hackAIR

OPEN PLATFORM

D7.2: Evaluation and impact assessment framework

WP7 – Pilot operation and evaluation

Document Information

Grant Agreement Number	688363	Acronym	hackAIR	
Full Title	Collective awareness platform for outdoor air pollution			
Start Date	1st January 2016	Duration	36 months	
Project URL	www.hackAIR.eu			
Deliverable	D 7.2 – Evaluation and impact assessment framework			
Work Package	WP 7 – Pilot operation and evaluation			
Date of Delivery	Contractual	30th August 2017	Actual	6 October 2017
Nature	Report	Dissemination Level	Public	
Lead Beneficiary	VUB			
Responsible Author	Gavin McCrory, Carina Veeckman, Laurence Claeys			
Contributions from	NILU, BUND, DRAXIS, CERTH			

Document History

Version	Issue Date	Stage	Description	Contributor
0.1	30/06/2017	Draft	First draft of deliverable	Gavin McCrory
0.2	03/08/2017	Draft	Input from technical and pilot teams	Gavin McCrory, Hai-Ying Liu (NILU), Christodoulos Keratidis (DRAXIS)
0.3	04/09/2017	Draft	Second draft of deliverable	Carina Veeckman
0.4	24/09/2017	Draft	Final draft	Carina Veeckman, Gavin McCrory, Laurence Claeys
0.5	26/09/2017	Draft	Input technical indicators	Philipp Schneider (NILU), Eleftherios Spyromitros-Xioufis (CERTH), Anastasia Moumtzidou (CERTH)
0.6	29/09/2017	Review	Internal review	Panagiota Syropoulou (DRAXIS), Sonja Grossberndt (NILU)
1.0	01/10/2017	Submission	Reworked deliverable, ready for submission	Laurence Claeys, Carina Veeckman

Disclaimer

Copyright message

Table of Contents

Executive summary....5
1 Introduction....6
1.1 Purpose of the evaluation framework....6
1.2 hackAIR's evaluation framework....7
2 Evaluating usability, UX & acceptance and pilots of the hackAIR solution....10
2.1 List of indicators – Track I....10
2.1.1 Usability and effectiveness indicators....10
2.1.1.1 Set 1: Acceptance indicators....10
2.1.1.2 Set 2: Experience indicators - Perceived enjoyment....12
2.1.1.3 Set 3: Experience indicators - User satisfaction....13
2.1.1.4 Set 4: Usability indicators - User performance....14
2.1.1.5 Summary of indicators....16
2.1.2 Pilot performance indicators....17
2.1.3 Technical performance indicators....18
2.2 Quantitative and qualitative evaluation tools....19
2.2.1 Standardized survey forms....19
2.2.2 Think-a-loud protocol....19
2.2.3 Logging framework....20
2.3 Planning and milestones....21
3 Impact Assessment....23
3.1 Impact assessment methodologies: a brief literature review....23
3.2 Impact assessment methodology for hackAIR....29
3.2.1 List of indicators....30
3.3 Evaluation tools....31
3.3.1 Self-assessment surveys or discussion rounds....31
3.3.2 Impact assessment workshop....31
3.3.3 Interviews....32
3.4 Impact assessment Planning and milestones....32
4 Evaluation summary for hackAIR....33
5 Conclusion....35
6 References....36
7 Annexes....37
7.1 Annex 1: Usability, UX & acceptance template for evaluating the hackAIR solution....37
7.2 Annex 2: think-a-loud....41
7.3 Annex 3: Pilot performance indicators....42

7.4 Annex 4: Technical performance indicators 42
7.5 Annex 5: Impact indicators list.....43

Table of Tables

Table 1: Measures of the Evaluation framework (set I): measures from TAM and UTAUT perspective.....11
Table 2: Measures of the Evaluation framework (set II).....12
Table 3: Measures of the Evaluation framework (set III).....14
Table 4: User performance tasks on the hackAIR platform 15
Table 5: Measures of the Evaluation framework (set IV).....15
Table 6: hackAIR evaluation indicators, definitions and origins.....16
Table 7: Technical performance indicators of the hackAIR platform.....19
Table 8: Logging framework of the hackAIR solution. 21
Table 9:hackAIR planning for usability, UX and acceptance evaluation.....22
Table 10: Milestones period I- evaluation of user experience and acceptance of the hackAIR solution. 22
Table 11: Milestones period II – evaluation of usability, UX and acceptance of the hackAIR solution. 22
Table 12: Milestones period III: evaluation of usability, UX and acceptance of the hackAIR solution. 22
Table 13: Rating scale for usability, UX, and user acceptance scores. 23
Table 14: IA4Si methodological framework: socio-economic impact assessment (Passani et al., 2016, Chapter 3) 28
Table 15: Set I: impact indicators of the hackAIR project. 31
Table 16:hackAIR planning for impact assessment.....33
Table 17:hackAIR complete set of evaluation indicators 34

Table of Figures

Figure 1: Timing of WP7 pilot monitoring and project evaluation activities.....6
Figure 2: Evaluation framework of the hackAIR project 8
Figure 3:TAM Model. Source: Davis (1989) 10
Figure 4:UTAUT Model. Source: Adapted from Venkatesh, Morris, Davis, & Davis (2003) 11
Figure 5: Theorisation of dominance between PE and PEOU. Source: Adapted from Sun and Zhang (2006) 12
Figure 6: Research model of user satisfaction and technology acceptance. Source: Adapted from Wixom & Todd (2005) 13
Figure 7: hackAIR indicators for usability, UX and acceptance 16
Figure 8: Multi-dimensional impacts of projects 25
Figure 9: IAIA impact assessment methodology. 26
Figure 10: hackAIR impact assessment methodology.....29
Figure 11: Gantt chart showing hackAIR overall evaluation timeline. Green = Track 1, Red = Track2/3. 34

Executive summary

This document outlines the evaluation framework of the hackAIR project. The evaluation framework has the objective to evaluate **(1)** the usability, user experience (UX), effectiveness and acceptance of the hackAIR platform, and **(2)** the established impact of the platform at the social and environmental level in the communities.
The hackAIR evaluation framework will identify proper evaluation methodologies, tools and indicators to answer the following research questions:

- What are the main determinants for using the hackAIR open platform, from the angle of user experience and user acceptance?
- What is the social and environmental impact for the involved communities by using the hackAIR solution?
- How effective were the behavioural change interventions for participating in air quality monitoring and adopting a more pro-environmental behaviour, and which factors influenced the process of behaviour change?

Based on these research questions, the hackAIR evaluation framework will exist of three different tracks. Track one and two are presented in this deliverable, whereas track three was outlined in D6.2: Behavioural change techniques for hackAIR community.

Track 1 of the hackAIR's evaluation framework contains the evaluation activities for assessing the usability, user experience (UX), effectiveness and acceptance of hackAIR solution (platform, mobile application and sensors) with end-users and pilot coordinators. In addition, some pilot indicators are defined to evaluate the working and roll-out of the pilots in the different cities/regions. In this track, mostly quantitative methodologies are used, such as online surveys (within the hackAIR application), automatically generated log files and short surveys after physical meetings (e.g. workshops).
Track 2 comprises of a set of evaluation activities to investigate the social and environmental impact for the involved communities by using the hackAIR solution. VUB is responsible for the identification of the social and environmental impact indicators. In this track, mostly qualitative methodologies, such as focus groups, impact assessment workshops and interviews, will be chosen in a phased approach to collect the feedback from multiple target groups.
For Track 3, three different types of experiments are chosen to measure behaviour change, namely before and after the participation in hackAIR workshops, during the usage of the gamification features on the platform, and when consulting the tips of the day. The latter two experiments will especially make use of the logging details on the platform with frequent users, and with users who do not use these specific features.

The structure of the deliverable is as following:
In Chapter 1 the deliverable is introduced. Chapter 2 focuses on "Track I – Evaluating usability, UX & acceptance and pilots of the hackAIR " and will identify of list of indicators through literature review, propose evaluations tools, a logging framework, and a time plan.
Chapter 3 focuses on "Track II Impact Assessment" and will present a brief literature review about former established impact assessment methodologies, such as IA4SI, list social and environmental impact indicators, evaluation tools and a time plan for the hackAIR project.
Chapter 4 gives an overview of the different tracks, evaluations tools and indicators.
It is advised to also check the annexes, as they contain practical material, such as questionnaires for setting up the evaluation activities.

- Annex I: Usability, UX and acceptance scales for the user survey
- Annex II: Think-aloud exercise
- Annex III: Pilot Performance indicators
- Annex IV: Technical Performance indicators
- Annex V: Impact indicators list

1 Introduction

This document outlines the evaluation framework of the hackAIR project that has the objective to evaluate **(1)** the usability, UX, effectiveness and acceptance of the hackAIR platform, and **(2)** the established impact of the platform at the social and environmental level in the communities. This deliverable is part of Task 7.4 "Evaluation and impact assessment", in which the scope and setup of the evaluation framework is defined in this document (T7.4.1). Next, according to the stipulated evaluation framework, user feedback from diverse end-user groups will be collected during the three implementation periods of the pilot sites (T7.4.2), resulting in one interim report (led by NILU, D7.4, M26), a final pilot implementation report (led by NILU, D7.5, M36) and a final impact assessment deliverable (led by VUB, D7.7, M36). The results of the behavioural change analysis, of which the strategy is outlined in D6.2, will also be reported in this later final impact assessment deliverable.

The timing and output of deliverables is outlined below in figure 1:

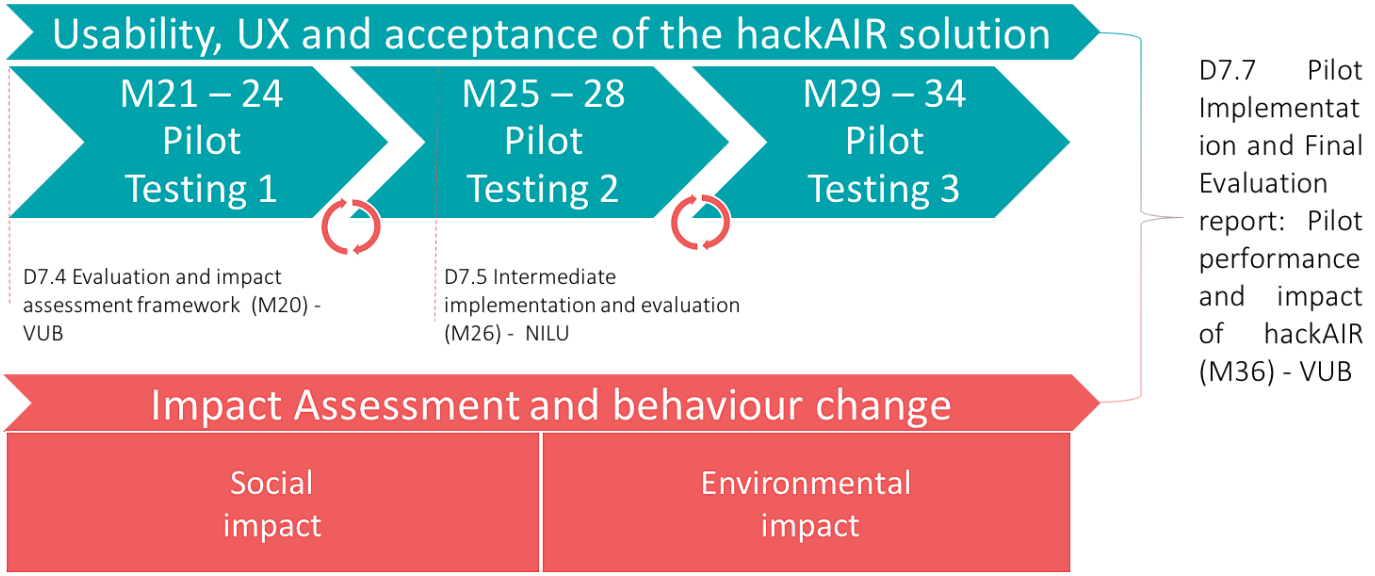

Figure 1: Timing of WP7 pilot monitoring and project evaluation activities

VUB is taking the lead for the development and coordination of the evaluation and impact assessment framework, as well as for the usability, user experience, effectiveness and acceptance assessment. DRAXIS coordinated the gathering of technical indicators. For the execution of the assessment framework different partners will have to take up responsibilities:

- NILU and BUND, the coordinators of the pilot sites in Norway and Germany to collect user feedback with end-users about the usability, user experience and acceptance of the hackAIR solution (platform, mobile application and sensors). Although the indicators and evaluation materials are developed by VUB, and the results will be reported in the pilots' deliverables (D7.4 and D7.5), pilot partners will have most contact with participants.
- DRAXIS for the logging of the usage of the platform and application (via e.g. Google analytics) and the development and integration of the survey feature into the hackAIR application for regular push services to the participants' profiles.

1.1 Purpose of the evaluation framework

The hackAIR evaluation framework will identify proper evaluation methodologies, tools and indicators to answer the following research questions:

- What are the main determinants for using the hackAIR open platform, from the angle of user experience and user acceptance? To answer this research question, VUB will identify usability and effectiveness indicators, such as ease of use, task completion rate, overall satisfaction scores, etc. and monitor this during the three pilot implementation periods. During the pilot implementation trials, standardized survey forms will be distributed to collect feedback from participants in the pilots. The aim is to achieve by the end of the trials the stated objectives in the DoA (p.5), i.e. being an 80% acceptance rate of the hackAIR platform by end-users, a user satisfaction level of 90% and to achieve a usability score of the hackAIR platform of 80%. Furthermore, pilot indicators are developed to evaluate the effectiveness of the tools after each implementation period. Last, also technical indicators will be taken into account that will be monitored by DRAXIS. NILU also developed KPI's that will be evaluated by consortium partners via an internal review process of the hackAIR tools.
- What is the social and environmental impact for the involved communities by using the hackAIR solution? To answer this research question, VUB will rely on the former established impact assessment methodology of the CAPS project, 'Impact Assessment for Social Innovation' (IA4SI) and will contextualise it to the hackAIR project. Specific indicators and tools will be chosen to evaluate the established social and environmental impact at micro and meso-level, i.e. changes that occurred at the individual level and the community level of using the hackAIR solution. To assess both the micro and meso level, two different target groups will be involved in the analysis: (1) end-users of the hackAIR solution to assess the micro-level, and (2) pilot coordinators through self-assessment to assess the organisational level.

Besides these two main research questions, the evaluation framework will also report about the following research question, of which the strategy was previously outlined in D6.2:

- How effective were the behavioural change interventions for participating in air quality monitoring and adopting a more pro-environmental behaviour, and which factors influenced the process of behaviour change? The outcomes of the behavioural change analysis will be reported in the final impact assessment deliverable, based on the outlined strategy of D6.2.

The evaluation framework will be sub-divided in parallel to the three periods of pilot implementation, and paired with a set of ex-ante and ex-post evaluations. By doing so, the evaluation can validate whether the (corrective) changes of a prior period are correctly implemented with subsequent periods (i.e. period 1 corrective actions for period 2), subsequently reflecting in an enhanced usability, user experience and impact results.

1.2 hackAIR's evaluation framework

Based on the former stipulated research questions, the hackAIR evaluation framework will exist of three different tracks. For track 1 and track 2, the following chapters in this deliverable will briefly discuss the current state of art in literature, identify tactics and propose adequate evaluation tools to perform the assessment. For further information about track 3, it is advised to read D6.2 "Behavioural change techniques".

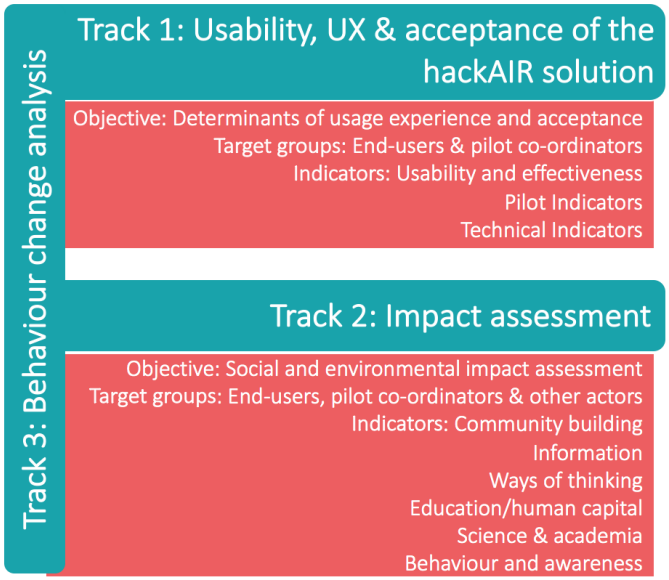

Figure 2: Evaluation framework of the hackAIR project

Track 1: Usability and user experience (UX) of the hackAIR solution

This evaluation track will set up activities to identify the main determinants for using the hackAIR solution (platform, mobile application and sensors), from the angle of user experience and user acceptance. In this track, a mix of qualitative and quantitative methodologies will be chosen to collect feedback from two different target groups. Both end-users of the hackAIR solution and pilot coordinators will be involved in these assessment activities. The evaluation activities for this track will run between M21 and M34 of the project, and will be periodically reported in D7.4 (M26) and D7.5 (M36), and will be organised through a phased-approach: period I (M21-M24), period II (M25-M28) and period III (M29 – M34). After each period, the evaluation results should validate whether the (corrective) changes of the former period are correctly implemented in the current period, and so reflect in an enhanced user experience. To do so, the same type of assessment scales will be used in standardized survey forms over the three periods in order to monitor the change over time.

The objectives to be reached within this track are the following:

- Acceptance of the hackAIR platform by end-users: 80%
  - The number of participants with health issues: at least 10% of the participants choose the option to stay informed about health-related air pollution information
  - The age balance of the participants: at least 10% of the participants are above the age of 50
  - The amount of basic and advanced/dedicated users in Norway: 5000/200
  - The balance between basic and advanced/dedicated users in Norway: 3000/90
- User satisfaction level: 90%
- Usability of the hackAIR platform: 80%

Track 2: Impact assessment

This evaluation track will set up activities to investigate the social and environmental impact for the involved communities by using the hackAIR solution. VUB is responsible for the identification of the social and environmental impact indicators. In this track, mostly qualitative methodologies, such as focus groups, impact assessment workshops and interviews, will be chosen to collect the feedback from two different target groups. Likewise track 1, both end-users in the communities (micro level impact) and pilot coordinators (meso level impact) will be involved in these assessment activities. The evaluation activities for this track will run between M21 and M34 of the project, and will be reported in the final impact assessment deliverable (D7.7, M36). The impact assessment activities will start with establishing a baseline of indicators for each pilot through a workshop with pilot coordinators. Based on these

validated indicators, the established impact will be assessed in each pilot location at the end of each implementation period (M24, M28 and M34). Together, the identified indicators in track 1 and 2 will complement each other in providing a systematic analysis of the performance, outcomes and subsequent impact of hackAIR both as a project, and as a collective awareness platform for air quality.

The objectives to be reached within this track will be evaluated in a more qualitative way, therefore, no measurable KPI's are defined.
The objectives are the following:

- An impact on online community building
- An impact on local community building
- An impact on access to and sharing of information
- An impact on changes in ways of thinking
- The impact of training provided by the project
- An impact on knowledge sharing
- An impact on citizen science
- An impact on user's environmental awareness, opinions and behaviour

Track 3: Behaviour change analysis

The strategy of the behaviour change analysis can be consulted in D6.2 Three different types of experiments are chosen to measure the change, namely pre-and post the participation in hackAIR workshops, during the usage of the gamification features on the platform, and when consulting the tips of the day. The latter two experiments will especially investigate the logging details on the platform with frequent users, and with users who do not use these specific features. As such, it will be able to measure their effectiveness in reaching behaviour change. The behaviour change analysis will also give input to the social and environmental impact indicators of track 2.

The objectives to be reached within this track are the following:

- Increased awareness: 85%
- Effectiveness in promoting behavioural change: 80%

2 Evaluating usability, UX & acceptance and pilots of the hackAIR solution

This chapter discusses track 1 of the hackAIR evaluation framework, i.e. the evaluation activities for assessing the usability, user experience and acceptance of hackAIR solution (platform, mobile application and sensors) with end-users and pilot coordinators.

2.1 List of indicators – Track I

2.1.1 Usability and effectiveness indicators

2.1.1.1 Set 1: Acceptance indicators

When designing a digital platform, accompanied by a companion application and integrating data collected through low-cost sensors, technological acceptance must be considered during multiple stages. It is the objective to evaluate the acceptance rate of the hackAIR solution, therefore the following literature review presents some models that explain the intention to use a new technology. Firstly, the Technology Acceptance Model (TAM) remains a central model in various domains to predict technology usage. The model's roots lie in the theory of reasoned action, as beliefs (ease of use and usefulness) and attitudes are linked towards behavioural intention, and so actual usage. Attitude is in turn hypothesized to be dependent on perceived usefulness and ease of use, which are both assessments of the consequences of using a system to accomplish a task. The relation between the 'ease of use' and intention to adopt is not very clear. Some studies such as Davis (1989, Figure 6) report that ease of use is not directly influencing the intention to adopt such as Davis while others (e.g. Warkentin, Gefen, Pavlou, & Rose, 2002) do report such as relation. Davis defines perceived usefulness as "(…) the degree to which a person believes that using a particular system would enhance his or her job performance" and perceived ease of use as "(…) the degree to which a person believes that using a particular system would be free of effort (…)". Hereby, it is predicted that the perceived ease of use has a causal influence on the perceived usefulness, as 'the easier a system is to use, the more useful it can be' (Davis, 1989).

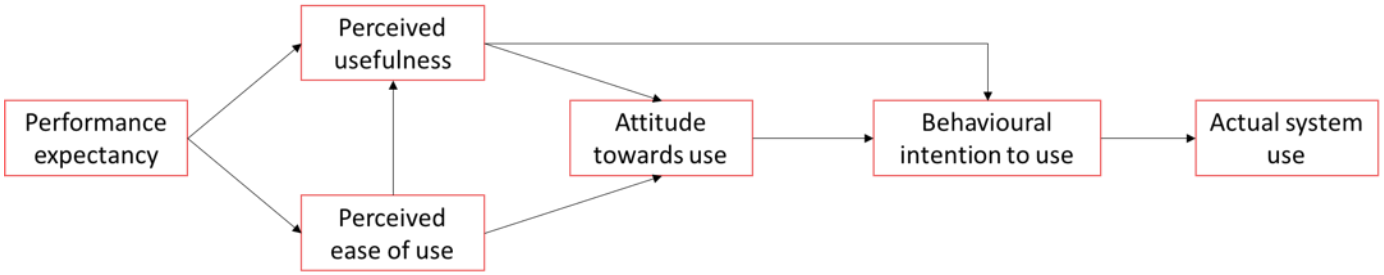

Figure 3:TAM Model. Source: Davis (1989)

In an attempt to bring together eight different existing models commonly applied to understand and predict user acceptance, one of which included the TAM model of 1989, the unified theory of acceptance and use of technology (UTAUT) was developed (Venkatesh, Morris, Davis, & Davis 2003). As illustrated in Figure 4 below, this theory identified four central constructs that directly influence user acceptance and usage behaviour: performance expectancy, effort efficiency, social influence and broader facilitating conditions. Performance expectancy stands as the clearest predictive construct for intention to use. Gender, age, experience and voluntariness of use represent conditional moderators for strength of causality between the four central constructs and intention. For example, the age of employees or workers is likely to moderate the importance placed on extrinsic rewards, a driver of performance expectancy. Gender, age and experience moderate all constructs in their influence on behavioural intentions to varying degrees, except facilitating conditions. Defined as "the degree to which an individual believes that an organizational and technical infrastructure exists to support a defined system" (p. 453), facilitating conditions are in fact considered

insignificant in influencing behavioural intention. Rather, when moderated directly by age and experience, facilitating conditions have an unconstrained and strong impact on usage behaviour.

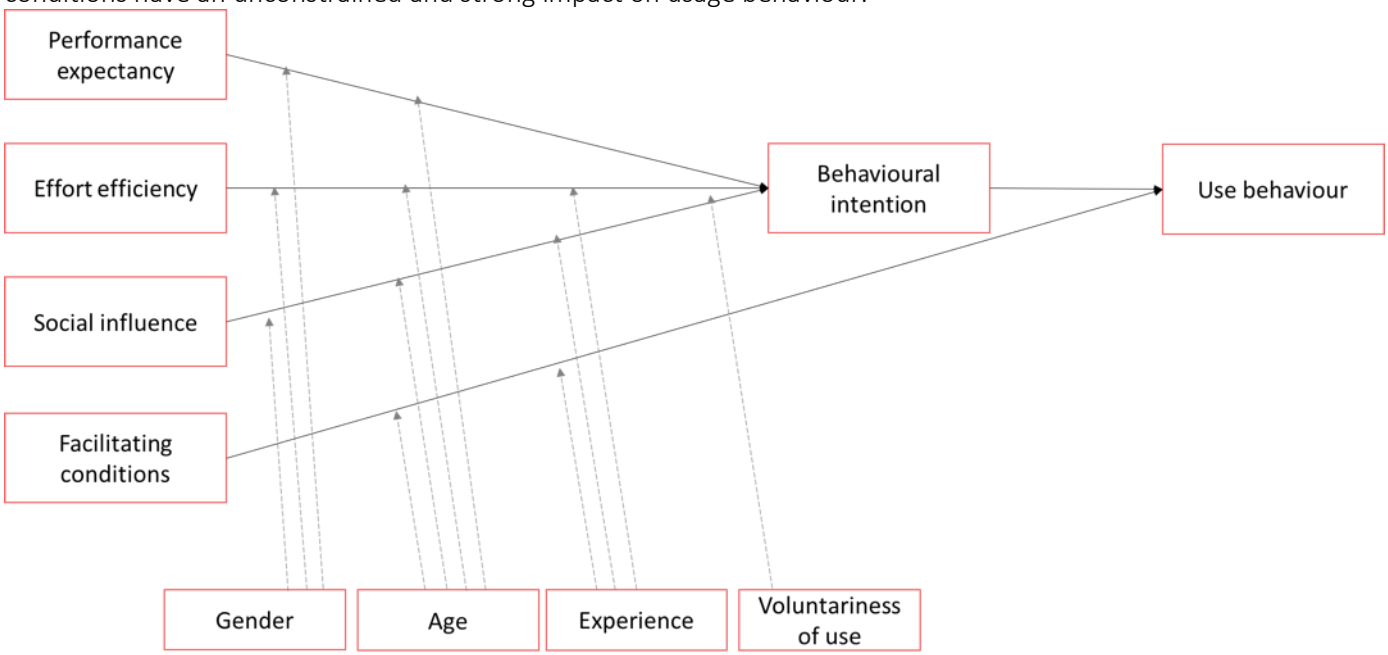

Figure 4:UTAUT Model. Source: Adapted from Venkatesh, Morris, Davis, & Davis (2003)

As mentioned above, the elements of both the original and modified TAM model were investigated and integrated into the UTUAT where considered relevant. For performance expectancy, perceived usefulness was included as an element. This is the degree to which a person believes that using a particular system would enhance his or her job performance. (Davis, 1989). In the context of effort expectancy, perceived ease of use was also considered relevant. This is the degree to which a person believes that using a system would be free from effort. Interestingly, attitude towards behaviour – an individual's positive or negative feelings about performing the target behaviour – was considered not to be causally responsible for an impact on usage in isolation. Previously, it was argued that an individual would hold a favourable attitude towards a specific behaviour or action if they believed that it would result in positive outcomes, and vice versa (Ajzen and Fishbein, 1975). As the hackAIR platform is intended to appeal to intrinsic motivations (see D6.1: Engagement strategy for community involvement), and given the centrality of this variable in the TAM model, we consider it reasonable to hypothesise that the attitude towards a specific technology will in fact have a direct influence on usage.

Given the influence of gender, age and experience on behavioural intention, we have grouped these three elements as one single measure, named 'user profile'. Table 1 below located user profile alongside attitude as social context measures. Additionally, perceived ease of use, perceived usefulness and intention to use have been included from TAM as measures of user acceptance for hackAIR.

Study	Measure[1]
Acceptance	Perceived ease of use
	Perceived usefulness
	Intention
(Social) Context	Attitude
	User profile

Table 1: Measures of the Evaluation framework (set I): measures from TAM and UTAUT perspective.

[1] All scales that will be used to measure the indicators are described in the annexes of this deliverable.

In order to guide our evaluation, our acceptance measures are connected based on the following hypotheses:
H1: Perceived usefulness has a significant impact on the intention to use the hackAIR solution
H2: Perceived ease of use has a direct impact on perceived usefulness, the easier it is thought to use the hackAIR solution, the greater the expected benefits
H3: Perceived usefulness and perceived ease of use have a direct impact on the attitude formation about the hackAIR solution
H4: Attitude is a key determinant of the intention to use the hackAIR solution

2.1.1.2 Set 2: Experience indicators - Perceived enjoyment

Given the centrality of the gamification and implications for the hackAIR platform, as well as the various implications for engagement and behaviour change objectives, the property of 'perceived enjoyment' (PE) was taken into consideration when gathering evaluation indicators. In an effort to extend and sharpen the causal links at play with technological acceptance, Sun and Zhang (2006) directed attention towards one specific relationship; the direction and strength of causality between perceived enjoyment (PE) perceived ease of use (PEOU). PE is defined as "the extent to which the activity of using computers is perceived to be enjoyable in its own right, apart from any performance consequences that may be anticipated" (Davis et al., 1992). Given the intrinsic quality of this quality, the potential role of PE in the acceptance of a gamification system cannot be underestimated. Acceptance and user experience studies have commonly experienced difficulties in both distinguishing between PE and PEOU due to their overlapping characteristics, as well as understanding the direction of causality between the two. Conceptually, Venkatesh (2000) contends that the primacy of intrinsic and extrinsic motivations as building blocks in motivational models, contrasted to the claims made by Davis in the TAM, have naturally led to contradictory findings. Using data from two case studies, the abovementioned research empirically established the dominance of PE in influencing perceived ease of use. This relationship, in extension of the 1989 TAM, has been illustrated below in Figure 4.

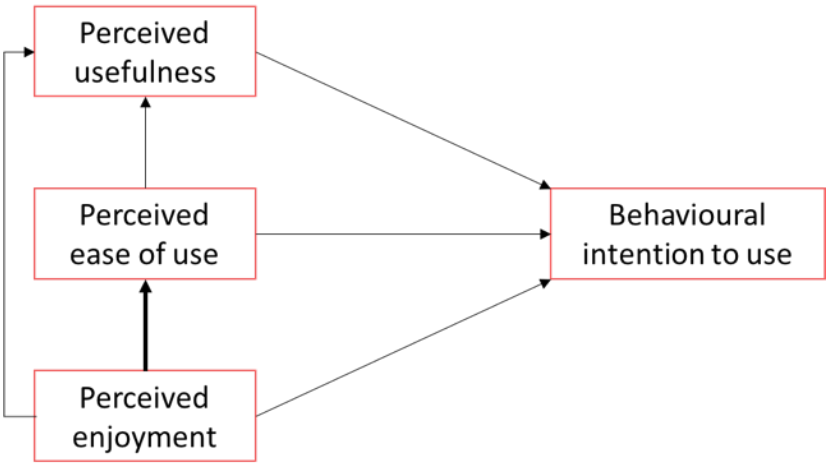

Figure 5: Theorisation of dominance between PE and PEOU. Source: Adapted from Sun and Zhang (2006)

Using this study as a precedent, PE was selected as a user experience measure and integrated into the evaluation framework.

Study	Measure
Experience	Perceived enjoyment

Table 2: Measures of the Evaluation framework (set II)

The causal direction of PE in influencing PEOU has also been connected to acceptance factors from the TAM, which will be evaluated through the following hypotheses:

H5: The perceived enjoyment of using the gamification approach of hackAIR has a direct impact on the overall perceived ease of use of the solution

H6: The perceived enjoyment of using the gamification approach of hackAIR has a direct impact on the overall intention to use the solution

H7: The perceived enjoyment of using the gamification approach of hackAIR has a direct impact on the overall perceived usefulness of the solution

2.1.1.3 Set 3: Experience indicators - User satisfaction

In contrast to the predictive characteristics of the TAM model and variants, user satisfaction literature has been traditionally associated with the development of the system and design properties of IT. Although developed in parallel research tracks, efforts have been made to combine the concepts underpinning user satisfaction and technology acceptance. Not least, work from Wixom & Todd (2005) stands as a commonly applied research model that combines both characteristics. Remaining consistent with the TAM model, this model proposes that behavioural beliefs and attitudes can predict behavioural intention to use. The causal link to user satisfaction is extended through the introduction of object based parameters, derived from existing user satisfaction literature. By doing so, the paper argues that the beliefs and attitudes towards a system, certain tasks or the respective brand can be in part influenced by their satisfaction with both the system and included information.

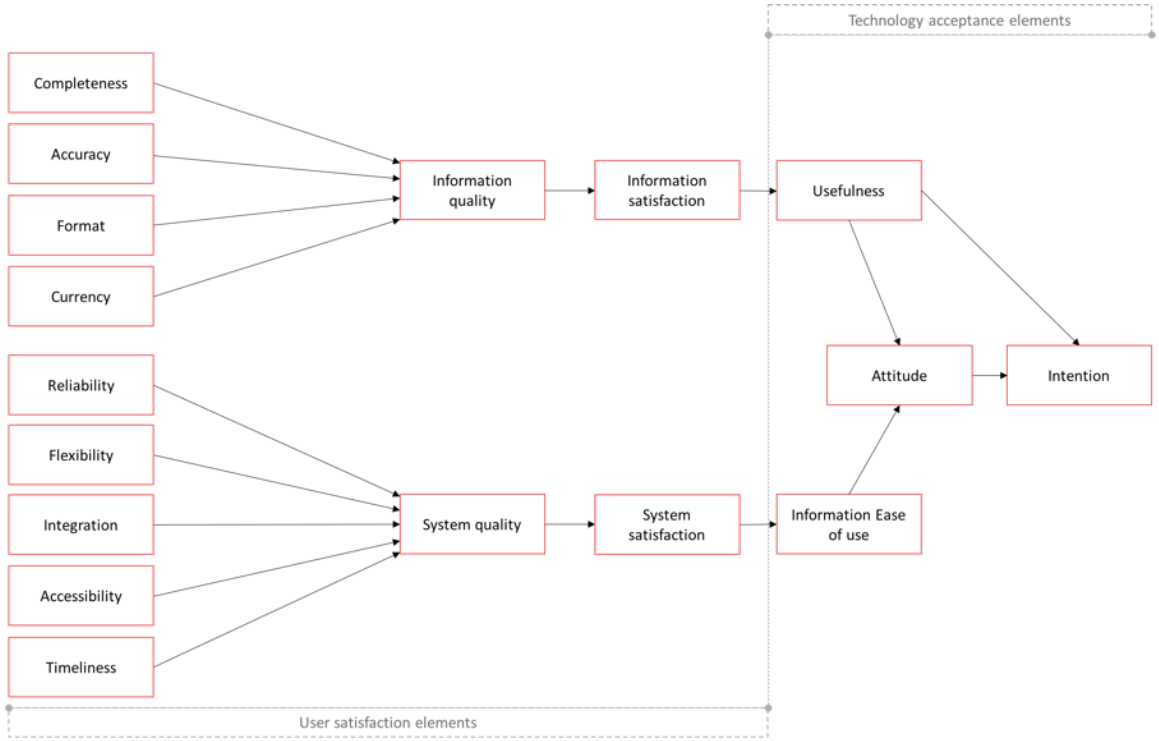

Figure 6: Research model of user satisfaction and technology acceptance. Source: Adapted from Wixom & Todd (2005)

As indicated on the left side of Figure 5, antecedents of system quality include *reliability*, *flexibility* (adaptive qualities), *integration* (data integration from sources), *accessibility* (data extraction and access) and timeliness (response time). Four elements also shape user perceptions of information quality: *completeness* (of information provided by the system), *accuracy* (if the information is viewed as correct), *format* (how well the information is presented) and *currency* (if the information is up to date).

Given its practical strength in combining both user perceptions, system characteristics and technological acceptance parameters, user satisfaction has been selected as a measure for the hackAIR evaluation framework. Secondly, this

measure allows for flexibility in assessing user satisfaction of individual or collective features on the hackAIR platform, some of which are particularly crucial in assessing impact at the end of the project.

Study	Measure
Satisfaction	User Satisfaction

Table 3: Measures of the Evaluation framework (set III)

User satisfaction is expected to interact with other parameters using the following hypotheses:
H8: User satisfaction influences perceived ease of use
H9: User satisfaction influences perceived usefulness
H10: User performance influences perceived ease of use

2.1.1.4 Set 4: Usability indicators - User performance

User performance is a strong indicator of the usability of a system, and will tell you which parts of the system need to be improved. If you know that users are not able to perform a certain action on the platform, then you know that users are not using the platform well and specific parts have to be redesigned. User performance can be measured through several performance metrics, in hackAIR it will be measured through "task success" and "errors on task":

- Task success: It measures how effectively users are able to complete a given set of tasks. This can be reported through binary success rates (1= success, 0= failure), or levels of success (complete success, partial success, failure). From the success rates, average scores can be calculated.
- Errors: reflect the mistakes during a task. Errors can be useful in pointing out confusing or misleading parts of an interface. Errors are not the same as bugs, a bug is an error or a failure in the software system that prevents it from behaving as pretended. Errors are encountered issues (e.g. failure in task completion, or a user does not know what to perform next), and a severity rating can be used for this:
- Low severity rate: Any issue that annoys participants, but does not play a role in task completion. The user is still able to complete the task
- Medium severity rate: Any issue that was encountered, but did not directly prevent task completion. It has a greater impact on efficiency and satisfaction.
- High severity rate: Any issue that leads to task failure. Significant impact on efficiency and satisfaction.

In the following table, the main features of the hackAIR solution are listed once more with the specific requirement number from D2.4 "Report on co-creation of services", and linked with specific user performance tasks. For the evaluation of the hackAIR solution, users should be able to successfully perform the following actions in terms of «task success», i.e. the extent to which users are effectively able to complete a given set of tasks in a user-friendly way (independent, and time constrained). The following task descriptions can be used with end-users when they are performing, either fictive or real-life, scenarios with the platform:

Feature on the hackAIR platform	User performance (task success)
The visualisation component: The web and mobile app should have a clear visualisation of the air quality, e.g. through a map (R18, also R1, R1e, R2, R17, R19)	• Users should be able to access the visualisation map with air quality data of their local area • Users should be able to filter (sensor data, subjective data, pictures, etc.) the different air quality sources on the visualisation map • Users should be able to access the timeline and evolution of the air quality levels

	• Users should be able to interpret the data (e.g. what is the difference between real-time and near real-time, what is particulate matter, etc.) • Users should be able to identify the initial source of the data • Users are able to compare the air quality in their city with other locations • Users are able to contribute with own data to the map
Gamification on the hackAIR platform	• Users should be able to successfully receive a badge based on their activity on the platform (automatically granted by the platform)
Subjective perceptions: Users should be able to upload how he/she perceives the air quality in their surrounds, by entering a rating on the air quality for a specific location (R56)	• Users are able to access the contribution page for perceptions • Users are able to submit a subjective rating • Users are able to see their contribution on the map
Personalised recommendations: (R32)	• Users can change their profile so they will receive personalised recommendations when they
Communication and interaction features: There is forum or message board to discuss relevant topics (R29, R30, R31)	• Users are able to ask a question to the hackAIR community • Users are able to see how many people in their neighbourhood are contributing with measurements
Features on the hackAIR mobile application	User performance (task success)
R52: Users should be able to contribute with geo-tagged and time-stamped sky pictures that they can submit via the hackAIR mobile app	• A user is able to take a correct picture of the sky to estimate air pollution • A user is able to send the photograph to the web app of hackAIR
Features of the hackAIR sensors	User performance (task success)
	• A user is able to install the sensor correctly • A user is able to send data to the hackAIR platform through the sensor

Table 4: User performance tasks on the hackAIR platform

Study	Measure
Usability	User performance

Table 5: Measures of the Evaluation framework (set IV)

2.1.1.5 Summary of indicators

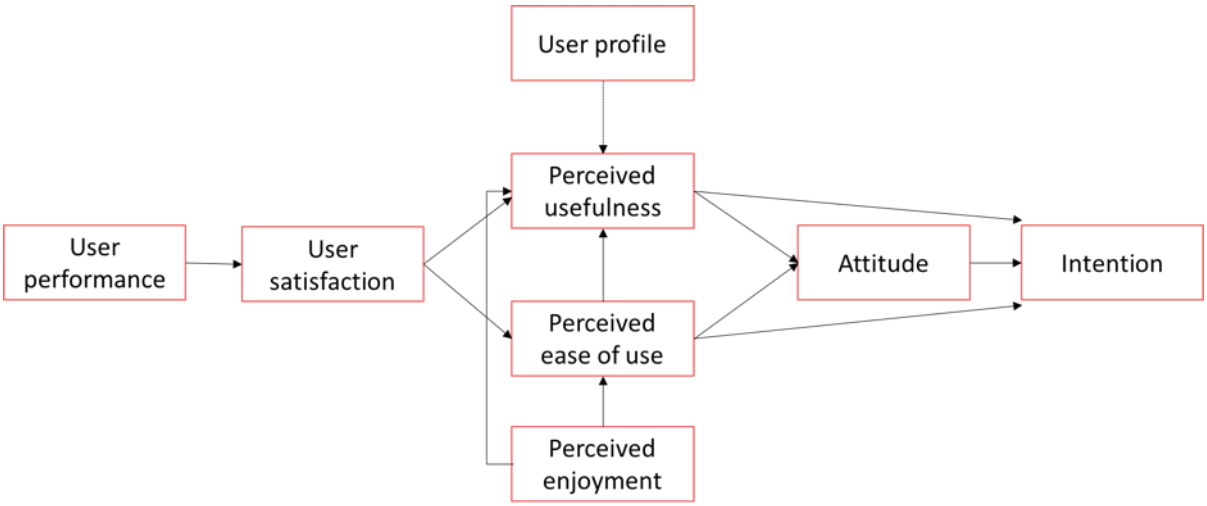

Figure 7: hackAIR indicators for usability, UX and acceptance

Indicators	Definition	Reference
Perceived ease of use (PEO)	The degree to which a person believes that using a particular system would be free of effort	TAM: Davis, 1989
Perceived usefulness (PU)	The degree to which a person believes that using a particular system would enhance his or her job performance	TAM: Davis, 1989
Intention (I)	The degree to which the system does not undermine the intention of its use	TAM: Davis, 1989
Attitude (A)	An individual's positive or negative behaviour towards innovation adaption	TAM: Davis, 1989
Perceived enjoyment (PE)	The extent to which the activity of using computers is perceived to be enjoyable in its own right, apart from any performance consequences that may be anticipated	HSAM model
User profile (UI)	Demographics such as age, gender, experience, etc.	UTAUT: Venkatesh, Morris, Davis, & Davis (2003)
User satisfaction	Users' subjective reactions to using the system	(Wixom & Todd, 2005)
User performance	The degree to which users can successfully accomplish a task, or set of tasks	(Albert & Tullis, 2013)

Table 6: hackAIR evaluation indicators, definitions and origins

2.1.2 Pilot performance indicators

NILU has developed a set of pilot pre-performance indicators (Key Performance Indicators, KPIs) to evaluate the level of the success for the hackAIR platform, including the indicators to evaluate the main toolkit towards its planned objectives. These indicators are described in D7.1, but they will only be assessed by hackAIR consortium members that developed the tools (DRAXIS and CERTH) and those who have tested the tools in practice and used them in their pilot activities (NILU and BUND). This feedback will be used as input for optimizing the hackAIR platform and tools before they are used to engage citizens.
In addition to this pre-pilot assessment, VUB also defined some pilot indicators that are related to the effective usage of the hackAIR application, but also of the different pilot activities and interventions.

Every pilot partner should gather following data:

Description	KPI	Measurement channel
Total number of participants on different points in time (unique and returning – evolution over time) (divided by location and age, and type of participant)	Sept/Oct 2017: 100 Dec 17/March 18: 1000 April/July 2018: 10000 Aug/Dec 2018: 10000	Log file
Amount of participants to workshops (divided by gender, location, type of workshop)	At least 7 workshops with in total at least 50 participants	Counting of participants by workshop coordinators & survey after/before workshop
Amount of participants to the measurement campaigns	At least 5 measurement campaigns with in total at least 30 participants	Counting of participants by pilot coordinators
Amount of implemented feature/bug requests of the participants or pilot coordinators released during de trial periods	70%	Release notes
Amount of interventions (communications, explanations, …) done by hackAIR experts to the citizens on their gathered data.	At least 25	Count of helpdesk interventions + face to face interventions
Amount of hardware sensors assembled and installed by participants	100	Log file

Table 6: hackAIR Pilot KPI assessment scale

2.1.3 Technical performance indicators

The following table lists the technical performance indicators that will be monitored and analysed by DRAXIS with the help of other technical partners if needed. The indicators focus both on the hackAIR platform, usage of the sensors, but also the fusion map, the environmental node discovery module, the social media monitoring tool and the image analysis module:

Objective description	Measurement process/unit	Indicator
Platform performance	Number of seconds between upload of picture to platform and visualisation of picture on the platform/app.	< 20 sec
Platform performance	Availability of the web platform pilot trial periods (uptime)	90%
Platform performance	Load time of fusion map (secs)	5 secs
Platform performance	Upload rate of AQ information (mins)	<10 mins
Platform performance	Page speed load (only source code)	<5s
Platform performance	Full page load time (average)	< 15s
Platform performance	Page speed load (only ping) with 150 concurrent requests	< 15s
Mobile App performance	Crash free sessions	95%
Mobile App performance	Non-fatal issues / Total sessions	< 1
Platform performance	Update frequency of the data fusion map	< 3 h
Platform performance	Averaging period of the data fusion map	<= 24 h
Platform performance	Uptime of the data fusion map	> 95%
Environmental node discovery module	Average number (during pilot execution) of image-based AQ estimations in the pilot countries (number of fresh measurements obtained in the last 24 hours)	> 50
Environmental node discovery module	Average number (during pilot execution) of AQ measurements (images + ground stations/sensors) in the pilot countries. (number of fresh measurements obtained in the last 24 hours)	> 100
Social media monitoring tool	number of discovered social media accounts relevant for communication strategy	> 500
Social media monitoring tool	number of "insight" sessions between communication managers and social media monitoring tool managers	1 per quarter
Image analysis module (i.e. sky detection)	Precision of the sky detection module (% of image-based AQ estimations coming from sky-depicting images)	

Ontology and Reasoning component	Execution time should be comparable to that of State-of-the-Art reasoners on a single task (time benchmark)	yes

Table 7: Technical performance indicators of the hackAIR platform.

2.2 Quantitative and qualitative evaluation tools

2.2.1 Standardized survey forms

Questionnaires will be used during all implementation trials to ask about user experience and acceptance. The question structure will use Likert scales, but questions can also be of an open nature, ranking of items and so forth. The questionnaire has the advantage of scale, price, cross-sectional, quick turn-around, and the data can be directly captured and stored using specific software (Qualtrics and SPSS). There are, however, some important drawbacks in employing questionnaires such as coverage error, low response rates and anonymity may be troubling sources of measurement errors. Another issue is non-response error that may be caused by factors such as the interface or other technical problems of the (survey) software.

However, in hackAIR, standardized survey forms will be used to reduce low response rates. These questionnaires will be very short, and will have a fixed set of questions that should be asked in each pilot implementation period to track progress about the user performance and user satisfaction. The data will be analysed using SPSS using various analysis techniques, such as descriptive analysis, regression analysis, crosstabs and Chi-squares, etc.

The standardized survey forms of which the scales that will be used during different phases are to be find in the annexes. Next to this, also log files and forms that should be filled in by pilot coordinators will be developed.

The survey form coordinated by VUB, to evaluate the usability, UX and acceptance indicators will be a form with a limited amount of questions in Likert-scale form. A set of fixed questions will be asked in each pilot implementation period to track progress. Additional questions can be asked to ask specific input about an updated feature, or specific user requirement. It is advised that this form has a link on the hackAIR platform, and if possible, is also distributed during other activities where you meet with end-users of hackAIR (e.g. during a workshop).

2.2.2 Think-a-loud protocol

A think-a-loud protocol is a method that requires participants to verbalize what they are doing and thinking as they complete a task, to understand the thought process of a subject as they use the hackAIR solution (e.g. revealing aspects of an interface that delight, confuse or frustrate). By thinking aloud while attempting to complete the task, users can explain their method of attempting to complete the task, and illuminate any difficulties they encounter in the process. The method is commonly used for either low- or high-fidelity (paper) prototypes, and can be used to evaluate the prototype, or to help identifying new features.
A think-aloud session is usually done with one participant and one observer and takes approximately 30 to 45 minutes. The moderator gives in advance a set of tasks that the user needs to complete on the hackAIR platform through a real-life scenario, e.g. register, set your profile settings and upload a sensor measurement to the hackAIR visualization map. During the completion of the scenario tasks, the moderator reminds the participant to verbalize what they are thinking as they work through the task. The focus of the test should be on what is happening, as opposed to why. It is important that the moderator does not help the participant (exceptions can be made if the user asks for help, or encounters an error). The moderator sits beside the participant, as this is a one-to-one exercise. Ideally, the actions on the screen of the participants can be (video) captured (e.g. software: Camtasia). Afterwards, some in-depth questions are asked about the general experience of the users (e.g. what were the main issues you experienced, what frustrated you, do

you have ideas for additional features, ...). It is advised to take notes during the task completion, so that difficulties or remarkable/unexpected events can be discussed afterwards.

2.2.3 Logging framework

The goal of a logging framework is to track user activities when using the hackAIR services. Logs are necessary to provide appropriate insights into authentication, authorization, gaining insights into critical user activity as well as for debugging and performance monitoring. More precisely, information can be logged based on user- driven events (e.g. 'User uploads a sensor measurements"). These user events can be logged via a set of web services, adopted or designed specifically for this purpose, and, in addition, by using scenario-type questions in the questionnaire user behaviour can be tested and validated. The logged information may include response time, click patterns, evaluate any errors experienced while using the platform etc. It is important that the logging is gathered and presented in a unified matter, because we are dealing with a distributed environment and systems (e.g. database server, web server). The second factor is that one must pay attention not to burden the system with too many logging events, which would increase the processing and reduce the overall system performance. The log data will be analyzed using statistical tools such as descriptive statistics.
The following tables describe the type of statistics we would like to log:

- General platform statistics (for each local instance)
- Registration statistics (for each local instance)
- Statistics for the visualisation map on the hackAIR platform
- Statistics for the mobile application
- Statistics for the sensors
- Statistics for the engagement and behaviour change tactics: badges, tips of the day, personalised recommendations

General platform and user statistics
Total number of users on different points in time (unique and returning – evolution over time)
Total number of visitors of each community instance (unique and returning, divided by age)
Country of origin of the visitors of each community instance
Total number of visitors that choose the setting 'show health information' in the hackAIR profile
Average duration sessions of each community instance
Most visited pages/least visited pages
Registration statistics
Number of registered users/members on each community instance (with possibility to make a difference in simple and advanced users and different age groups)
Number of registered users activating the game module
Statistics for the visualisation map
Total number of sensor measurement contributions
Average duration of a sensor to provide measurement to platform
Total number of subjective perceptions
Total number of user-generated photographs
Total number of user-generated photographs that are validated as qualitatively valuable for usage in the air pollution measurement for the fusion map

Statistics for the mobile application
Total number of active users
Total number of contributions (uploading of photographs) made through the mobile applications
Statistics for multi-device/ modal usage
Number of participants that combine (login from): - smart phone & computer usage of platform - smart phone & tablet usage of platform - smart phone & tablet usage of the platform - tablet & computer usage of the platform
Statistics for the sensors
Total number of active sensors (at different points in time)
Total number of contributions made through the sensors
Statistics for the engagement and behaviour change tactics
Average duration sessions for 'tip of the day'
Total number of users/members that is using 'tip of the day' (on/off in settings)
Average duration sessions for personalised recommendations
Total number of users/members that is using 'personalised recommendations' (on/off in settings)
Total number of missions completed (complete rate vs opened missions)
Total number of badges rewarded
Total number of users/members participating in the gamification feature (on/off in settings)

Table 8: Logging framework of the hackAIR solution.

2.3 Planning and milestones

In the following table an overview is given of three pilot implementation periods and the evaluation activities that will be set up:

	Evaluation tool	User group	Evaluation measures
PERIOD I M21-M24	All evaluation tools: focus is on qualitative input	All end-user groups (starting with early adopters, and friendly users)	ALL MEASURES, with a focus on user performance
D7.4 "Intermediate pilot implementation report": listing of issues according to severance Issue implementation on hackAIR platform by the technical team			
PERIOD II: M25-M28	All evaluation tools – focus is on quantitative input	All end-user groups	ALL MEASURES, with a focus on user satisfaction & acceptance

PERIOD III: M29 – M34	All evaluation tools: focus is on quantitative input	All end-user groups	ALL MEASURES, with a focus on user satisfaction & acceptance
D7.5 "Final pilot implementation report" D7.7 "Pilot performance and impact assessment"			

Table 9:hackAIR planning for usability, UX and acceptance evaluation

For the three pilot implementation periods, milestones were defined for user performance, user acceptance and user satisfaction rates. These three indicators are key performance indicators of the project (see DoW p.5), and should result in user performance scores of 80% (as the most important indicator for usability in our framework), user acceptance rates of 80%, and satisfaction scores averagely around 90%. The following table presents the envisioned progress over the three periods:

Milestones period I (M21-M24)
hackAIR platform can be used for friendly user testing
Internal stakeholders and friendly users discover and use the first version of the platform
User performance rates are on average around 50%
User acceptance rates are on average around 50%
User satisfaction scores are around 50%

Table 10: Milestones period I- evaluation of user experience and acceptance of the hackAIR solution.

Milestones period II (M25-M28)
The hackAIR solution is up and running for each pilot and is fully functional
All types of end-user groups are able to use the hackAIR solution for various needs
User performance rates are averagely between 50%-70%
User acceptance rates are averagely between 50%-70%
User satisfaction scores are averagely between 50%-70%

Table 11: Milestones period II – evaluation of usability, UX and acceptance of the hackAIR solution.

Milestones period III (M29 – M34)
The hackAIR platform is operational for a large base of users
All types of end-user groups are able to use the hackAIR solution for various needs
User performance rates level up to 80%
User acceptance rates level up to 80%
User satisfaction scores level up to 90%

Table 12: Milestones period III: evaluation of usability, UX and acceptance of the hackAIR solution.

To define the completion of milestones, the following scale can be used:

Excellent or complete success: 100 - 90% user satisfaction, acceptance and performance scores
Good to very good implementation: 90-70% user satisfaction, acceptance and performance scores
Fair or moderate: 70-50% user satisfaction, acceptance and performance scores
Low or relatively poor: less than 50% user satisfaction, acceptance and performance scores

Table 13: Rating scale for usability, UX, and user acceptance scores.

3 Impact Assessment

3.1 Impact assessment methodologies: a brief literature review

This chapter discussed the second track of the evaluation framework, i.e. the impact assessment methodology. In the following sections, a brief literature review is performed about the methodology. The literature review is structured along several main questions:

What is an "impact assessment"?

In general, conducting an impact assessment refers to "the process of identifying the future consequences of a current or proposed action"(Vanclay, Esteves, Aucamp, & Franks, 2015). This means that from planning a certain action, intervention or project, one can reasonably foresee, predict and measure what the consequences are. As such, impact assessments are mostly done in advance of the action. An impact assessment that is conducted in advance is also called an "ex-ante assessment", with a prediction about the likelihood of the to be established impact of a planned intervention.

On the other hand, one can also choose to perform an assessment after the event, which is called an "ex-post assessment". This will then take the form of an evaluation of a particular project or policy.

Depending on the type of impact you are focussing on, one can choose to perform different types of impact assessments, e.g. social impact assessments, socio-economic impact assessments or environmental impact assessments. As a research field, social impact assessments (SIA) have been primarily conducted within sociology, rural sociology, environmental sociology and human geography, and were developed as a practice in the 1970s as a response to certain environmental regulations that were taken (Joyce & MacFarlane, 2001).

Seen the scope and objectives of the hackAIR project, the main focus of the impact assessment methodology will rely on the prediction and evaluation of mostly social impacts, i.e. the social consequences that are likely to happen from using the air quality monitoring tools to measure the local air quality and from the way communities are empowered to take direct action. It is opted to perform the assessment both pre-and post the project, seen the hackAIR project addresses diverse audiences through different tools. As such, it is hard to predict which specific impacts will be established in the two pilot communities. Furthermore, within the field of social innovation - in which the hackAIR project is active - there are not many yet established methodologies to measure outcomes and impacts in a robust way (Bund, Hubrich, Schmitz, Mildenberger, & Krlev, 2013), with the exception of the IA4SI methodology that will be discussed in the following sections. The impact assessment methodology will also specifically focus on the individual level (micro), and community level (meso), seen long-term and sustaining impacts of the project at the macro level are not able to occur during the project lifetime.

What are the objectives of an impact assessment?

Impact assessments are commonly employed as tools to either predict or assess a variety of associated impacts that can be attributed to action. Often incorporated into a methodology and applied in various stages, impact assessments should not be limited to solely monitoring activities; as integrated evaluation methodologies, they offer promise at policy, project or organizational levels.

As a monitoring activity, having tools in place to predict or periodically assess impacts allows managers or central actors to remain reflective of negative or positive consequences (Vanclay, 2012). As such, they offer opportunities to respond and avoid potential feedbacks from an action at a given time or place. As an evaluation activity, impact assessments provide a basis for assessing the multiple interactions and associated benefits that an action has produced during its lifetime. It is this form of evaluation that that will take place in the context of hackAIR – enabling the project to reflect on and assess the resultant benefits of hackAIR as a digital social innovation.
As a more recent development, impact assessments are now considered to be effective in evaluating not only negative, but positive implications. As argued by Vanclay (2003, p.6), the goal of impact assessments in their own right is to "bring about a more ecologically sustainable and equitable environment". One way in which this is demonstrated is through the growing appreciation of impact assessments that are methodologically built upon participatory processes. This can be done by including the perspectives of heterogeneous actors and local knowledge in evaluation activities. In fact, impact assessments are becoming more socially oriented in their focus by often incorporating locally-developed impact indicators.

What is 'impact'?

The European Commission's INFOREGIO UNIT defines impact as «a consequence affecting direct beneficiaries following the end of the their participation in an intervention, or after the completion of public facilities...» (European Commission, 2009). The established impact can come in various types and characteristics:

- Impact can be either short or long-term: some impacts can be observed immediately and are rather labelled as "outcomes", while other impacts only occur in the longer term. Longer-term impacts are usually referred to as sustainable results.
- Impact can either appear direct or indirectly
- Impacts can be either positive or negative
- Impact can be either expected or unexpected
- Impacts can be observed at the micro (individual level), meso (community level) and macro level (society level); sometimes impact on the micro and meso level are rather preferred to be labelled as "outcomes" as they occur before the impact at macro/society level is established.

According to Ebrahim & Rangan (2010), impact refers to results at the end of a project that benefit the society and communities at large, and is the consequence of project activities, outputs and outcomes. As such, impacts may not be confused with outcomes. Outcomes are mainly short-term effects, while impacts are long-term effects. Examples of outcomes in hackAIR are for instance the specific changes in behaviours through the delivery of the tools created for measuring air quality. A long-term effect would be if end-users of the hackAIR project create a habit out of using the tools in their daily life.
Different types of impacts can be generated by a project. The interventions done by a project can be very diverse, and as such create a multidimensional nature of impacts (Centre for Good Governance, 2006):

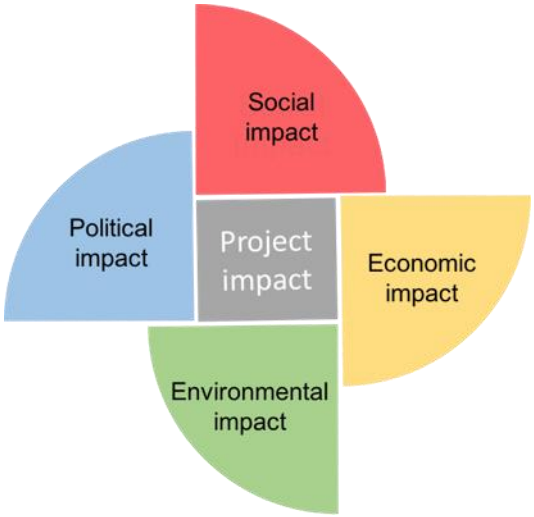

Figure 8: Multi-dimensional impacts of projects

The types of impacts that are of main concern for the hackAIR project are the 'social impact' and the 'environmental impact'. Social impact can be defined as "all the issues associated with a planned intervention (i.e. a project) that affect or concern people whether directly or indirectly" (Vanclay, Esteves, Aucamp, & Franks, 2015). Social impacts can be considered as changes either at the perceptual level, or the corporeal level. Almost anything can be a social impact as long as it is valued by or important to a specific group of people. Examples of social impacts are changes to people's way of life (e.g. how they live, work, play, interact, etc.), to their community (e.g. the interaction, cohesion, stability, etc.), to their culture (e.g. beliefs, customs, values, etc.), to their environment (e.g. the quality of the air), to their health and well-being, etc. Environmental impacts can be understood as positive or negative changes to the environment because of a certain action or project, e.g. pollution, urban regeneration, loss of biodiversity, etc. Environmental impacts can also be social impacts because people depend on the environment for their livelihoods, and because people can be attached to places.

Who are the actors of an impact assessment?

An impact assessment methodology should include the involved actors who might experience a perceived change in their lives from the proposed action, intervention or from the participation in a project. The impact should be identified, assessed and managed by these affected individuals or communities. Therefore, it is advised by the International Association of Impact Assessment (Vanclay et al., 2015), to set up participatory processes and deliberative spaces in which the impacts are discussed with the community members. Related to the hackAIR project, the following types of questions should be discussed with the pilot communities:

- Which types of impacts, either negative or positive, are most likely to be foreseen from participating in the project?
- Which impacts do you think will be established on the individual level, and on the community level?
- In which way do you think you will be impacted by your involvement in the project?

What are the phases of an impact assessment?

There are different proposed models for performing impact assessments. The one that is shown in the Figure below is from the IAIA (Vanclay et al., 2015). Their social impact assessment methodology consists out of five consecutive stages, which might overlap. They perceive SIA as a learning process, in which initial assumptions and preliminary understandings about impacts are validated and updated through an on-going process of consultations with project stakeholders, and especially the impact communities. For this latter mentioned reason, it was chosen to highlight the impact assessment phases of the IAIA:

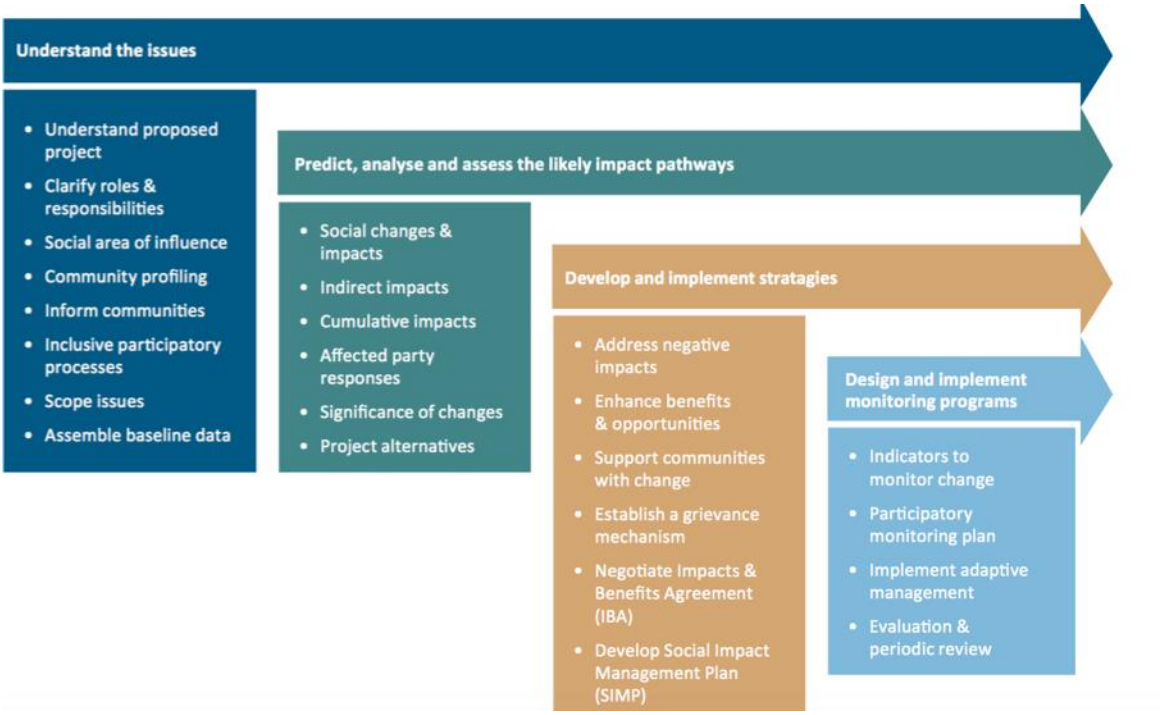

Figure 9: IAIA impact assessment methodology.

As seen from the Figure above, their impact assessment methodology is quite extensive. Seen the scope of the hackAIR impact assessment methodology, not all of these steps will be executed, also since most of the steps in phase I 'understand the issues' are already clear to project members and performed in other work packages activities (e.g. understand proposed project, gaining an understanding of the community through profiling, etc.) and since phase III is mostly about providing mitigation measures in case of negative impacts, most of these tasks will not be included in the hackAIR impact assessment methodology. However, the main pillars from their methodology will also be included in hackAIR: identification of social area of influence (i.e. identification of social impact indicators for the involved communities), assembling of baseline data (i.e. validation of the indicators), analysing and assessing social changes and impacts (i.e. collecting outcomes about how the involved communities will respond), and have indicators in place that monitor the change over time through evaluation and periodic review.

In the next section, the IA4SI methodology will be discussed that offers more practical guidance for setting up an impact assessment methodology for social innovation.

What is the IA4SI impact assessment methodology?

The "Impact Assessment for Social Innovation" project (short, IA4SI) was a coordination and support action that was granted funding in 2013. The objective of the project was to build a methodological framework to assess specific areas of impact of digital social innovations, and in particularly in the domain of Collective Awareness Platforms for Social Innovation (CAPS). The IA4SI project analysed 15 CAPS projects in total, and could as such establish a validated methodological framework for assessing impacts of social innovation. As such, the project contributed to the research field, as many methodologies for assessing the outputs and impacts of social innovations are still at an early development stage (Bund et al., 2013), and because social innovation is intended to produce positive changes (Phills, Deiglmeier, & Miller, 2008).

The developed methodological framework is a self-assessment methodology for evaluating projects in the field of digital social innovation through a mix of qualitative and quantitative methodologies, and has been validated with 15 other CAPS projects of which the results were gathered and analysed at the aggregated level. The IA4SI methodology uses eight synthetic indices: four of them are related to key areas of impact (social impact, economic impact,

environmental impact and political impact), and four are related to attributes of the innovation developed (efficiency, effectiveness, fairness, sustainability). Each index is composed of a number of dimensions, or sub-indices, and a number of variables:

Impact area	Sub-categories	Variables
Social impact	Impact on community building and empowerment	**Online community building** (e.g. number of users at the beginning and end of the project; amount of time spent by users at the beginning and end of the project, communication on the platform, network density, etc.) **Online community empowerment** (e.g. network diversity, number of groups created by users, tools for inclusion, tools for privacy management, etc.) **Local community building** (e.g. the capacity to enlarge the local community, better self-organisation, reduced power asymmetries) **Impact on social innovation CAPS community** (e.g. capacity to spread the social innovation, tools for networking among CAPS, number of collaborations with other projects within the domain, and outside the domain, etc.)
	Impact on information	**Access to and sharing of information** (amount of available information, improvement of access of info for users, capability to influence information asymmetries) **Quality of information** (instruments provided to assess the quality of information) **Data management policies** (policy in terms of standardization, content licenses)
	Impact on ways of thinking/opinions and changes in behaviour for individual and collective behaviour and lifestyles	**Changes in ways of thinking** (topics were opinion change is expected, activities performed to achieve the expected change, opinions and behaviours, number of people participating) **Change in behaviour** (topics were changes in behaviours are expected)
	Impact on education and human capital	**Training provided by the project** (hours of training provided, training efficiency, topics covered by training activities, tools for education developed by the project) **Impact on human capital** (impact on users e-skills, number of activities supporting the acquisition of digital competences, digital literacy, reduction of digital device, the self-assessment to improve skills of people employed in the consortium and of its users) **Impact on change in training curricula, educational policies**
	Impact on science and academia	**Knowledge production** (scientific impact: number of researchers, number of articles, number of patent applications, number of IPRs, project level of interdisciplinarity) **Knowledge sharing** (use of open access, sharing through social media, dissemination through project website, sharing through events) Impact on research processes

	Impact on employment	Impact on job creation Impact on European employment and within the social innovation sector (e.g. number of spin-offers created by the project) Impact on working practices and routines
Environmental impact	The environmental impact of the project itself	CO2 compensation, greenhouse gases production, reduction of air pollution in percentage,
	The impact on user's environmental behaviour	Project's capacity to provide easier access to low carbon technologies Number of compensation activities performed by users since their engagement with the project Project's capability to contribute to the change in users participated to environmental related actions Project's capacity to increase users' sensitivity towards the issue of air pollution
Economic impact	Impact on user's economic empowerment	Impact on access to finance (project's capability to increase the access to finance, impact through crowd funding) Impact on entrepreneurship (project's capability to support enterprises, and creation of new activities) Income generation
	Economic value generation by the project	Cost-benefit and return on investments, new market opportunities, new business models, competitiveness of the project
	Impact on ICT driven innovation	Product innovation (e.g. increasing the efficiency of existing technologies, technology readiness level of the outputs) Process innovation, organizational innovation, user-driven and open innovation
Political impacts	Impact on civic and political participation	Instruments developed by the project offering new channels for civic participation Capacity of the project for increasing citizen participation in civic-society Increase in number of grassroots organisations Participation of citizens in (signature) campaigns and boycotts
	Impact on citizens/user's political awareness	Time spent by users to be informed about political issues Time spent by users in persuading others about the political issues Changes in the political topics addressed by users
	Impact on policies and institutions	Project's capability to influence policies and institutions, CAPS users impact on policies and institutions

Table 14: IA4Si methodological framework: socio-economic impact assessment (Passani et al., 2016, Chapter 3)

Each of the above domains are relevant for the hackAIR project, however, seen the scope areas of the hackAIR it was opted to only focus on the social and environmental impact domain. The chosen variables are discussed in chapter 3.2.1.

3.2 Impact assessment methodology for hackAIR

The impact assessment methodology for hackAIR includes several stages whereby indicators will be developed and refined. Firstly, the IA4SI impact assessment methodology is used as inspiration for understanding impact assessments, before a first set of impact indicators are distilled. As outlined in the DoW, these indicators will allow an evaluation of both social and environmental indicators. Figure 10 outlines the various stages of the methodology, including the methods employed, actors targeted and estimated timing.

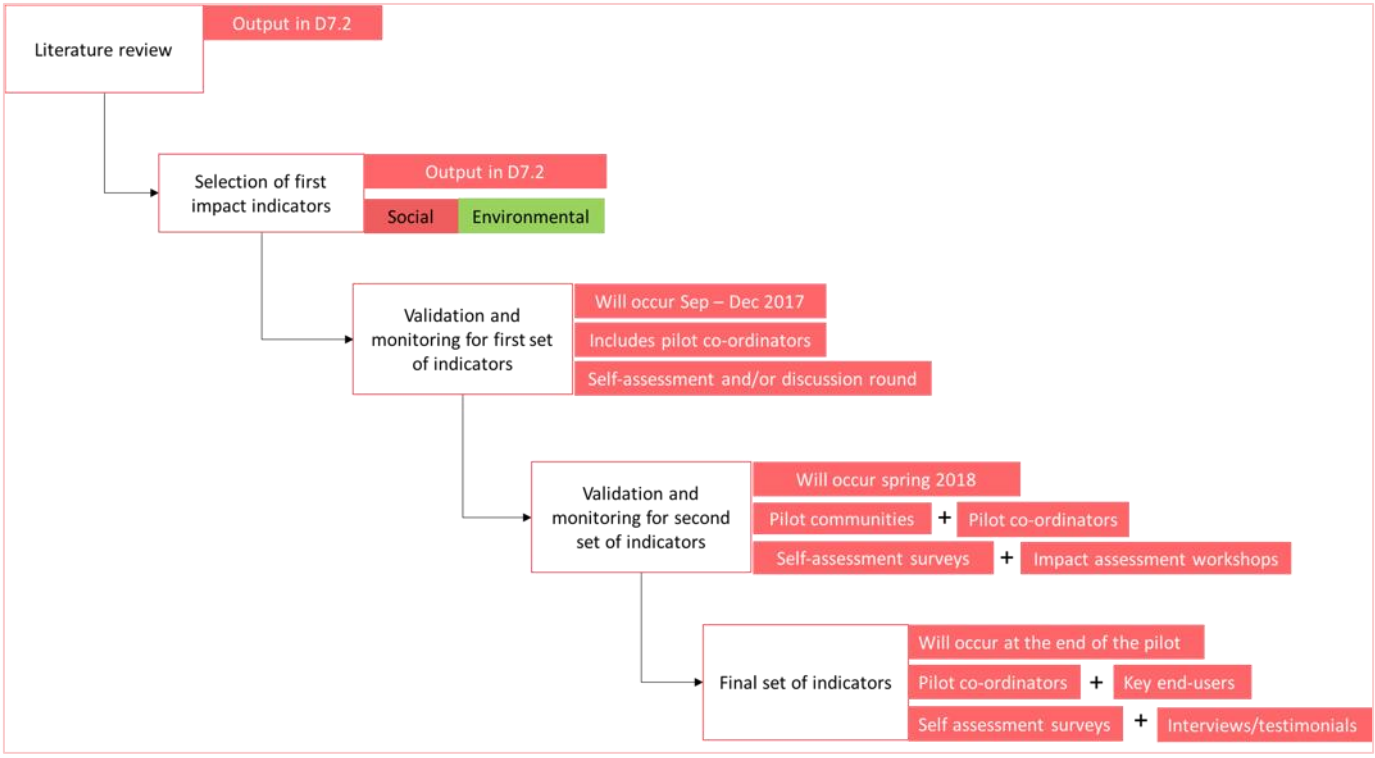

Figure 10: hackAIR impact assessment methodology.

A first set of indicators will be adopted from the IA4SI methodology, and described in sub-chapter 3.2.1. This represents a first output in this deliverable for iteration. Secondly, it is important that indicators are expanded and validated internally, to determine their applicability in the context of hackAIR. Through self-assessment rounds – where pilot partners are able to reflect on the indicators and provide systematic feedback to iterate – End 2017 and/or discussion rounds, a provisional first set of indicators will be agreed upon. Self-assessment can be collected via an online or collective survey.

In Spring 2018, a second monitoring and validation round will take place. This time, pilot communities will have a central focus by critically considering the way that we measure impact in the project. It is envisioned that feedback will be collected via workshops with hackAIR users in both pilot communities. This setting can be combined with the on-going usability evaluation of the hackAIR platform. As such, it is expected that the refinement of indicators takes place in M25/26. It is important in this stage that partners seriously consider the input from citizens, ensuring that the participatory methods used are most effective. This methodology is designed to allow for new indicators to emerge from multiple different avenues; it might be the case that hackAIR users prioritize certain features or indicators more or less than initially expected.

At the end of the project, a final evaluation of indicators will take place. In a similar fashion to the previous process in spring of 2018, it is expected that both pilot co-ordinators and users are able to provide their experiences and perspectives of hackAIR. Pilot input can be collected via self-assessment surveys. During this round, policy-makers, teachers, members of health organizations or environmental members can also be considered. User interviews and/or

testimonials will be employed as a way to generate a deep understanding of user perspectives in a novel way. It is in this round that broader users and stakeholder groups are encouraged to provide their perspectives. In doing so, the final stage of the hackAIR impact assessment will adopt storytelling and narratives, in conjunction with systematic self-assessment feedback to communicate and assess the impacts of the hackAIR project.

3.2.1 List of indicators

As mentioned above, this section will outline the first set of chosen impact indicators for the hackAIR project. In annex 2, each impact indicator has been operationalized for use in a self-assessment survey.

Impact area	Sub-category	Variables	hackAIR expected impact
Social impact	Impact on community building and empowerment	Online community building	Number of users engaged on the platform (beginning/end) Amount of time spent on the platform (beginning/end) Degree of communication on the platform (beginning/end)
		Local community building	Capacity to enlarge the local community in Germany and Norway
		Impact on social innovation CAPS community	Number of collaborations with other projects within the domain, and outside the domain
	Impact on information	Access to and sharing of information	Amount of available information about AQ on the platform Perceived improvement of access of info for users about AQ through hackAIR hackAIR's capability to influence information asymmetries about AQ between aware and non-aware citizens
	Impact on ways of thinking/opinions and changes in behaviour for individual and collective behaviour and lifestyles	Changes in ways of thinking	Number of citizens participating in the engagement activities of hackAIR Increased level of awareness Change in opinion about AQ
	Impact on education and human capital	Training provided by the project	Number of workshops provided Number of different topics covered by training activities Tools for education developed by the project about AQ
	Impact on science and academia	Knowledge sharing	Use of open access Sharing through social media Dissemination through project website Sharing through events
		Citizen science	Proportion of open science contributions Number of local measurement initiatives developed Consortium attitudes towards citizen science

			Number of activities for citizens to contribute to data Number of activities where citizens can analyse and interpret data[2]
Environmental impact	Project's capacity to impact individual and collective behaviour	The impact on user's environmental behaviour	Number of compensation activities performed by users since their engagement with the project Project's capability to contribute to the change in users participated to environmental related actions towards AS Project's capacity to increase users' sensitivity towards the issue of air pollution

Table 15: Set I: impact indicators of the hackAIR project.

3.3 Evaluation tools

3.3.1 Self-assessment surveys or discussion rounds

Self-assessment surveys represent a data gathering technique whereby participants are required to internally assess and approximate the expected or current impacts of a proposed action. As such, self-assessments can take place at organizational (i.e. reflecting on internal consortium capacities), or output levels (i.e. estimating expected outputs or benefits). In the case of hackAIR, self-assessment surveys will be distributed to pilot co-ordinators and gather data at the level of the output.

Self-assessment surveys follow a similar structure to standardized surveys (see 2.2.2), and carry similar advantages of scale, price, speed and analytical rigour. Furthermore, they accommodate data capture and analysis using integrated software such as qualtrics, or via external analysis software as SPSS. In order to capture project perspectives, both open and closed question will be used. Closed questions or standardized scales allow for aggregated analysis or relational comparisons, whereas open questions will target more descriptive elements or insights. Given the size of the hackAIR consortium, it is unlikely that statistically significant results will be generated via self-assessment surveys. Having said that, any results can be used to establish consensus or complement descriptive feedback. Annex X below discloses a template pilot survey, to illustrate the types of questions that will be posed during data collection rounds.

As an extension of survey results, group-based discussion rounds are particularly effective in collectively exploring the contributions of hackAIR as a consortium.
As mentioned above, self-assessment surveys will be used on three different occasions: in September – December 2017, in Spring 2018 and at the project end. Discussion rounds are estimated to take place once (in Sep-Dec 2017), however this could possibly change depending on the outcomes of each data collection round. As an initial step, one discussion round is planned for hackAIR partners and co-ordinators. Having said that, it is possible that this method is used as a way to validate impact indicators with other stakeholders.

3.3.2 Impact assessment workshop

Impact assessment workshops are evaluation activities that are effective in defining and develop lists of impact indicators. In this context, the aim of workshops is to monitor and validate the variables identified by the hackAIR team. Furthermore, impact workshops allow for the incorporation of perspectives from hackAIR users that are not directly connected to the consortium. Often accompanying a format similar to the world café and other variants, their

[2] Citizen science indicators adapted from OECD (2016): A framework to monitor open science trends in the EU. Source: www.oecd.org/sti/063%20-%20OECD%20Blue%20Sky%202016_Open%20Science.pdf

strength lies in structuring and integrating the perspective of multiple actors in a participatory environment. The assumption of such workshops is that by encouraging an informal environment, participants will be more likely to connect with each other, as well as with hackAIR partners. The setting follows a thematic sequence, with participants progressing from tables of their choice after a set period of time. In this way, it would be possible to split up and connect social or environmental impact indicators across respective tables. A facilitator should be present at each table, with the role of guiding conversations and drawing conclusions at the end of each session.
A world café and other variants are commonplace in organizational and reconciliation settings, and have also gained traction in transdisciplinary sustainability research. By using the outputs of the workshops, it is possible to better define its indicators also by developing a wider and more focused list. Impact assessment workshops (through world café's or other formats) are expected to take place during the spring of 2018 in pilot locations, and can be combined with usability and user experience activities due for the same time.

3.3.3 Interviews

A method similar to focus groups, the interview allows us to collect deep understandings of phenomena, meaningful themes, practices, and relationships from an interviewees' own perspective. In the context of hackAIR, it is expected that interviews will present a worthwhile opportunity to learn about both foreseen and unforeseen impacts and bottlenecks in a way that is connected to pilot locations, as well the project as a whole.
The interview design will follow standard practices that accompany qualitative research. In terms of application, interviews will last approximately one hour, be audio recorded and transcribed. All interviews will be semi-structured in nature and will be conducted following an interview guide. This approach allows for the method to remain flexible by using a set of core questions, whilst at the same time approaching broad themes in a loose format. Questions will follow a similar structure and validation procedure as the other methods used in this project. In the context of data analysis, collected data will be systematically coded and thematically analysed.
hackAIR interviews will take place in the final stages of the project (phase 5, 2018 end – evaluation and integration), and predominantly target users of the hackAIR platform. Given the pilot specific nature of hackAIR and languages in Germany in Norway, pilot leads can choose to organise interviews as a way to obtain specific information. Beyond this, VUB will further explore the option of interviewing other actors of hackAIR. Experts, health organizations or policy makers will most likely hold a high interest in the context and domain specific impacts of hackAIR.

3.4 Impact assessment Planning and milestones

In the following table an overview is given of three pilot implementation periods and the impact assessment activities that will be set up:

	Evaluation tool	User group	Evaluation measures	Output
M21-M24	Self-assessment surveys Discussion rounds	Consortium partners & co-ordinators	Social impact Environmental impact	Provisional indicators set 1
M25-M28	Self-assessment surveys Impact assessment workshop	Pilot communities Pilot-co-ordinators	Social impact Environmental impact	Indicators set 2
M29-M34	Self-assessment surveys Interviews & testimonials	Pilot co-ordinators Pilot communities Key actors	Social impact Environmental impact	Impact assessment using final set of indicators

M36	D7.7 “Pilot performance and impact assessment”

Table 16:hackAIR planning for impact assessment

4 Evaluation summary for hackAIR

Table 15 lists all selected indicators from the three evaluation tracks. These include: 1) Usability, UX, effectiveness and acceptance indicators, 2) pilot indicators, 3) technical performance indicators and 4) impact indicators (social and environmental indicators).

Usability, UX and acceptance indicators	Pilot indicators	Impact indicators
Perceived ease of use (PEO)	Total number of users on different points in time (unique and returning – evolution over time) (divided by location and age)	Online community building
Perceived usefulness (PU)	Amount of participants to workshops (divided by gender, location, type of workshop)	Local community building
Intention (I)	Amount of participants to the measurement campaigns	Impact on social innovation CAPS community
Attitude (A)	Amount of implemented feature/bug requests of the participants or pilot coordinators released during de trial periods	Access to and sharing of information
Perceived enjoyment (PE)	Amount of interventions (communications, explanations, …) done by hackAIR experts to the citizens on their gathered data.	Changes in ways of thinking
User profile (UI)	Amount of hardware sensors assembled and installed by participants	Training provided by the project
User satisfaction		Knowledge sharing
User performance		Citizen science
		The impact on user’s environmental behaviour
Technical performance indicators		
Number of seconds between upload of picture to platform and visualisation of picture on the platform/app.		
Availability of the web platform pilot trial periods (uptime)		
Load time of fusion map (secs)		
Upload rate of AQ information (mins)		
Page speed load (only source code)		
Full page load time (average)		
Page speed load (only ping) with 150 concurrent requests		
Crash free sessions		

Non-fatal issues / Total sessions
Update frequency of the data fusion map
Averaging period of the data fusion map
Uptime of the data fusion map
Average number (during pilot execution) of image-based AQ estimations in the pilot countries (number of fresh measurements obtained in the last 24 hours)
Average number (during pilot execution) of AQ measurements (images + ground stations/sensors) in the pilot countries. (number of fresh measurements obtained in the last 24 hours)
number of discovered social media accounts relevant for communication strategy
number of "insight" sessions between communication managers and social media monitoring tool managers
Precision of the sky detection module (% of image-based AQ estimations coming from sky-depicting images)
Execution time should be comparable to that of State-of-the-Art reasoners on a single task (time benchmark)

Table 17:hackAIR complete set of evaluation indicators

The following Gantt tracks the parallel evaluation and assessment tracks across the duration of the hackAIR project. In addition, D7.4 and D7.7 are included in order to highlight where findings will be distilled for input into periodic and end-evaluation reports. As mentioned in chapter 2, usability, UX and acceptance Periods 1, 2 and 3 will integrate all mentioned data collection tools. Discussion rounds, self-assessment surveys ad interviews/testimonials have been split in Figure 11 due to their different timings.

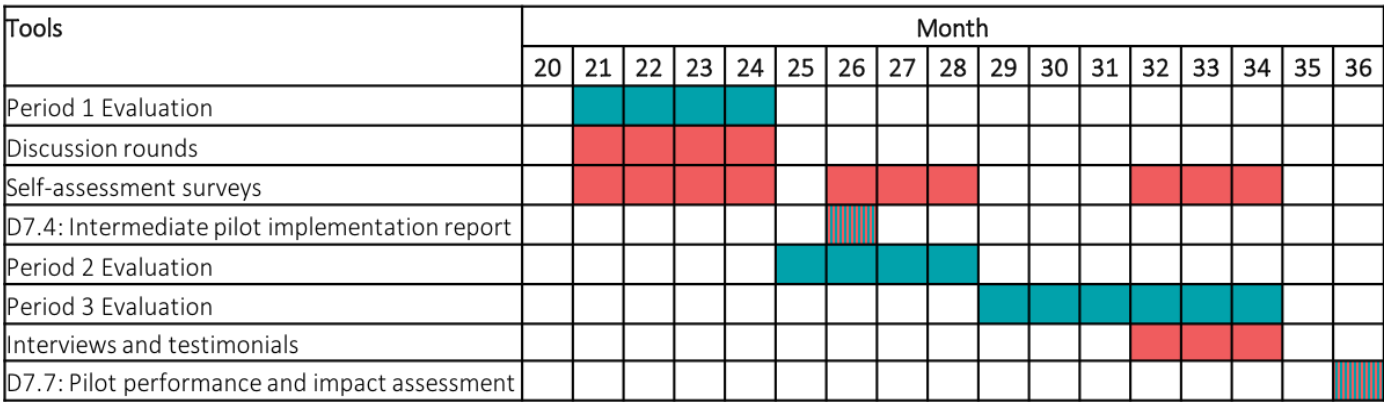

Figure 11: Gantt chart showing hackAIR overall evaluation timeline. Green = Track 1, Red = Track2/3.

5 Conclusion

The objective of this deliverable was to outlines the evaluation framework of the hackAIR project that has the objective to **evaluate (1) the usability and effectiveness of the hackAIR platform, (2) the pilot, (3) the technical components, and (4) the established impact of the platform at the social and environmental level in the communities.**

Chapter 2 was dedicated to a literature review and subsequent selection of usability, UX and acceptance indicators. Key technical indicators and pilot indicators were included to fulfil the objectives set out in evaluation track 1. Chapter 3 comprised a brief overview of impact assessment methodologies, before zooming in closer on the I4ASI framework. This is a hybrid methodology devised to assess project impacts using participatory processes. Secondly, the chapter identified social and environmental Impact indicators of relevance for the fulfilment of evaluation tracks 2 and 3. In both chapters, the evaluation tools, timings and milestones were outlined. Chapter 4 combines the key times and indicators from other chapters, charting the next steps in the hackAIR evaluation methodology.

The next steps are the effective implementation of the assessment framework within the pilots locations and in the technological framework to be able to measure and analyse all the indicators. The results of the evaluation framework will be reported in D7.4 (intermediate pilot implementation and evaluation report) and D7.7 (Pilot implementation and final evaluation report: pilot performance and impact of hackAIR).

6 References

Albert, W., & Tullis, T. (2013). *Measuring the user experience: collecting, analyzing, and presenting usability metrics*. Newnes.

Bund, E., Hubrich, D.-K., Schmitz, B., Mildenberger, G., & Krlev, G. (2013). *Blueprint of social innovation metrics: contributions to an understanding of opportunities and challenges of social innovation measurement* (Deliverable of the project "The theoretical, empirical and policy foundation for building social innovation in Europe (Tepsie").

Centre for Good Governance. (2006). A Comprehensive Guide for Social Impact Assessment. Retrieved September 8, 2017, from https://www.scribd.com/document/77519948/A-Comprehensive-Guide-for-Social-Impact-Assessment

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 319–340.

Ebrahim, A. S., & Rangan, V. K. (2010). The limits of nonprofit impact: A contingency framework for measuring social performance. *Harvard Business School General Management Unit Working Paper*, (10-099), 10–099.

European Commission. (2009). Impact assessment guidelines SEC (2009) 92. Retrieved September 8, 2017, from http://ec.europa.eu/smart-regulation/impact/commission_guidelines/docs/iag_2009_en.pdf

Joyce, S. A., & MacFarlane, M. (2001). Social impact assessment in the mining industry: current situation and future directions. *London: International Institute for Environment and Development (IIED)-Mining, Minerals and Sustainable Development*, 8–10.

Passani, A., Bellini, F., Prampolini, A., Vanobberghen, W., Firus, K., & Dulskaia, I. (2016). *IA4SI Methodological framework – Final version. Deliverable 2.2 of the project "IA4SI – Impact assessment for Social Innovation."*

Phills, J. A., Deiglmeier, K., & Miller, D. T. (2008). Rediscovering social innovation. *Stanford Social Innovation Review*, *6*(4), 34–43.

Vanclay, F. (2002). Conceptualising social impacts, 22, 183–211.

Vanclay, F., (2003). International Principles for Impact Assessment, 5517(July). http://doi.org/10.3152/147154603781766491

Vanclay, F., Esteves, A. M., Aucamp, I., & Franks, D. (2015). *Guidance for assessing and managing the social impacts of projects*. International Association for Impact Assessment. Retrieved from https://www.iaia.org/uploads/pdf/SIA_Guidance_Document_IAIA.pdf

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 425–478.

Wixom, B. H., & Todd, P. A. (2005). A theoretical integration of user satisfaction and technology acceptance. *Information Systems Research*, *16*(1), 85–102.

7 Annexes

7.1 Annex 1: Usability, UX & acceptance template for evaluating the hackAIR solution

To conclude a testing a survey will be handed over to the participants (offline or online). We will ask the participants their opinion about the platform through a questionnaire that will take approximately 3 minutes. The survey could exist out of the full set of questions, but will most likely exist of a selection of the questions that are presented in this annex.

1.Questionnaire to measure PERCEIVED EASE OF USE (based on Davis, 1989)

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
Learning to work with the hackAIR application would be easy for me	O	O	O	O	O
I would find it easy to get the hackAIR application to do what I want it to do	O	O	O	O	O
It would be easy for me to become skilful at using the hackAIR application	O	O	O	O	O
I would find the hackAIR application easy to use	O	O	O	O	O

2. Another possibility to measure the PERCEIVED EASE OF USE and the INTENTION OF USE at once is to use the System Usability Scale (SUS) - a simple, ten-item scale giving a global view of subjective assessments of usability.

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
1. I think that I would like to use this system frequently	O	O	O	O	O
2. I found the system unnecessarily complex	O	O	O	O	O
3. I thought the system was easy to use	O	O	O	O	O
4. I think that I would need the support of a technical person to be able to use this system	O	O	O	O	O
5. I found the various functions in this system to be integrated	O	O	O	O	O
6. I thought there was too much inconsistency in this system	O	O	O	O	O
7. I would imagine that most people would learn to use this system very quickly	O	O	O	O	O
8. I found the system very cumbersome to use	O	O	O	O	O
9. I felt very confident using the system	O	O	O	O	O
10. I needed to learn a lot of things before I could get going with this system	O	O	O	O	O

3. Questions to measure PERCEIVED USEFULLNESS (based on Davis, 1989)

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
Using the hackAIR application would make it easier to be informed about the quality of the air	O	O	O	O	O
Using the hackAIR application would make it easier to learn about the quality of the air	O	O	O	O	O
Using the hackAIR application would make it easier to change my behaviour related to the quality of the air	O	O	O	O	O

4. Question to measure PERCEIVED ENJOYMENT

To what extent do you agree with the following statements:

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
I find using the hackAIR application to be enjoyable	O	O	O	O	O
The actual process of using the hackAIR application is pleasant	O	O	O	O	O
I have fun using the hackAIR applications	O	O	O	O	O

5. Questions to measure INTENTION

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
I think that I would like to use the hackAIR solution frequently	O	O	O	O	O
I intend to use the hackAIR solution to check information about air quality	O	O	O	O	O
I will continue using the hackAIR solution after the testing period	O	O	O	O	O

6. Question to measure ATTITUDE

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
I think that I would recommend the hackAIR solution to others	O	O	O	O	O
My attitude towards using the hackAIR solution is very favourable	O	O	O	O	O
Overall, using the hackAIR solution is a pleasant experience	O	O	O	O	O

7. Questions to measure the USER PROFILE

What is your age category?	○ Below 15 years old (1) ○ 15 - 20 (2) ○ 21 - 30 (3) ○ 31 - 40 (4) ○ 41 - 50 (5) ○ 51 - 60 (6) ○ 61 - 70 (7) ○ Above 70 years old (8)
What is your gender?	○ Female ○ Male ○ Other

8. Questions to measure USER SATISFACTION

Please give a score on 1 (not satisfied) to 10 (fully satisfied) about your experience with the system:

Not satisfied									Fully satisfied
	1	2	3	4	5	6	7	8	9	10
Overall, how satisfied are you with the hackAIR system?	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
How satisfied are you with the registration procedure (creation of account)?	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
How satisfied are you with the application download experience from the Appstore/Playstore?	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
How satisfied are you with the support of the system (helpdesk, manual, Q&A)?	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
How satisfied are you with the game aspects of the system?	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍

9. Questions to measure USER PERFORMANCE / TASK SUCCESS

Please indicate to what extent you were able to complete each testing task:

	I was able to successfully complete the task	Partial success	I was not able to complete the task	I wasn't asked to execute the task
Task 1: Log in to the hackAIR application	❍	❍	❍	❍
Task 2: Take and upload a picture to the hackAIR application	❍	❍	❍	❍
Task 3: Create my own sensor (type sensor: ……)	❍	❍	❍	❍
Task 4: See your own contribution (sensor data, picture) on the hackAIR map	❍	❍	❍	❍
Task 5: Look up and understand the air quality level in your location	❍	❍	❍	❍
Task 6: Communicate in the hackAIR application with one or more other participants	❍	❍	❍	❍

7.2 Annex 2: think-a-loud

Example of a possible story in the think-a-loud protocol is the following:
- Navigate to the hackAIR platform
  - Register and log in on the hackAIR platform
  - Explore and search for air pollution levels in your local area
  - Take and upload one air pollution measurement using photograph

-Write down notes, and conduct after-scenario assessment – record usability issues, record user satisfaction, record 1) bottlenecks for task completion, and record errors. Can be validated via 'dry-run' with consortium

Observation parameters and task completion status:

Task 1: Register and log in on the hackAIR platform				Task completion:	Incomplete/complete	
Frustrated	Smiling/satisfied	Surprised	Excited	Concentrated	Impatient	Neutral
Interesting comments/suggestions expressed by the participants/ unexpected events:						
1. 2. 3.						

Task 2: Explore and search for air pollution levels in your local area				Task completion:	Incomplete/complete	
Frustrated	Smiling/satisfied	Surprised	Excited	Concentrated	Impatient	Neutral
Interesting comments/suggestions expressed by the participants/ unexpected events:						
1. 2. 3.						

Task 3: Contribute with one measurement in your local area through a photograph				Task completion:	Incomplete/complete	
Frustrated	Smiling/satisfied	Surprised	Excited	Concentrated	Impatient	Neutral
Interesting comments/suggestions expressed by the participants/ unexpected events:						
1. 2. 3.						

In-depth discussion:
- How satisfied are you overall about the hackAIR platform and mobile application?
- Which aspects can be further improved in terms of ease of use?
- Which aspects did you find useful?
- Did you enjoy using the hackAIR platform and mobile application?
- Do you think you will use the hackAIR platform and mobile application again?

7.3 Annex 3: Pilot performance indicators

Description	KPI	Measurement channel
Total number of users on different points in time (unique and returning – evolution over time) (divided by location and age)	Sept/Oct 2017: 100 Dec 17/March 18: 1000 April/July 2018: 10000 Aug/Dec 2018: 10000	Logfile
Amount of participants to workshops (divided by gender, location, type of workshop)	At least 7 workshops with in total at least 50 participants	Counting of participants by workshop coordinators & survey after/before workshop
Amount of participants to the measurement campaigns	At least 5 measurement campaigns with in total at least 30 participants	Counting of participants by pilot coordinators
Amount of implemented feature/bug requests of the participants or pilot coordinators released during de trial periods	70%	Release notes
Amount of interventions (communications, explanations, …) done by hackAIR experts to the citizens on their gathered data.	At least 25	Count of helpdesk interventions + face to face interventions
Amount of hardware sensors assembled and installed by participants	100	Logfile

7.4 Annex 4: Technical performance indicators

Objective description	Measurement process/unit	Indicator
Platform performance	Number of seconds between upload of picture to platform and visualisation of picture on the platform/app.	< 20 sec
Platform performance	Availability of the web platform pilot trial periods (uptime)	90%
Platform performance	Load time of fusion map (secs)	5 secs
Platform performance	Upload rate of AQ information (mins)	<10 mins
Platform performance	Page speed load (only source code)	<5s
Platform performance	Full page load time (average)	< 15s
Platform performance	Page speed load (only ping) with 150 concurrent requests	< 15s
Mobile App performance	Crash free sessions	95%
Mobile App performance	Non-fatal issues / Total sessions	< 1
Platform performance	Update frequency of the data fusion map	< 3 h

Platform performance	Averaging period of the data fusion map	<= 24 h
Platform performance	Uptime of the data fusion map	> 95%
Environmental node discovery module	Average number (during pilot execution) of image-based AQ estimations in the pilot countries (number of fresh measurements obtained in the last 24 hours)	> 50
Environmental node discovery module	Average number (during pilot execution) of AQ measurements (images + ground stations/sensors) in the pilot countries. (number of fresh measurements obtained in the last 24 hours)	> 100
Social media monitoring tool	number of discovered social media accounts relevant for communication strategy	> 500
Social media monitoring tool	number of "insight" sessions between communication managers and social media monitoring tool managers	1 per quarter
Image analysis module (i.e. sky detection)	Precision of the sky detection module (% of image-based AQ estimations coming from sky-depicting images)	
Ontology and Reasoning component	Execution time should be comparable to that of State-of-the-Art reasoners on a single task (time benchmark)	yes

7.5 Annex 5: Impact indicators list

Sub-category	Variables	Indicator questions
Impact on community building and empowerment	Online community building	To what extent to you feel users have adequately engaged during the hackAIR project? (1 = Very little extent- 5 = a great extent)
		How much time have users spent on the hackAIR platform
		What is your opinion of the degree of communication on the platform
		How has hackAIR helped you expand your local community?
Impact on information	Access to and sharing of information	What is the amount of available information about AQ on the platform?
		How have users experienced an improvement of access of info about AQ through hackAIR?
		What is your opinion on hackAIR's ability to influence information asymmetries about AQ between aware and non-aware citizens?
Impact on ways of thinking/opinions and	Changes in ways of thinking	How many citizens have participated in the engagement activities of hackAIR?

changes in behaviour for individual and collective behaviour and lifestyles		How has hackAIR changed the opinions and perspectives of citizens about topics related to air pollution?
Impact on education and human capital	**Training provided by the project**	Number of workshops provided? Number of different topics covered by training activities? How do you think hackAIR has influenced education and training about air quality?
Impact on science and academia	**Knowledge sharing**	To what extent do you think hackAIR has shared knowledge using principles of open science and citizen science? What proportion of publications from hackAIR have been published using open access journals? How has hackAIR shared project or phenomenon knowledge through social media? What impact do you think offline engagement activities has had on knowledge sharing?
Project's capacity to impact individual and collective behaviour	The impact on user's environmental behaviour	How many compensation activities have been performed by users since their engagement with the project?
		How would you assess hackAIR's ability to contribute to the change in users that have participated in air quality activities?
		How would you assess hackAIR's capacity to increase users' sensitivity towards the issue of air pollution?