:

* the main source of information regarding air pollution is the Internet,
* the quality of the information available regarding the air quality is appreciated in a positive way,
* only 21% of respondents use mobile apps to inquire regarding the air quality in their living area,
* the frequency of information is rare, "when they have the opportunity to read about this topic",
* the frequency of notifications, preferred, regarding the information presenting the air quality is weekly or in the case in which serious episodes of pollution are manifested.

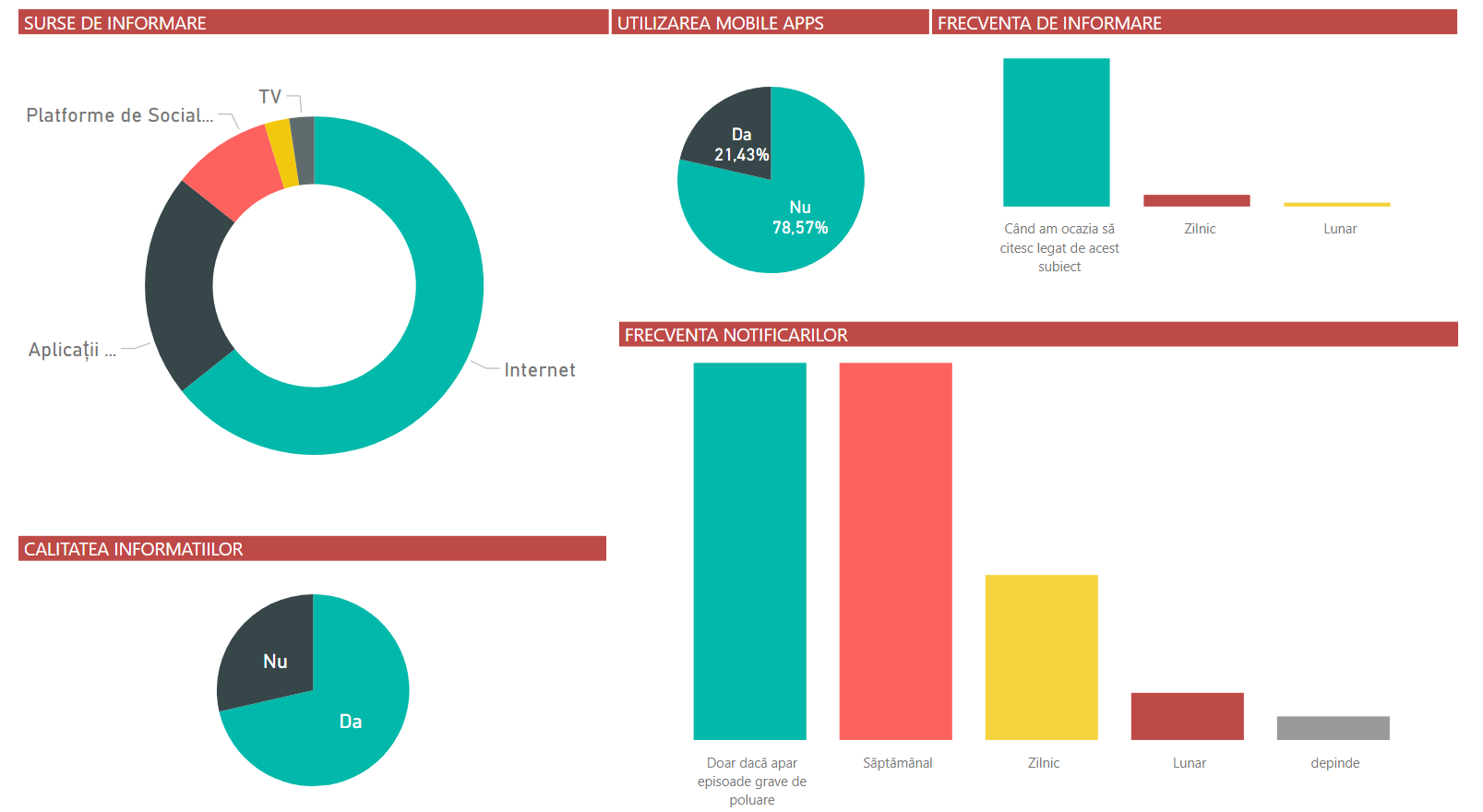


Figure 3. Graphic representation of the answers of the ESTABLISH Workshop questionnaire (section 3 of the questionnaire).

The centralized answers from the fourth section of the questionnaire offer a perspective over the graphical representation of air quality information, we observed the following aspects:

* The preferred display tools are the graphs in scale, colored, where the red color represents poor air quality, and the green color expresses good air quality.

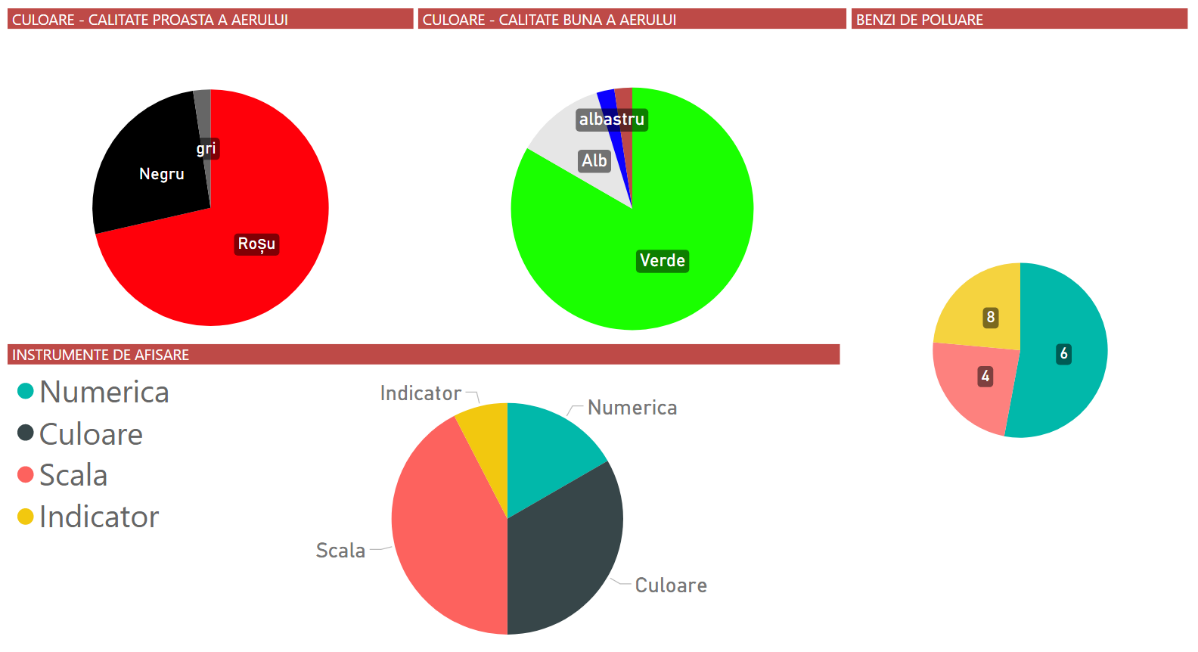


Figure 4. Graphic representation of the answers of the ESTABLISH Workshop questionnaire (section 4 of the questionnaire).

The centralized answers from the last section of the questionnaire present the following statistics:  
- most of the participants at the survey appreciated the display and the availability on the platform of visual instruments that present almost real-time information regarding air quality, but also the patient management component,

- 85% of the respondents considered the platform useful for the implementation of rehabilitation pilot,  
- 57% have smart watches and 60% of them would agree with data sharing in order to benefit from smart health medical services,

- the preferred payment method to access the rehabilitation ESTABLISH platform and its services is a monthly subscription.

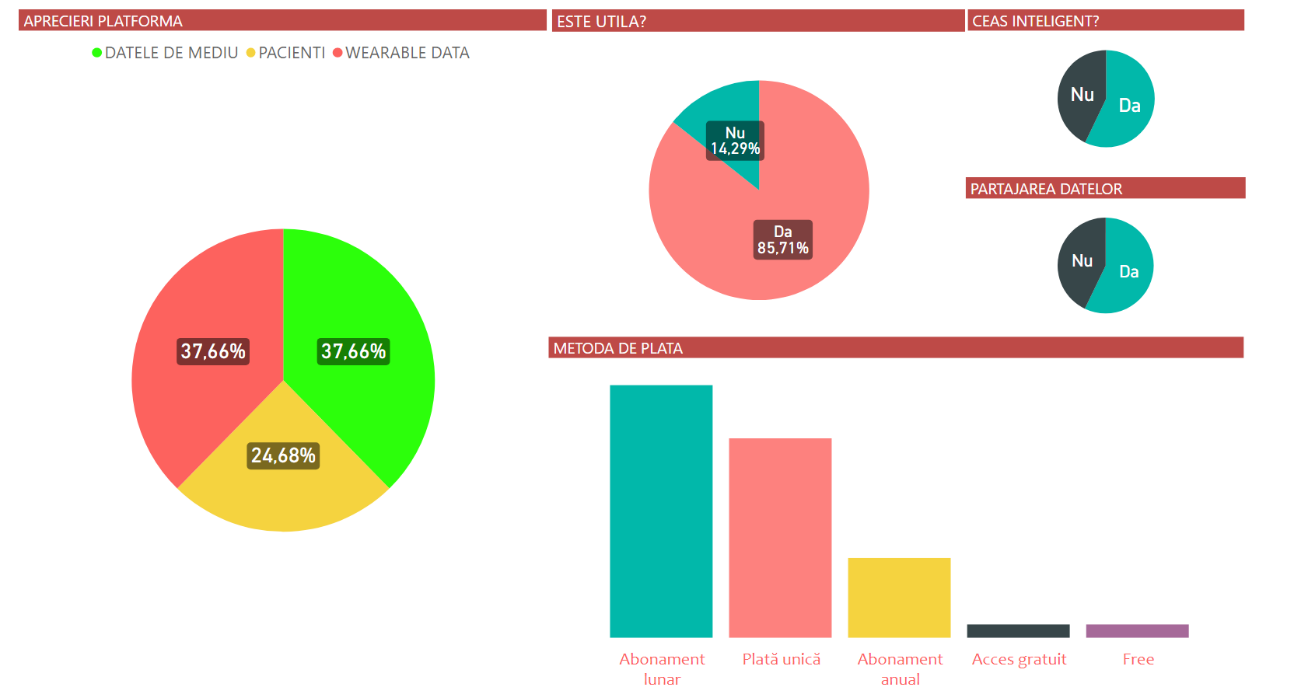


Figure 5. Representation in graphical form of the answers of the ESTABLISH Workshop questionnaire (section 5 of the questionnaire).

Based on the information presented **in D5.4 Report on techniques and algorithms for simulation**

The decision support component uses a complex structure of hardware and software subcomponents.  
  
The hardware subcomponents of the decision support tool are:

* the Backend Infrastructure,
* the Environmental sensors - Libelium Plug and Sense (outdoor) and Libelium Gas Board (indoor),
* the Gateway equipment – Meshlium IoT Gateway.

1. The software sub-components consist of:

* prerequisites for data analysis and data presentation: LAMP package (Linux, Apache, MySQL, PHP) was used.

The steps taken to implement the decision support tool are:

* 1. Data extraction from the Meshlium's MySQL database.
  2. After collecting the data from the GW database, we applied a specific query in order to calculate the Hourly Moving Average, for each of the three parameters (NO2, PM10, PM2.5).
  3. Afterwards, we applied the AQI calculation algorithm, basically we achieve the Kmax identification and AQI index calculation.
  4. We store the values obtained in the previous phase.
  5. Perform the Graphical display of the values of the AQI index for the last 24 hours and presentation of the decision support tool with the specific notification[[3]](#footnote-3).

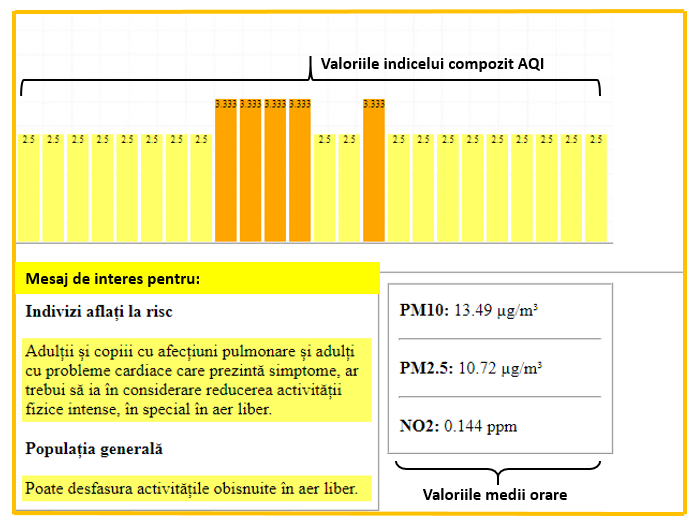


Figure 6. Graphical presentation of the decision support component (tool).

# Indoor air quality improvement at school

## Parameters and methods

This section describes how evidence is collected for each parameter (Step 2) selected for **Indoor air quality improvement at school** pilot. The following tables also specify the types of measures, meters, scales and criteria/quality standards (Step 3) used in the evaluation.

### Meeting objectives

|  |
| --- |
| **Description** |
| Indoor air quality improvement at school pilot objectives   1. Converting sensor data into actionable information in terms of healthier and safer environment. 2. Improving well-being (or maintaining the good level) 3. Collect requirements for the service that advances better indoor air quality and health 4. Develop and intelligent service that produces tailored real-time feedback combining information from indoor air sensors, wearable devices and self-reports 5. Study the correlation between objective measurements (indoor air quality, physiological measurements, concentration tasks) and subjective evaluations (symptoms, indoor air quality, stress) |
| **Methods** |
| **Researcher** evaluates the success of each goal by 100%, if goal was fully met and 0% if not. Then the average percentage of the successfully met objectives is calculated. |
| **Accomplishment Scale** |
| 1) Bad (0-20%)  2) Poor (21-40%)  3) Medium (41-60%)  4) Good (61-80%)  5) Excellent (81-100%) |
| **Results** |
| 1. Sensor data was visualised. However, the people were not exactly sure how to use the information. The interpretation and suggestions were missing 🡪 50% 2. On average, the reported symptoms were slightly changed to the worse (12/31symptoms improved, 6 improved and 3 remained the same) and the participants were hopeful that it had positive effects 🡪 40% 3. Requirements were successfully collected via questionnaires and also additional interviews 🡪100% 4. Tailored real-time feedback was given from the indoor sensors and number of the self-reported symptoms but data from wearables was not visualised 🡪 60% 5. Correlation between objective measurements and subjective evaluations were calculated 🡪 100%   🡪 average 70% 🡪 **4 (good)** |

### Systems’ performance

|  |
| --- |
| **Description** |
| Evaluation of the systems’ overall performance in terms of completeness of sensor set, robustness, quality of produced data, big data network and data analytics. How well different components work in real-world environment. |
| **Methods** |
| **Researcher** evaluates the performance of the system as a whole on a scale 1-5 (1=bad, 5=Excellent).  Also, the **participant** evaluated the system performance at the end. SUS scale is used and the output gives values from 0 to 100. |
| **Accomplishment Scale** |
| *Expert* evaluation  1) Bad  2) Poor  3) Medium  4) Good  5) Excellent  *Participant* evaluation  1) Bad (SUS average 0-20)  2) Poor (SUS average 21-40)  3) Medium (SUS average 41-60)  4) Good (SUS average 61-80)  5) Excellent (SUS average 81-100) |
| **Results** |
| **Researcher** evaluation for the system performance is **4 (good)**: the sensor set was complete (5), system was fairly robust (3), quality of the produced data was good (4), network was working mostly ok (5), data analytics medium (3), components worked ok in real-world environment (3).  **Participant** evaluation for the whole system on average was 24.5 (SUS scale 0-100) 🡪 **2 (poor)** |

### Privacy and security

|  |
| --- |
| **Description** |
| How well the system privacy and security is ensured?   * network security: encryption * interface security: authentication/access control * data security: backup, encryption, data isolation * legal issues security: data storage location, data sanitization, data management after project lifecycle, etc. |
| **Methods** |
| **Researcher** evaluates the security categories/attributes described above with scale 1 to 5. 5 is given f the category security is fully covered and 1 if not. Evaluation of the privacy and security as a whole is given as an average of different category evaluations. |
| **Accomplishment Scale** |
| 1) Bad (average between 1-1.8)  2) Poor (average between 1.9-2.6)  3) Medium (average between 2.7-3.4)  4) Good (average between 3.5-4-2)  5) Excellent (average between 4.3-5) |
| **Results** |
| **Researcher** evaluation privacy and security issues is **4 (good)**:   * network security: encryption (5) * interface security: authentication/access control (3) * data security: backup, encryption, data isolation (data back-up and encryption existed provided by Azure and limited access by admin) (5) * legal issues security: data storage location, data sanitization, data management after project lifecycle, etc. (data storage locates in Ireland in secure place, data was pseudonymised when stored, the data is stored behind password) (4) |

### Feedback system performance

|  |
| --- |
| **Description** |
| Feedback system is realized as visualizations of the pilot data implemented with Power BI. The pilot participants are provided with graphical, as well as, numerical values measured by the indoor air quality sensors and collected with self-reporting application. The usability of the visualisations is evaluated. |
| **Methods** |
| Usability aspects of the visualization system are evaluated by the pilot **participants** at the end of the project with selected items from System Usability Scale (SUS) and other usability statements with 5-point Likert scale (1= strongly disagree, 5=strongly agree).  The questions included were:   1. I thought the visualisation application was easy to use (SUS) 2. I would imagine that most people would learn to use this visualisation application very quickly (SUS) 3. I felt very confident using the visualisation application (SUS) 4. The visualisation application was working robustly 5. It was easy to understand the visualisations 6. It was easy to find the information I wanted from the visualisations 7. The login to the visualisation app was easy through the link 8. Loading the visualisations took long time   The average of answers for each question is calculated and then average of the question averages is calculated. |
| **Accomplishment Scale** |
| 1) Bad (average between 1-1.8)  2) Poor (average between 1.9-2.6)  3) Medium (average between 2.7-3.4)  4) Good (average between 3.5-4-2)  5) Excellent (average between 4.3-5) |
| **Results** |
| **Participant** evaluation for the feedback system performance on average was 3.2 🡪 **3 (medium)** |

### System satisfaction

|  |
| --- |
| **Description** |
| Evaluate how satisfied the participants were with the system as a whole? |
| **Methods** |
| Pilot **participants’** satisfaction for the system as a whole is collected with NSP (net promoters score) on a scale 0-10 at the end of pilot. Score gives output based on all answers between -100 to 100. |
| **Accomplishment Scale** |
| 1) Bad (when NPS -100- 61)  2) Poor (when NPS -60 – -21 )  3) Medium (when NPS -20- +20)  4) Good (when NPS 21-60)  5) Excellent (when NPS 61-100) |
| **Results** |
| **Participant** evaluation for the system satisfaction was 0 (NPS -100- +100) 🡪 **3 (medium)** |

### Sensor system performance

|  |
| --- |
| **Description** |
| Evaluation of the sensor systems’ performance. Expert/researcher evaluation of each sensor system in terms of battery life/power consumption, data quality, measurement accuracy, etc.   1. air quality sensors 2. physiological sensor 3. self-reporting application |
| **Methods** |
| **Researcher** evaluates the performance of each module by scale from 1 to 5. 5 is given if the module works as planned and 1 if not. Evaluation of the sensor systems as a whole is given as an average of different module evaluations. |
| **Accomplishment Scale** |
| 1) Bad (average between 1-1.8)  2) Poor (average between 1.9-2.6)  3) Medium (average between 2.7-3.4)  4) Good (average between 3.5-4-2)  5) Excellent (average between 4.3-5) |
| **Results** |
| **Researcher** evaluation for the sensor system performance was **4 (good)**:   1. air quality sensors (battery life/power consumption, data quality, measurement accuracy) (5) 2. physiological sensor (battery life/power consumption, data quality, measurement accuracy) (2) 3. self-reporting application (battery life/power consumption, data quality, measurement accuracy) (5) |

### Visualization framework

|  |
| --- |
| **Description** |
| Visualizations of the pilot data are implemented with Power BI. The pilot participants are provided with graphical, as well as, numerical values measured by the indoor air quality sensors and collected with self-reporting application.  Evaluation of how well the visualizations meet the end-users’ needs and how well it performs data search, interpretation, and comparing of air quality parameters. |
| **Methods** |
| Pilot **participants** evaluate how useful the visualisations are for them with scale 1 to 10, 1 being fully useless and 10 being extremely useful. Average answer is calculated. |
| **Accomplishment Scale** |
| 1) Bad (when 1-2)  2) Poor (when 3-4 )  3) Medium (when 5-6)  4) Good (when 7-8)  5) Excellent (when 9-10) |
| **Results** |
| **Participant** evaluation for the visualisation framework was on average 7.5 🡪 **4 (good)** |

### Management platform

|  |
| --- |
| **Description** |
| Management of system’s entities – operators, sensors, gateways, organizing logs, warnings, remote access possibilities, entering of data, user-friendliness, etc. |
| **Methods** |
| **Researcher’s** rough estimation of the platform’s   1. costs / year / user vs target costs (660-1600€/ year/user[[4]](#footnote-4)) (azure, power BI, network connection, sensors, air purifiers, installation, maintenance…), 2. the level of automatic warnings related to data values, data losses, lost data connections etc should be considered, 3. Remote access. |
| **Accomplishment Scale** |
| 1) Bad (average between 1-1.8)  2) Poor (average between 1.9-2.6)  3) Medium (average between 2.7-3.4)  4) Good (average between 3.5-4-2)  5) Excellent (average between 4.3-5) |
| **Results** |
| **Researcher** evaluation for the management platform was   * Costs (azure=~200€/year plus 150€/year data storage cost for wrist devices), power BI=120€/year, network connection 60€/year, sensors (wrist device: 200€, noise meter: 167€, mcf88:360€, THP sensors ~50€ each, 4G router ~150€) =~400€, air purifiers, installation, maintenance…): costs would reach to 1400 🡪 3 (? 1=costs > 2000€, 2= costs 1999-1500€, 3=costs 1499-1000€, 4=costs 999-500€, 5=costs <500€) * Automatic warnings (abnormal data values, data losses, lost data connections): There was automatic warnings of connection loss for one sensor type (noise meter) (2) (1=no automatic warnings, 5= automatic warnings for all parameters, their abnormal values, data loss, and connection loss) * Remote access: the status of the sensors could be checked remotely (4) (1=no remote access to anything, 5= remote access to all measurement devices) * Average **3 (good)** |

### Big data management / processing

|  |
| --- |
| **Description** |
| Azure components were used to implement data acquisition, storing and processing for the pilot implementation. How well the implementation was realised? Were there data losses, problems with data storing? Were there batch processing or real-time/stream-based processing? |
| **Methods** |
| **Researcher’s** rough estimation on how well the big data management tools perform, their suitability and robustness in scale 1 to 5: 5 being implementation works as planned’ and 1 being ‘constant problems’. |
| **Accomplishment Scale** |
| 1) Bad  2) Poor  3) Medium  4) Good  5) Excellent |
| **Results** |
| **Researcher** evaluation for the big data management was 5 (excellent): big data management tools were suitable and robust. However, getting consistent connection from sensor GW to the cloud was sometimes difficult depending on the client SW stack used. Local data caching on the GW was implemented as a remedy. After this, there were no data losses due to the tools used. |

### Use case implementation

|  |
| --- |
| **Description** |
| Evaluate the progress of all the components-subsystems to be developed and the integration between them  Main components to be evaluated are:   * air quality sensors * physiological sensor * self-reporting application * air purifiers * recommendations/automatic HVAC adjustment * concentration tests (pupils) * visualizations/feedback module * data collection backend |
| **Methods** |
| **Researcher** evaluates implementation of each module by scale 1 to 5. 5 is given if the module is implemented as planned and 1 if not. Evaluation of the sensor systems as a whole is given as an average of different module evaluations. |
| **Accomplishment Scale** |
| 1) Bad (average between 1-1.8)  2) Poor (average between 1.9-2.6)  3) Medium (average between 2.7-3.4)  4) Good (average between 3.5-4-2)  5) Excellent (average between 4.3-5) |
| **Results** |
| **Researcher** evaluation for the use case implementation was **4:**   1. air quality sensors: installed and integrated to the system (5) 2. physiological sensor: installed and integrated to the system (5) 3. self-reporting application: installed and integrated to the system (5) 4. air purifiers: installed and integrated to the system (5) 5. recommendations/automatic HVAC adjustment: not implemented (1) 6. concentration tests (pupils): concentration tests on paper but not in electronic form (2) 7. visualizations/feedback module: installed and integrated to the system (5) 8. data collection backend: done (5)  * average **4,1 (good)** |

### Operability

|  |
| --- |
| **Description** |
| Evaluate the operability of the pilot components considering how simple it is to operate and maintain and whether there was problems. Main components to be evaluated are:   * air quality sensors * physiological sensor * self-reporting application * air purifiers * concentration tests (pupils) * visualisations |
| **Methods** |
| **Researcher** evaluates the operability of each of the above-mentioned component with scale 1 to 5, 1 being bad, 5 being excellent. Average is calculated. |
| **Accomplishment Scale** |
| 1) Bad (average between 1-1.8)  2) Poor (average between 1.9-2.6)  3) Medium (average between 2.7-3.4)  4) Good (average between 3.5-4-2)  5) Excellent (average between 4.3-5) |
| **Results** |
| **Researcher** evaluation for the operability was **3**   1. air quality sensors: Noise sensor required visit every now and then at the pilot site but other sensors were working quite well (4) 2. physiological sensor: the physiological sensor was cumbersome to operate and didn’t work for a while in the beginning. Also, data transfer from the sensor required fairly detailed procedures from participants using the sensor + phone, which led to some data loss (2) 3. self-reporting application: self-reporting application worked without problems (5) 4. air purifiers: air purifiers worked practically by themselves (5) 5. concentration tests (pupils): tests were pen and paper and they required management from the teacher and lots of time when digitising them. Test was otherwise really suitable. (2) 6. visualisations: Sharing visualisations required some operations, it was easy to open basic view but many features were hidden (3)  * average **2,7 (good)** |

### Data sources utilization

|  |
| --- |
| **Description** |
| Are data sources used in optimal way? Do we have overloaded data sources? Do we have space for system optimization in this respect? |
| **Methods** |
| **Researchers** assess how much of the collected data (%) is utilized in the current feedback system. |
| **Accomplishment Scale** |
| % of the data source utilization.  1) Bad (0-20%)  2) Poor (21-40%)  3) Medium (41-60%)  4) Good (61-80%)  5) Excellent (81-100%) |
| **Results** |
| **Researcher** evaluation for the data sources utilisation was **3 (good)**  The visualisation is made from the indoor air quality data and self-reports, but not the physiological data, or concentration data 🡪 50% |

### Storage, recording of the data

|  |
| --- |
| **Description** |
| What percentage of the stored data is useful and reliable? How much of the collected data can be used in further analysis. |
| **Methods** |
| **Researcher’s** rough estimation of the good quality data in scale from 1 to 5, 5 being ‘all data was useful and reliable’ and 1 ‘the data was useless/ impossible to be used’. |
| **Accomplishment Scale** |
| 1) Bad (average between 1-1.8)  2) Poor (average between 1.9-2.6)  3) Medium (average between 2.7-3.4)  4) Good (average between 3.5-4-2)  5) Excellent (average between 4.3-5) |
| **Results** |
| **Researcher** evaluation for the storage and recording of data was **3 or 4**.   1. air quality sensor data (noise, humidity, pressure, temperature…): (3/4) 2. physiological sensor: data was too noisy and thus useless for stress detection via inter-beat interval estimation. (1) 3. self-reporting application: self-reporting data was useful and as reliable as self-reports can be (5) 4. concentration data (pupils): There were only few discrepancies between the reported number of papers and pupils, some reported results were corrected after rechecking the results (4)  * average **3.3/3.5 (medium)** |

### Stakeholders’ feedback

|  |
| --- |
| **Description** |
| How satisfied the participants were with the pilot as a whole |
| **Methods** |
| **Pilot participants’** satisfaction for the entire system as a whole is collected though NSP (net promoters score) on a scale 0-10 at the end of pilot. Score gives output based on all answers between -100 to 100. |
| **Accomplishment Scale** |
| 1) Bad (when NPS -100- 61)  2) Poor (when NPS -60 – -21 )  3) Medium (when NPS -20- +20)  4) Good (when NPS 21-60)  5) Excellent (when NPS 61-100) |
| **Results** |
| **Participant** evaluation of the entire system satisfaction was 0 on NPS scale 🡪 **3 (medium)** |

### IoT Interoperability

|  |
| --- |
| **Description** |
| Ease of using other types of sensors and communication protocols, or the ease of using several IoT platforms at the same time. |
| **Methods** |
| **Researcher’s** rough estimation (%) on how much of the current implementation allows integration of other types of sensors/communication protocols/IoT platforms. |
| **Accomplishment Scale** |
| 1) Bad (0-20%)  2) Poor (21-40%)  3) Medium (41-60%)  4) Good (61-80%)  5) Excellent (81-100%) |
| **Results** |
| **Researcher** evaluation for the IoT interoperability was **4 (good)**. Currently, sensors may communicate via MQTT, AMQP or HTTP using Azure client SDKs or by tagging protocol data. Client SDKs exist for various clients, protocol tagging available for any device. Some bridge components from other platforms to Azure exist already (e.g., ARM Pelion / LwM2M), and others can be built using, e.g., virtual machines. Getting data out to other systems is also possible in both stream and batch mode, but less streamlined than with Azure data processing chains. |

### Reliability of simulations / predictions

|  |
| --- |
| **Description** |
| Using objective IEQ sensor data to classify personal perceptions of teachers regarding “negative” or “positive” productivity, stress and IEQ. |
| **Methods** |
| Employing machine learning method (Support Vector Machine, SVM) with person-specific training (*i.e.*, model of each individual was trained using only his/her data) to classify teachers’ perceptions of work productivity, stress and IEQ. |
| **Accomplishment Scale** |
| 1) Bad (0-20%)  2) Poor (21-40%)  3) Medium (41-60%)  4) Good (61-80%)  5) Excellent (81-100%) |
| **Results** |
| Experimental results with real life data, collected in four classrooms of Finnish elementary school during 18 weeks, show that IEQ sensor data allows to classify with high accuracy perceptions of teachers regarding their work productivity (91%), stress (81%) and IEQ (92%).  **Researcher** evaluation for employing machine learning classifier (SVM) to classify personal perceptions of teachers regarding “negative” or “positive” productivity, stress and IEQ was **5 (excellent)**, because the average classification accuracy was over 81%. |

### Portability

|  |
| --- |
| **Description** |
| Understanding portability, how easy it is to adapt and apply the pilot to other cities or other buildings. How coupled the software systems is developed? |
| **Methods** |
| **Researcher’s** rough estimation (%) on how much of the current implementation can be directly applied to other cities/buildings. |
| **Accomplishment Scale** |
| 1) Bad (0-20%)  2) Poor (21-40%)  3) Medium (41-60%)  4) Good (61-80%)  5) Excellent (81-100%) |
| **Results** |
| **Researcher** evaluation for the portability was **5 (excellent),** system can be implemented in another place very well. The sensors are not specific to any particular building type, and self-report questionnaires can be configured using Google Sheets. Data collection and processing is done in cloud (no server HW required). Switching the cloud back-end to another one is expected to require some porting work. However, this should be just fairly a straightforward adaption task by switching the cloud connector client interface to the new target one. Deployed format in date exchange is generic JSON which is application and software implementation independent. |

## Reporting and decision making

The following table presents the results of the parameters evaluation done by client, system end user or expert/specialist, as specified in the previous section, and appropriate decisions/proposed actions drawn from the findings (Step 4). Table 3 will be updated with the actual result in the next version.

Table 5. System evaluation results

|  |  |  |  |
| --- | --- | --- | --- |
| Parameter | Parameter evaluation | Result  (1=Bad, 2=Poor, 3=Medium, 4=Good, 5= Excellent) | Action |
| Meeting objectives | By researcher | 4 |  |
| Systems’ performance | Both by researcher and pilot participant | 4 (by researcher)  2 (by participant) | System should be improved to bring more concrete value for the end user. Would require intelligent HVAC systems and perhaps even giving information of what good air quality consists of. |
| Privacy and security | By researcher | 4 |  |
| Feedback system performance | By pilot participant | 3 |  |
| System satisfaction | By pilot participant | 3 |  |
| Sensor system performance | By researcher | 4 |  |
| Visualization framework | By pilot participant | 4 |  |
| Management platform | By researcher | 3 |  |
| Big data management | By researcher | 5 |  |
| Use case implementation | By researcher | 4 |  |
| Operability | By researcher | 3 |  |
| Data sources utilization | By researcher | 3 | Physiological sensor data is not useful; either a more accurate sensor should be used or the data discarded altogether. |
| Storage, recording of the data | By researcher | 3 or 4 |  |
| Stakeholders’ feedback | By pilot participant | 3 |  |
| IoT Interoperability | By researcher | 4 |  |
| Reliability of simulations / predictions | By researcher | 5 for predicting IEQ, stress and productivity | We had only 4 test subjects. Prediction accuracy (IEQ, stress and productivity) should be confirmed in populations that are more diverse and in different working environments. |
| Portability | By researcher | 5 |  |

## Restrictions of the evaluation processes

The parameters are evaluated by researcher, pilot participants or both. Most of the parameters are evaluated by a researcher only (11/16). Also, many research methods include researchers’ decision and no objective measures are necessarily used but the evaluations is researchers’ estimation.

The pilot participant number was four and therefore it sets also limitations for the generalisability of the end user evaluations. The good thing is that the pilot participants have interacted with the system 3 to 4 months, so they have had time to form opinion on it.

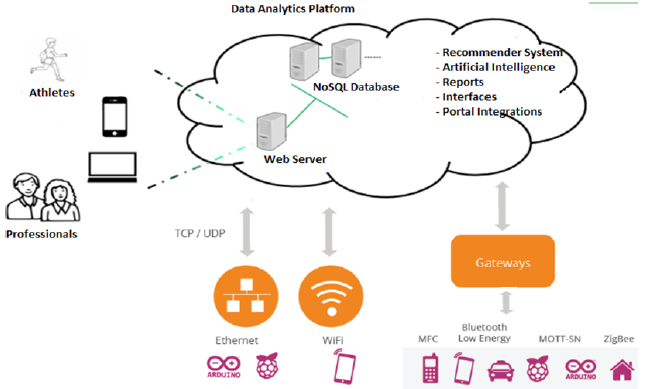
# Tracking of athletes with wearable sensors

## Parameters and methods

The development of wearable devices has given rise to new fields such as data analytics, reporting, developing a recommender system, while removing the problem of data collection. Wearable sensors, sensors that enable measurement of environmental factors and data from mobile applications can be analysed to give suggestions as health and life assistant to the athletes. In this sense, the pilot, which aims to close an important gap, will present innovations in the sense of solving the problems experienced with its goals. The main objective of this pilot is to define an IoT infrastructure that integrates different sensor technologies. This infrastructure will include the several capabilities such that:

* Environmental factors, body basic state parameters will be collected and analysed with machine learning and text mining methods.
* According to the results, the services and the applications that are able to provide guidance and direction can be developed.

ESTABLISH will be implemented as a PaaS (Platform as a Service), see Figure 7 below.



*Figure 7. Platform as a Service.*

The platform will have the ability to receive and store data from any device that has the capability of communicating and will have the ability to make suggestions based on the prescription (Rule sets) that will be created by users based on their profile. Platform will include content based recommender system for recommendation. Thus, it will be a platform that can be used by amateur or professional individual and team athletes who want to record, report and manage their sports activities, as well as individuals who use lifestyle facilitating devices.

In the platform there is also another data type which is environmental measurement data such as air quality and temperature of ambient. This data type will be got from a state-owned organization which is authority on environmental issues. This organization publishes environmental measurements such as temperature and air quality on a neighbourhood basis. If the athlete is outdoors, the published measurement data will be used. If the athlete is in the indoor environment, the measurements will be taken from the sensors that measure the indoor air quality and temperature.

For example, an athlete may keep his training program on his profile on the platform as a recipe. Wearable sensors can be used to transmit values that measure body behaviour in his/her sports activities. The athlete will be able to see these values in meaningful reports and will be able to take recommendations on behalf of coaching according to the pre-entered rules and the results obtained from the analysis of the transferred data. The main components of the project are:

1. Extracting meaningful information by analysing collected data from mobile applications and IOT devices.

2. Integration of semantic data using Big Data platform,

3. Content-aware adaptation and automation of the IOT infrastructure,

4. Development of a suggestion system based on the results of data analysis.

5. Development of mobile and web applications software for tracking data, accessing analysis results and tracking recommendations.

To realize the project, two parts should be deployed Local Zone and Cloud Zone. Local zone includes some data analytics and data aggregation and manipulation steps, data acquisition from sensors. Data can be collected in local storage mechanisms or data can be stream immediately to Cloud Zone via IoT Gateway. There are Authentication and Authorization mechanisms on the data transfer and monitoring. A REST API will be implemented to data transfer between zones. The API will include Login and several pushData services.

Cloud zone will include some Big Data Analytics Components such as Cassandra, Machine Learning Libraries & Tools, Spark, in-memory databases, Kafka etc. The core data analytics will realize in this site of project. Of course there will be many open source technologies for the monitoring of the result and information delivery to related people such as Apache Web Server, Dashboards, Reporting tools etc.

## Reporting and decision making

In the Establish Cloud zone, there are some analytics on the data that is collected from environment, athletes and their trainings. In Turkish use case the data type is sensor data, so predictive analysis, classification and some clustering algorithms are applied. The aim of the analytics can be listed as follow.

* Making predictions about athletes‘ performance
* Determine wrong / correct training techniques
* Tracking athletes
* Determine impact of environmental factors on athletes ‘performance

To realize Turkish use case, two parts should be deployed Local Zone and Cloud Zone. Local zone includes some data analytics and data aggregation and manipulation steps, data acquisition from sensors. Data can be collected in local storage mechanisms or data can be stream immediately to Cloud Zone via IoT Gateway. There are Authentication and Authorization mechanisms on the data transfer and monitoring. A REST API is implemented to data transfer between zones. The API includes Login and several pushData services.

Cloud zone includes Big Data Analytics Components such as Cassandra, Machine Learning Libraries & Tools, Spark, in-memory databases, Kafka etc. The core data analytics realize in this site of project. Of course there are many open source technologies for the monitoring of the result and information delivery to related people such as Apache Web Server, Dasboards, Reporting tools etc.

Data storage layer of Establish solution for Tracking of athletes with wearable sensors pilot includes Cassandra as a NoSQL database.

Decisions are made by dashboard. In the dashboard user can see the results of the sensor measurements:

* Sensor 1: Motion sensor
* Sensor 2: Blood pressure sensor
* Sensor 3: Brain activities sensor
* Sensor 4: Focus for attention sensor
* Sensor 5: Sleep sensor
* Sensor 6: Body Temperature sensor

There are three actuators that are tracked:

* Actuator 1: Electrocardiogram sensor
* Actuator 2: Cardiac fitness
* Actuator 3: Stress

Decisions are made by sports experts by comparing the results of the sportsman performance in its history. Similarly, athletes should be encouraged to make a decision by evaluating options and being allowed to make mistakes and purposely inspect these mistakes so that it does not occur in the future. For this atmosphere to be reached by athlete, the coach must provide the change and environment. And the biggest impact on whether athletes end up taking responsibility is how the athlete is coached.

Coaches tried the dashboard to see their decisions, where they use questioning techniques to draw out their athlete’s thinking. To use questioning technique means not relying on a “traditional” style of coaching which is usually based around the directive style of coaching. If coaches mainly rely on this style and give their athletes the answers to most of the problems they face, then the athlete never learns to address issues themselves. They need to learn and practice decision making. If an athlete does not get the opportunity to learn and practice decision making in training, they do not get it right in a competition. Overall, coaches who use an athlete-centred approach have a better chance of developing athletes who have self-awareness and who have the abilities to make great choices in both training and then competition. So our system has been reviewed by some non-professional sport center coaches.

## Restrictions of the evaluation processes

Besides Turkgen and Semantik Ar-Ge’s development processes and testing of the use-cases, we have implemented the use-case to a non-professional sport center. Two coaches tested the athlete tracking system on non-professional sportspeople who subscribe to this center. The first evaluation results were positive and the center liked to test the measurement of the sportspeople by smartwatches. This use-case was also welcomed by the sportspeople to see their past performance and evaluate their performance by professional coaches who can comment on their performance.

# Conclusion

This document has provided validation and evaluation of the 3rd use case of the ESTABLISH project - “Promoting independence of specific vulnerable groups”, in which have been developed three pilots: “Rehabilitation decision support” – implemented in Romania, “Indoor air quality improvement at school” – implemented in Finland, and “Tracking of athletes with wearable sensors” – implemented in Turkey. These use cases have been implemented within Czech and Korean pilots, respectively.

Validation results were very positive and help to make final assessment of the system readiness to enter the market. From the business model point of view the evaluation results support development of the proposed business models discussed at stakeholders’ meetings and workshops.

# References

1. ESTABLISH full project proposal.
2. D3 1 Architectural and design specification
3. D3.3 Test Plan
4. D2.1 State-of-the-art, use case definitions and user requirements
5. D2.2 Technical requirements and high level specifications
6. D4.1 Architecture of Data Management Platform
7. D4.3 Data Privacy and Security on the Platform
8. BetterEvaluation Rainbow Framework, <https://www.betterevaluation.org/plan/manage_evaluatio>n
9. Evaluation Methodology, by Marie Baehr, Elmhurst College, <https://www.webpages.uidaho.edu/ele/scholars/Practices/Evaluating_Projects/Resources/Evaluation%20Methodology.pdf>
10. <https://www.issco.unige.ch/en/research/projects/ewg95/node24.htm>l
11. FP7-318338, Requirement Analysis and Evaluation Framework, Akrivi Katifori (UoA), Rudolf Schlatte (UiO), <http://optique-project.eu/wp-content/uploads/2012/11/D1.2.pdf>, Deliverable D1.2
12. Abuhussein et al 2012, Evaluating security and privacy in cloud computing services: A Stakeholder's perspective, <https://ieeexplore.ieee.org/document/6470836/>
13. Net Promoter Score (NPS), <https://en.wikipedia.org/wiki/Net_Promoter>

1. https://iaqscience.lbl.gov/si/benefits-rates [↑](#footnote-ref-1)
2. <https://app.powerbi.com/view?r=eyJrIjoiMmZhYTRjMTYtMGFkZS00MjI5LTgxYjMtNzQ0N2Q4Y2Y4OTU4IiwidCI6IjA0NzRjNDUzLTkxNjgtNDViNy05ODU5LWYxY2JlNmU2NWUyNCIsImMiOjl9> [↑](#footnote-ref-2)
3. <http://agile.ro/intern/meshlium-establish/index.php> [↑](#footnote-ref-3)
4. https://iaqscience.lbl.gov/si/benefits-rates [↑](#footnote-ref-4)