



## D6.2 - Initial Evaluation Framework Specifications

Deliverable ID	<b>D6.2</b>
Deliverable Title	<b>Initial Evaluation Framework Specifications</b>
Work Package	<b>WP6</b>
Dissemination Level	<b>PUBLIC</b>
Version	<b>1.0</b>
Date	<b>2019-06-25</b>
Status	<b>Final</b>
Lead Editor	<b>Christian Fernández (BQ)</b>
Main Contributors	<b>Brunella Caroleo (LINKS Foundation) , Yannick Bachteler (FIT)</b>

**Published by the GOEASY Consortium**

Document History

Version	Date	Author(s)	Description
0.1	2019-03-21	FIT	First Draft with TOC
0.2	2019-05-16	FIT	Sections 4.1, 4.2, 4.3, 4.4 added
0.3_BC	2019-05-22	LINKS	Section 5 added
0.3_FIT	2019-05-22	FIT	Sections 2, 3 and 4 completed
0.4	2019-05-22	BQ	Versions merged
0.4_BC			
0.4_BC_FIT	2019-05-28	FIT	Pilot partner feedback included
1.0	2019-06-20	BQ	Final version of the document

Table of Contents

Document History .....2

Table of Contents .....2

1 Introduction..... 3

    1.1 Scope ..... 3

    1.2 Related documents..... 3

2 Evaluation Methodology .....4

3 Evaluation Objectives..... 5

4 Evaluation Framework.....6

    4.1 Timeline..... 6

    4.2 Review of KPIs – End of period 1 ..... 6

    4.3 Review of KPIs – End of period 2 .....10

    4.4 Formative evaluation tasks.....10

    4.5 Alpha-version evaluation .....10

5 Impact Assessment Approach.....15

6 Conclusions.....17

Acronyms ..... 18

List of figures.....18

List of tables.....18

References ..... 18

Appendix.....19

    A. Overall timeline of evaluation activities in GOEASY (M18 to M36).....19

## 1 Introduction

This task will define the overall methodological framework along qualitative and quantitative metrics for the evaluation towards requirements and project objectives. Since Data Management aspects are a key issue in the evaluation process, in the early stages, this task will focus on the documentation of procedures used by the project to handle data collected in test sites.

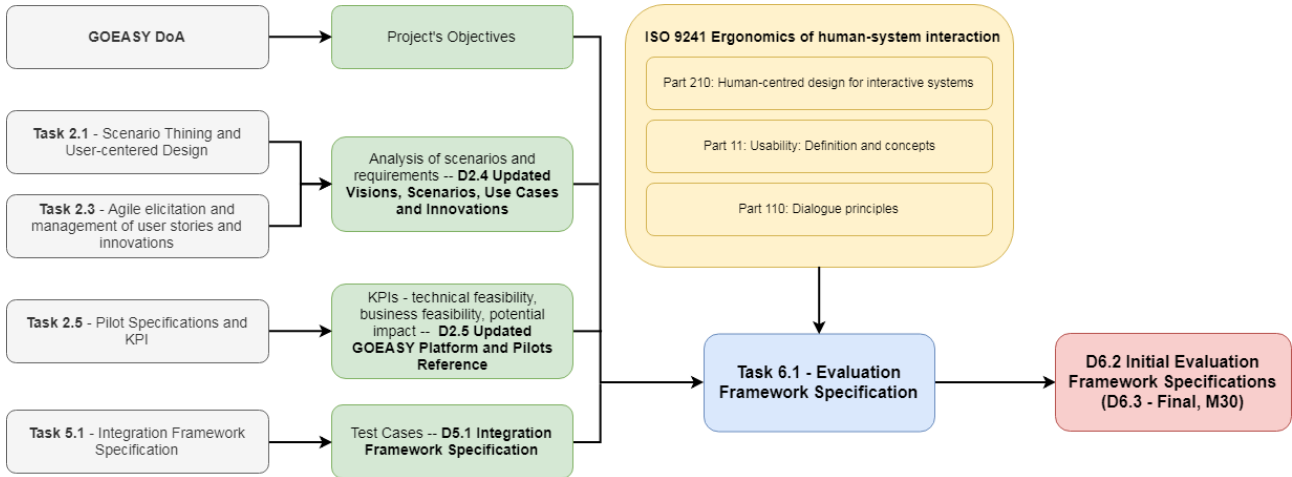


Figure 1 – Input sources for Task 6.1

### 1.1 Scope

The task will focus on the definition of the evaluation objectives, the methodology to be followed, the evaluation phases and their timing. Evaluation indicators for user acceptance will be defined based on ISO 9241 standards.

The task will set up instruments which are needed for the uniform collection of data during the evaluation (e.g. online questionnaires, data collection forms, direct interviews, quantitative evaluations, etc.). The evaluation framework will provide guidance for carrying out the evaluation activities and for making the decisions about redesign, error correction, start of implementation etc., on the basis of the evaluation results. It will act as baseline on how, when and by whom evaluation is going to take place.

### 1.2 Related documents

ID	Title	Reference	Version	Date
[RD.1]	Description of Action/ Grant Agreement			
[RD.2]	D2.1 Initial Visions, Scenarios and Use Cases		1.0	02/2018
[RD.3]	D2.4 Updated Visions, Scenarios, Use Cases and Innovations		1.0	03/2019
[RD.4]	D5.1 Integration Framework Specification		1.0	07/2018
[RD.5]	D2.5 Updated GOEASY Platform and Pilots Reference		1.0	05/2019

## 2 Evaluation Methodology

Evaluation is part of the implementation of the user-centred design process, described in [RD.2]. The main aim is to assure that the GOEASY services and the pilot applications developed follow common quality standards and meet the needs and requirements of stakeholders, especially end users.

The evaluation framework described in this deliverable focuses on the alpha-version evaluation, which should be conducted in the months 22 to 24. The *D6.3 Final Evaluation Framework Specification* will focus on the beta-version evaluation, conducted in the months 31 to 35. However, the evaluation phases follow the same methodology.

While summative evaluations are planned in the alpha- and beta-version evaluation phase, formative evaluations may take place at any time (see Section 4.4). Formal and detailed documentation will only be provided for the summative evaluations and will be written down in *D6.4 Alpha Version Deployment and Evaluation Report* and *D6.5 Beta Version Deployment and Evaluation Report Plan* accordingly. Reports or results from formative evaluations carried out in the meantime may be included in these deliverables.

As part of the UCD approach, the planned summative evaluations will make use of qualitative user evaluations. User tests will allow to judge the usability and user experience, including user acceptance of the solutions. In order to evaluate it, the analysis of results will be based on ISO 9241 indicators. While ISO 9241-210:2010 describes the process itself, ISO 9241-110:2006 provides concrete indicators required for a usable product. Following ISO 9241-210 (International Organization for Standardization, 2010), evaluations enable to:

- Collect information about user needs,
- Gather feedback about strengths and weaknesses of the design,
- Judge, if user requirements are fulfilled,
- Gather base data or compare different designs.

Regardless of the concrete objective (listed above), an evaluation always consists of certain steps (International Organization for Standardization, 2010):

- (1) Prepare evaluation, also considering the overall timeline of the project.
- (2) Conduct evaluation to gather comprehensive results regarding the focused objective.
- (3) Analyse results, define key issues and possible solutions.
- (4) Share possible solutions in an appropriate way with the development team.

Those phases are also shown in Figure 2, which describes the timeline for the alpha-version evaluation in Section 4.1. In GOEASY, step (4) is split into two, since the results will not only be shared with the development team, but with any project stakeholder and, due to its dissemination level (public), any person interested in GOEASY:

- (4) a) Feedback results to development
- b) Finalize report D6.4

There are several methods that can be applied for an evaluation following the approach of the user-centred design. In addition to a concrete goal, an evaluation is also always dedicated to a specific topic. In the project, this allocation is done on the basis of the use cases defined in [RD.3]. However, this does not mean that an evaluation must be planned for each defined use case. Some use cases must be combined to enable the user to accomplish a task (e.g. see UC-AW-03: to define a route, further use cases are triggered in the back, without the user noticing).

Regarding the evaluation for user acceptance, analysis will mainly rely on ISO 9241-110:2006, which defines “dialogue principles related to the ergonomic design of the dialogue between user and interactive system, and does not consider any other aspect of design such as marketing, aesthetics or corporate design” (ISO, 2018). Moreover, it is worth noting that GOEASY applications will affect the pilot context not only in the short term, but also in the long term. For this reason, in this deliverable (Section 5) the evaluators also investigate the added value of GOEASY applications in terms of long-term impacts on society, economy and environment, in tune with the “triple sustainability” tenets.

### 3 Evaluation Objectives

The evaluations objectives in GOEASY result from different points of view. Based on the technical and strategic objectives described in the Description of Action (DoA), Key Performance Indicators were defined which are used to review project's progress. Further KPIs are derived by the project objectives focusing on the long-term impacts of GOEASY.

Further evaluation objectives are given by the pilot applications and their users. It is important, that the user achieves his goal in an efficient, effective and satisfying way. This can only be reliably evaluated with potential users (defined in [RD.3]) involved. Gathering their feedback about strengths and weaknesses of the design allows to increase the experienced usability and user experience as much as possible.

## 4 Evaluation Framework

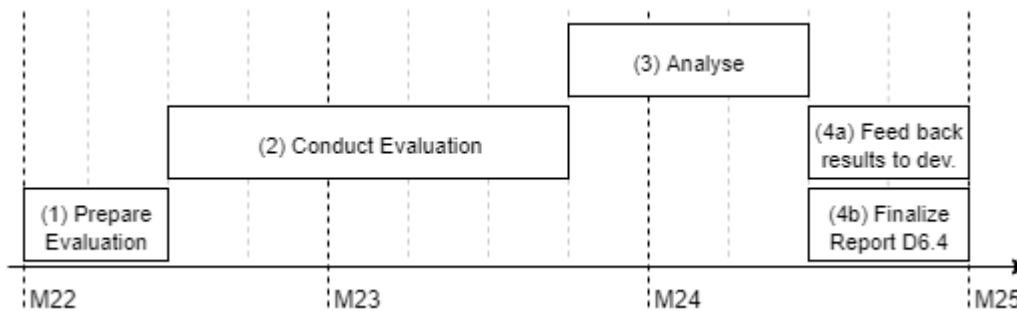
### 4.1 Timeline

During the project's lifetime multiple evaluation activities will take place. Beside the summative evaluation of alpha-/ beta-version and the formative evaluation in between, also the review of the KPIs and activities from Tasks *T5.5 Integration of Platform Components and Applications* and *T5.6 Scalability and e-Security stress-tests* have to be considered in planning:

- Review of the KPIs for period 1 (M19) and period 2 (M36)
- Alpha-version evaluation (summative evaluation; M22-M24)
- Beta-version evaluation (summative evaluation; M31-M35)
- Integration and stress-test based on T5.5 and T5.6 (reported in M19 & M34)
- Formative evaluation, including continuous review of requirements (at any time)

A graphical representation of the overall timeline of evaluation activities, including relevant milestones and deliverables, is attached in Appendix A.

Based on the milestones, the project timeline considers three months for the alpha-version evaluation, starting in month 22, if *MS10 Alpha-version Deployed and Ready for Evaluation* is achieved in M21. And ending in M24 with *MS11 Alpha-version Evaluation Phase Completed*, which includes the submission of *D6.4 Alpha Version Deployment and Evaluation Report*. This results in the following high level schedule:



**Figure 2 – Timeline for alpha-version evaluation phase.**

The beta-version evaluation is planned to last 5 months. The high level timeline will be shown in the final version of this deliverable (*D6.3*) in order to consider lessons learned and experiences collected in the alpha-version evaluation phase.

### 4.2 Review of KPIs - End of period 1

The success criteria defined in the DoA are listed as measurable KPIs in [RD.5]. The consortium will assess the progress of the KPIs as described in Table 1. The progress achieved at the end of period 1 will be presented in the mid-term review (June 2019) and documented in *D1.2 Periodic Activity, Management and Financial Reports* (M19).

**Table 1 – Evaluation Plan for KPIs at the end of P1.**

Objective ID	How to evaluate	Measurement	Responsible
TO1	<p>Integration test of running components.</p> <p>Precondition:</p> <ul style="list-style-type: none"> <li>- at least two components are running</li> <li>- test data is available</li> </ul> <p>Following components are expected to be deployed for the alpha-version and will be tested in an integration test:</p> <ul style="list-style-type: none"> <li>- Test 1: ComponentA, ComponentB, ComponentC</li> <li>- Test 2: ComponentX, ComponentB, ComponentZ</li> <li>- ...</li> </ul>	<p>Five-point scale, defined as following:</p> <p>1 = integration not started [either C1, C2 or test data is not ready]</p> <p>2 = integration started [integration with at least one component → C1 can send a request to C2]</p> <p>3 = partially integrated [C2 accepted request]</p> <p>4 = integrated [C2 responds to C1]</p> <p>5 = fully integrated [response matches specification of test data; integration test passed]</p>	CNET and GAPES (each for one test-level scale LBS)
TO2	Are multi-constellation measurements included?	Binary: Yes / No	BQ
	Are Galileo navigation data available?	Binary: Yes / No	LINKS
TO3	Scalability testing: How	<p>Five-point scale (aggregated time transformed to scale), defined as following:</p> <p>1 = not scalable [no component is scalable]</p> <p>2 = partially scalable [at least one of the required components is not scalable]</p>	GAPES

	<p>Step 1: Identification of components that could be bottlenecks</p> <p>Step 2: Conservative load tests for the identified components</p>	<p>3 = scalable [each required component is scalable enough but do not satisfy requirements, e.g. response time (GOES-89)]</p> <p>4 = highly scalable [every required component is scalable enough and satisfy most of the requirements]</p> <p>5 = fully scalable [every requirement regarding scalability is satisfied]</p>		
TO3	Is there a Smart City Platform federated with GOEASY services?	Binary: Yes / No	CNET	
	Is there a Collective Awareness Platform federated with GOEASY services?	Binary: Yes / No		
	Is there an Internet of Things Platform federated with GOEASY services?	Binary: Yes / No		
TO4	Comparison of actual vs. planned use cases (defined in [RD.3]) deployed.	Five-point scale, defined as following:		CNET and GAPES (each for their alpha-version application)
		1 = development not started		
		2 = development started	[at least a paper prototype is developed]	
		3 = use case implementation started	[at least one use case is implemented]	
		4 = all use cases implemented	[not all components are integrated with GOEASY yet]	
5 = integrated with GOEASY	[every component is integrated and integrated test passed]			
TO5	How many open enablers are adopted or integrated in the GOEASY platform (cloud/ data-oriented/ smart city enabler)?	Actual number of enablers.	CNET, GAPES	
SO1	How many users have been engaged in user-centred focus groups and evaluations?	Actual number of engaged users in UCD activities.	FIT	
	How many citizen engagement mechanisms were developed by GOEASY?	Actual number of developed engagement mechanisms (not yet evaluated).	CNET, GAPES	



SO2	<p>Are the following components implemented as designed?</p> <ul style="list-style-type: none"> <li>- Privacy-Aware DBMS (Data Access Mgr., Data Anonymizer, Aggregator, Public and Encrypted Data Storage)</li> <li>- &amp; e-Security Infrastructure (Position Alteration Detection Library, Trusted Collection and Exchange of Position Information)</li> </ul>	<p>Binary: Yes / No</p>	FIT, LINKS
SO3	<p>How many related business stakeholders are engaged in user-centred focus groups and business evaluations?</p>	<p>Actual numbers of engaged business stakeholders.</p>	CNET, COT, GAPES
SO4	<p>Preparatory activities including user recruitment for alpha-version tests</p>	<p>Binary: Yes / No</p>	CNET, COT, GAPES
	<p>How many users are engaged in GOEASY pilots?</p>	<p>Actual number of engaged users in pilots.</p>	CNET, COT, GAPES
	<p>Are security aspects evaluated in real-size pilots? Requires security analysis.</p>	<p>Binary: Yes / No</p>	FIT
	<p>Are scalability aspects evaluated in real-size pilots? Requires scalability analysis and stress-tests.</p>	<p>Binary: Yes / No</p>	FIT
	<p>Are projected costs and revenues evaluated in real-size pilots?</p>	<p>Binary: Yes / No</p>	BQ, CNET, GAPES

### 4.3 Review of KPIs - End of period 2

The evaluation of the criteria at the end of period 2 will follow the same structure as described in Section 4.2. Details will be provided in *D6.3* in order to make sure that any changes will be considered that may occur in the meantime. Currently, it is planned to evaluate the achievements in the beta-evaluation phase and report the results in *D1.4 Final Activity, Management and Financial Reports* that is due in M36.

### 4.4 Formative evaluation tasks

Formative evaluations can take place at any time in the project and are strongly linked with the agile development approach. These kind of evaluations will be conducted as expert evaluations that means that only project members or colleagues with the appropriate knowledge will participate. As an example: to evaluate a new design for the AsthmaWatch, not a potential end-user but a usability engineer from FIT will do the evaluation in a short session.

As described in [RD.3] the user stories are managed in JIRA and are updated accordingly to the current state of implementation. The default way is: Open → In Progress → Resolved → Closed. A developer changes the state to “Resolved” if the user story is implemented and the developer is satisfied. For quality assurance a user story cannot be closed by a developer. Only the requirements engineer is allowed to close a user story. The requirements engineer has to verify that there is no more work to be done and has to verify the completion in order to close the user story.

To check whether a user story has been completely and correctly implemented, each user story is tested following a test plan that defines a test objective, test case and its procedure and acceptance criteria for each user story. This plan was initially published in [RD.4] and is continuously updated if changes have to be applied to a requirement (next update is expected in *D5.5*). The moment of execution of the test is usually linked to the integration of the appropriate component. If the completion has been verified by the test, the user story will be marked as Closed by the requirements engineer. Tests will be executed by the requirements engineer and the component owner together.

With integration testing, the component owner is also requested to review non-functional requirements (NFRs) that are linked to the component. NFRs are documented as user stories in Jira and listed in [RD.3] and define system attributes for availability, legal, performance, platform support, reliability scalability, security and usability & humanity. If it is proven, that a linked NFR is fulfilled for that component, it gets documented in Jira (commented<sup>1</sup>) and reported along with the integration and stress-test results in *D5.5 Updated Integration and Stress-test report*.

Referencing the Description of Action, the evaluation framework specification tells “how, when and by whom evaluation is going to take place”. Following applies to formative evaluation:

- How: Following the test plan (published in [RD.4]).
- When: If the user story and depending stories are resolved by the developer.
- By Whom: Requirements engineer and component owner jointly.

A formative evaluation with potential users was already applied in the GOEASY project. In February 2019, five sessions were conducted in Florence to evaluate an interactive (non-coded) mock-up from ApesMobility. This mock-up was developed in multiple iteration from GAPES supported by FIT as usability experts. The involvement of potential end users enabled GAPES to gather useful insights before the implementation (coding) started. If the design would have been evaluated later (e.g. in the alpha-version evaluation phase), costs to adjust the application based on the insights would have been higher. This is one of the reasons why the agile development approach combined with the user-centred design is applied in this project.

### 4.5 Alpha-version evaluation

During the alpha-version evaluation, the pilot applications and especially the integration of platform components are evaluated. Furthermore, the progress of the implementation of non-functional requirements is checked.

---

<sup>1</sup> NFRs are usually required for more than one component. We decided to not duplicate a NFR for each component to keep the clarity. Thus, we use the comment functionality until it is tested and fulfilled for any related component. When that comes, the user story will be resolved and closed.

## (1) Prepare Evaluation

Following tasks have to be completed for each of the evaluation described in subsequent sections:

- Recruiting participants and schedule date and time for their session
- Review of deployed components
- Definition of task list
- (optional) Prepare mock-up in case component(s) are not deployed and running as expected
- GOEASY internal test run to check smooth execution of the planned evaluation sessions

## (2) Conduct Evaluation

Descriptions of how and by whom evaluations are conducted are given in the subsequent sections, grouped based on use cases in ApesMobility (see Section 4.5.1), AsthmaWatch (see Section 4.5.2) and GOEASY platform (see Section 4.5.3).

## (3) Analyse

An important change with regard to the data storage policies and regulations happened in the first half of the scope of the GOEASY project with the approval of the GDPR, in May 2018.

The General Data Protection Regulation (GDPR) is a legal framework that sets guidelines for the collection and processing of personal information from individuals who live in the European Union (EU). Since the Regulation applies regardless of where websites/software solutions are based, it must be complied by all sites that attract European visitors, even if they don't specifically market goods or services to EU residents.

The GOEASY project and platform is fully GDPR compliant along the whole chain of data exchange and is going to take care of the user's privacy to keep some sensitive data like location and daily habits safe and anonymous.

The information collected (subject to GDPR regulations) in all evaluation sessions will be added to Confluence for further analysis. During the analysis, key issues and concrete recommendations will be identified. Results of the analysis will also be reflected in user stories that could be newly created, updated or closed as out of scope – any change of a user story will be reasoned in the comment area of the appropriate user story.

Analysis will be led by FIT, supported by the concerned pilot partner and/or component owner.

## (4) a) Feedback results to development & b) Finalize report D6.4

As described in (3) Analyse, results will be reflected in user stories. This ensures that changes will be considered by the development.

The report *D6.4 Alpha Version Deployment and Evaluation Report* will contain a detailed description of the setup and execution of the evaluation, including the actual number of participants, key issues identified and applied changes to the solution.

### 4.5.1 Pilot Evaluation: ApesMobility

It is expected that following ApesMobility use cases will be deployed (fully or partly) for the alpha-version evaluation:

- UC-APES-07 – Join challenge
- UC-APES-05 – Certify location
- UC-APES-06 – Offer challenge
- UC-APES-08 – Manage location tracking (partly)
- UC-APES-09 – Manage permission for data usage (partly)

UC-APES-10 – Manage stored location data

These use cases and the corresponding components in the background, will be evaluated in user tests during the alpha-version evaluation. Except of UC-APES-06, all use cases will be tested in the same session, which is possible since the potential user is represented in one user role. UC-APES.06 has to be tested by a different user role and thus, in an own session. Further details are provided in Table 2 and Table 3.

**Table 2 – Evaluation of ApesMobility use cases with citizens.**

Evaluation with citizens (UC-APES-07, UC-APES-05, UC-APES-08, UC-APES-09, UC-APES-10)	
<b>Method(s)</b>	Qualitative evaluation by user testing with thinking aloud, focussing on usability and user experience.
<b>Description</b>	Participant will get a task list that guides her/him through the app. While fulfilling the tasks, the participant is asked to think out loudly any thought that comes to her/his mind.  One participant per session.
<b>Participants</b>	3-5 citizens, who are potential users (whether or not knowing greenApes platform). Test environment is independent of the city.  Participants are recruited by COT and GAPES.
<b>Required components</b>	ApesMobility, GOEASY Local Trust Manager, Position Alteration Detection, End-to-end position authentication, (greenApes – existing application)
<b>Conduction by</b>	GAPES, supported by FIT

**Table 3 – Evaluation of UC-APES-06: Offer Challenge.**

UC-APES-06 – Offer challenge	
<b>Method(s)</b>	Qualitative evaluation by user testing with thinking aloud, focussing on usability and user experience.
<b>Description</b>	Participant will get a task list that guide her/him through the app. While fulfilling the tasks, the participant is asked to think out loudly any thought that come to her/his mind.
<b>Participants</b>	3-5 potential users from municipality or organization.  Participants are recruited by COT and GAPES.
<b>Required components</b>	greenApes (existing application)
<b>Conduction by</b>	GAPES, supported by FIT

Based on the use case description provided in [RD.3], the scope of use cases for the alpha-version will be limited as following:

- UC-APES-08 – Manage location tracking:  
Start and stop tracking is expected to be deployed. Setting a new alert cannot be tested in the alpha-version.
- UC-APES-09 – Manage permission for data usage:  
This use case is updated to allow that the ApesMobility user can link to an greenApes account. This will be tested in the alpha-version.

In case that any required component is not deployed as expected, it will be replaced by a mock-up that allows the conduction of the evaluation as planned. In this way we can guarantee to gather feedback from potential users regarding the usability and user experience.

#### 4.5.2 Pilot Evaluation: AsthmaWatch

For the AsthmaWatch application, it is expected that following use cases will be deployed (fully or partly) for the alpha-version evaluation:

- UC-AW-01 – Mark conditions on map (partly)
- UC-AW-02 – View conditions on map
- UC-AW-12 – Manage my profile (partly)
- UC-AW-07 – Enter health info (partly)
- UC-AW-08 – Monitor health info

These use cases and the corresponding components in the background, will be evaluated in user tests during the alpha-version evaluation. UC-AW-01 and UC-AW-02, as well as UC-AW-12, UC-AW-07 and UC-AW-08 are related and will be tested in the same session. Further details are provided in Table 4 and Table 5.

**Table 4 – Evaluation of UC-AW-01: Mark conditions on map & UC-AW-02: View conditions on map.**

<b>UC-AW-01 – Mark conditions on map &amp; UC-AW-02 – View conditions on map</b>	
<b>Method(s)</b>	Qualitative evaluation by user testing with thinking aloud, focussing on usability and user experience.
<b>Description</b>	Before starting the evaluation session Users sign a consent form. Participant will get a task list that guides her/him through the app. While fulfilling the tasks, the participant is asked to think out loudly any thought that comes to her/his mind.  One participant per session.
<b>Participants</b>	3-5 potential users (preferred asthma patient, alternatively non-asthma patient). City is not relevant, the task list just have to take care of that the participant is going to check the map of Copenhagen or Stockholm, since sensor data is coming in from there.  Participants are recruited by CNET.
<b>Required components</b>	AsthmaWatch, Data Access Manager, Public Data Storage
<b>Conduction by</b>	CNET, supported by FIT

**Table 5 – Evaluation of UC-AW-12: Manage my profile, UC-AW-07/ UC-AW-08: Enter/ Monitor health info.**

<b>UC-AW-12 – Manage my profile &amp; UC-AW-07 – Enter health info &amp; UC-AW-08 – Monitor health info</b>	
<b>Method(s)</b>	Qualitative evaluation by user testing with thinking aloud, focussing on usability and user experience.
<b>Description</b>	Participant explores by her/his own the profile area of the application. While exploring the app, the participant is asked to think out loudly any thought that comes to her/his mind.

	One participant per session.
<b>Participants</b>	3-5 asthma patients Participants are recruited by CNET.
<b>Required components</b>	AsthmaWatch
<b>Conduction by</b>	CNET, supported by FIT

Based on the use case description provided in [RD.3], the scope of use cases for the alpha-version will be limited as following:

- UC-AW-01 – Mark conditions on map:  
Measurement points are marked on map. Air pollution zones cannot yet be displayed.
- UC-AW-12 – Manage my profile:  
Adding thresholds is possible.
- UC-AW-07 – Enter health info:  
At least manually entering health information will be possible.

In case that any required component is not deployed as expected, it will be replaced by a mock-up that allows the conduction of the evaluation as planned. In this way we can guarantee to gather feedback from potential users regarding the usability and user experience.

#### 4.5.3 Evaluation of GOEASY platform use cases

For the GOEASY platform, it is expected, that the use case UC-GEP-01 will be deployed and ready for evaluation in the alpha-version.

**Table 6 – Evaluation of UC-GEP-01: Store sensor data to GEP.**

UC-GEP-01 – Store sensor data to GEP	
<b>Method(s)</b>	No explicit user testing.
<b>Description</b>	From a user perspective, the use case is tested by UC-AW-02: View Conditions on Map (see Table 4). From a technical perspective, the developer will check the communication at the UDP server that receives the data from the sensor and check the database to see the data is stored.
<b>Participants</b>	No potential user required.
<b>Required components</b>	Mobile Sensing Gateway encrypted and public database, Data Access Mgr, Data Anonymizer
<b>Conduction by</b>	CNET, supported by FIT

## 5 Impact Assessment Approach

GOEASY assessment approach should clearly state the steps and the methodology to assess the results of GOEASY pilot applications. In this regard, special attention is paid to the development of a set of proper indicators measuring the results of GOEASY initiative.

The results of an initiative/action/project can be classified temporarily, according to their temporal effects in the 'results chain'. The results chain (Figure 3) is the linear representation of the theory of change and show how the initiative will trigger different levels of change from activities to impact.

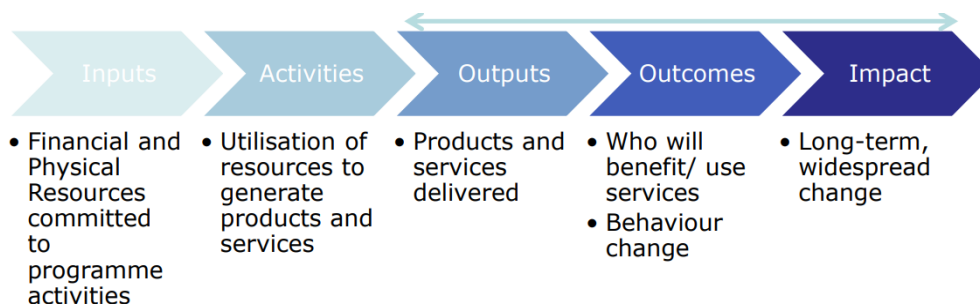


Figure 3: The results chain (European Commission, 2013).

As far as results are concerned, there can be indicators of outputs (short term), outcomes (medium term), and impacts (long term) (World Bank, Independent Evaluation Group, 2012):

- *output* indicators add more details in relation to the product ("output") of the activity, e.g. the number of smart meters distributed, the number of electric buses in the system;
- *outcome* indicators measure the intermediate results generated by project outputs. These indicators refer to the reason why it was decided to conduct certain interventions in the first place. They are the result of both the "quantity" ("how many") and quality ("how well") of the activities implemented. Often they are 'coverage indicators' measuring the extent to which the target population has been reached by the project;
- *impact* indicators measure the quality and quantity of long-term results generated by programme outputs, e.g. measurable change in quality of life, reduced energy use, reduced air pollutant emissions and (even a more distant impact) improved air quality.

Since indicators are measured to indicate progress toward goals, as an overall guiding principle suitable indicators should be selected by capturing the essence of GOEASY objectives, independently from the fact whether indicators are available / already in use for the targeted phenomenon or not. As a matter of fact, as detailed in the reference document of the European Commission (European Commission, 2009), it is essential to link objectives with future monitoring and evaluation activities: without clear objectives you cannot monitor and evaluate whether an innovative action is on track.

With this in mind, the output indicators have been derived from GOEASY technical objectives (TO), while the outcome indicators have been derived from GOEASY strategic objectives (SO), in accordance with the content shown in Table 1.

In order to complete the results chain, a possible set of impact indicators for the GOEASY city mobility case is provided in Table 7, covering different impact areas according to the 'triple sustainability' approach: society, economy, and environment.

Both output and outcome indicators could be affected in the GOEASY timeframe, while impact indicators are likely to be affected on the long term, which might take a few years.

Table 7: Proposal of impact indicators for GOEASY city mobility case.

Impact area	Impact indicator	What to measure?
Society	IA1. Behavioural change	Which portion of people have changed their mobility habits towards more sustainable ones thanks to GOEASY city mobility application?

		It could be measured by assessing the change of the modal split in the pilot city case.
Society	IA2. Urban planning	To what extent has GOEASY city mobility application contributed to, or inspired, changes in the current urban mobility policies? It could be measured by interviewing the local decision makers (e.g. mobility manager).
Economy	IA3. Economic activity	To what extent has the involvement of the commercial partners of GOEASY city mobility case contributed to the economic activity of the pilot city? It could be measured by interviewing the pilot managers.
Environment	IA4. Climate change	To what extent do CO <sub>2</sub> emissions from road traffic decrease in the pilot thanks to GOEASY city mobility application? It could be measured by the environmental monitoring stations in the pilot city.
Environment	IA5. Air pollution	To what extent do PM <sub>2.5</sub> emissions from road traffic decrease in the pilot thanks to GOEASY city mobility application? It could be measured by the environmental monitoring stations in the pilot city.
Environment	IA6. Noise pollution	To what extent does the noise level decrease in the pilot thanks to GOEASY city mobility application? It could be measured by the environmental monitoring stations in the pilot city.



## 6 Conclusions

Thanks to the implementation of the methodology mentioned during this writing, we can guarantee that the management of the information collected is truthful, honest and loyal. In fact, the information collected through online surveys and the assessment of end users will not be used for any purpose other than improving those services.

The evaluation framework described in this deliverable focuses on the alpha-version evaluation, which should be conducted in the months 22 to 24. The *D6.3 Final Evaluation Framework Specification* will focus on the beta-version evaluation, conducted in the months 31 to 35. However, the evaluation phases follow the same methodology.

## Acronyms

Acronym	Explanation
Dx.y	Deliverable number x.y of the GOEASY project
DoA	Description of Action
KPI	Key Performance Indicator
MSx	Milestone number x of the GOEASY project
Mx	Month number x of the GOEASY project
NFR	Non-Functional Requirement
TO	Technical Objective
SO	Strategic Objective
UCD	User-Centred Design
GDPR	General Data Protection Regulation

## List of figures

Figure 1 – Flow of Task 6.1.....	3
Figure 2 – Timeline for alpha-version evaluation phase.....	6
Figure 3 – Overall timeline of evaluation activities in GOEASY (M18-M36).....	19

## List of tables

Table 1 – Evaluation Plan for KPIs at the end of P1. ....	7
Table 2 – Evaluation of UC-APES-07: Join challenge & UC-APES-05: Certify location.....	12
Table 3 – Evaluation of UC-APES-06: Offer Challenge.....	12
Table 4 – Evaluation of UC-AW-01: Mark conditions on map & UC-AW-02: View conditions on map.....	13
Table 5 – Evaluation of UC-AW-12: Manage my profile. ....	13
Table 6 – Evaluation of UC-GEP-01: Store sensor data to GEP. ....	14

## References

- International Organization for Standardization. (3 2010). *ISO 9241-210:2010 Ergonomics of human-system interaction -- Part 210: Human-centred design for interactive systems*. Von ISO - International Organization for Standardization: <https://www.iso.org/standard/52075.html> abgerufen
- ISO. (2018). *International Organization for Standardization*. Von International Organization for Standardization: <https://www.iso.org/standard/38009.html> abgerufen

Appendix

A. Overall timeline of evaluation activities in GOEASY (M18 to M36)

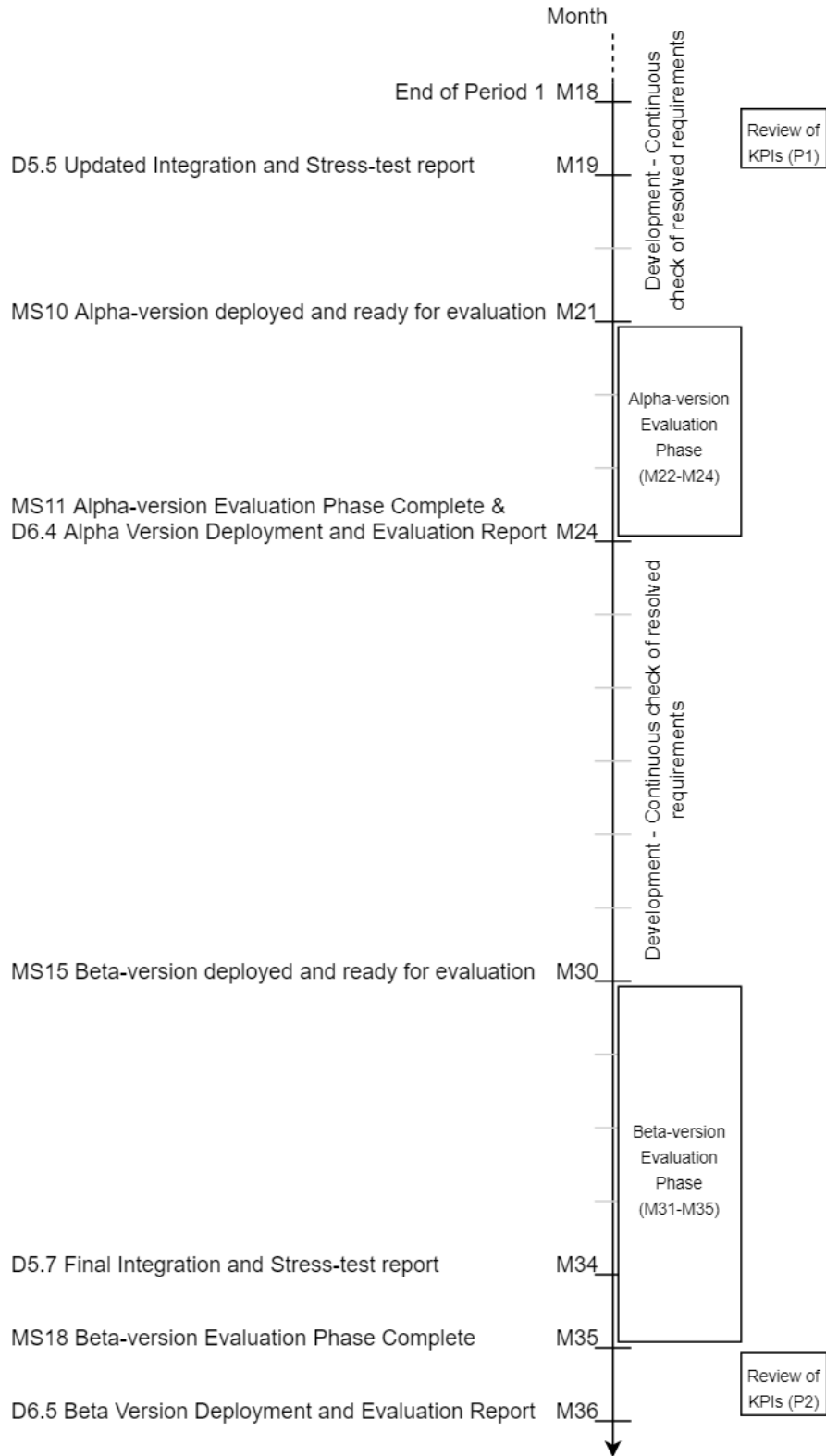


Figure 4 – Overall timeline of evaluation activities in GOEASY (M18-M36).