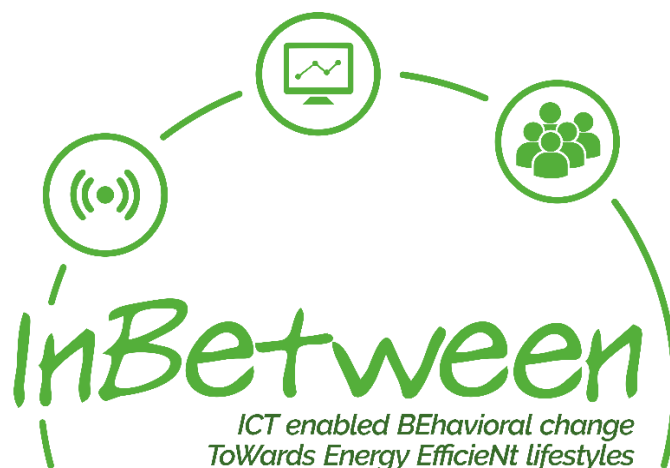




### D3.13– A SET OF USER ORIENTED EVALUATION CRITERIA FOR TECHNOLOGY EFFECTIVENESS

Due date of deliverable: February 2020

Actual submission date: April 2020



**LEADER: IDC**

#### DISSEMINATION LEVEL

PU	Public	X
CO	Confidential, only for members of the consortium (including the Commission Services)	

#### HISTORY

Version	Description	Lead author	Date
V0	First draft	IDC	16/02/2020
V1	After reviewers' comments	IDC, SON, VIL	21/02/2020
V2	Additional review	IDC, RINA	26/02/2020
V3	Additional review	RINA	04/03/2020
Final	Final version	RINA	06/04/2020

#### DISCLAIMER

The work presented in this document has been conducted in the context of the H2020 of the European community project InBetween (n° 768776). The partners in the project are: Rina Consulting S.p.A., Acciona Construcción S.A., AIT Austrian Institute of Technology GmbH, Develco Products, The Interdisciplinary Center Herzliya, Institute Mihajlo Pupin, Vilogia S.A, Sonnenplatz Großschönau GmbH. The content of this report does not reflect the official opinion of the European Union. Responsibility for the information and views expressed in the therein lies entirely with the author(s).

## EXECUTIVE SUMMARY

The InBetween project applies user-centric approach, where throughout the project we have developed and applied different methodologies to collect qualitative and quantitative data about the users' views and expectations from the platform. In this report we synthesize previous reports (D3.7, D3.9, D3.11), and propose guidelines, methodologies and tools for user-oriented evaluation criteria. Full guidelines will be presented in D3.14.

## TABLE OF CONTENTS

<b>1</b>	<b><i>Introduction</i></b> .....	<b>3</b>
<b>2</b>	<b><i>A methodology for understanding users' view on the app usefulness</i></b> .....	<b>4</b>
2.1	Semi-structured interviews.....	5
2.2	Workshop and survey.....	5
<b>3</b>	<b><i>A user-centric methodology for evaluating actions and behavior change</i></b> .....	<b>6</b>
3.1	KPIS DEFINITION PROCESS .....	6
3.2	USER ENGAGEMENT KPIS .....	8
3.3	ENERGY USAGE KPIS .....	9
3.4	SECURITY AND HEALTH KPI .....	9
3.5	BENCHMARK FOR COMPARISON.....	9
<b>4</b>	<b><i>A user-centric methodology for associating notifications and user behavior</i></b> .....	<b>10</b>
4.1	The notifications .....	10
4.2	User perception of the notification .....	11
4.3	Associating notifications and action .....	11
<b>5</b>	<b><i>Conclusions</i></b> .....	<b>12</b>
<b>1</b>	<b><i>Annex 1</i></b> .....	<b>12</b>

## LIST OF FIGURES

Figure 1: An overview of the user-centric approach	4
Figure 2 Methodology rationale	6

## LIST OF TABLES

Table 1: Summary of selected KPIs	7
Table 2: Notifications and text	11

## 1 INTRODUCTION

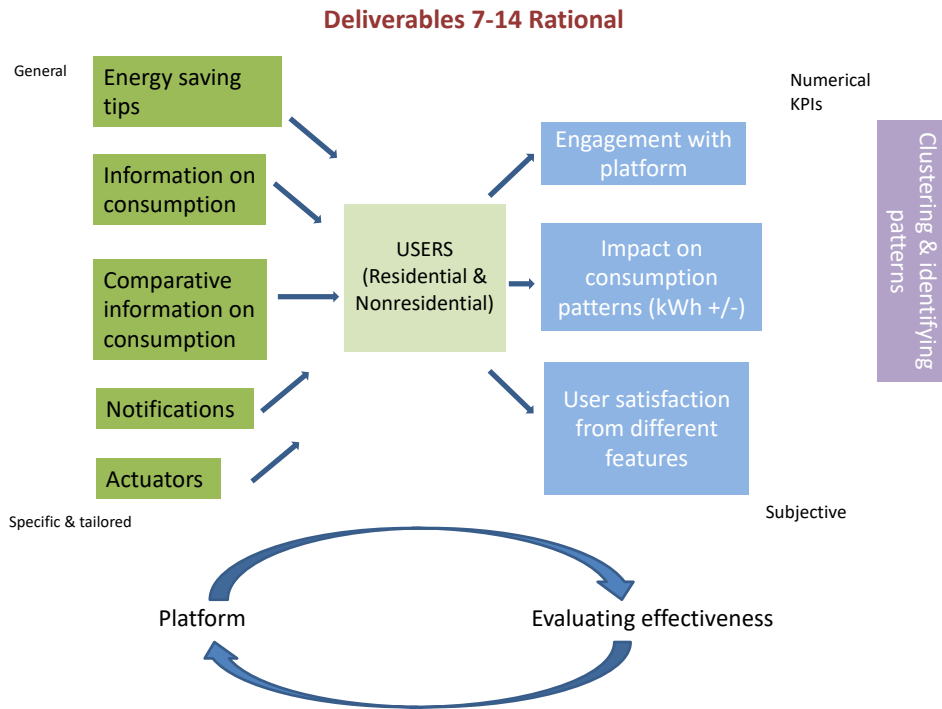
The InBetween project applies user-centric approach. This means that throughout the project we developed and applied (given the constraints) different methodologies to collect qualitative and quantitative data about the users' views and expectations from the platform. A comprehensive set of relevant KPIs was identified in D1.5 using the available literature. Based on the information available at the time, an assessment was made which KPIs would be calculable from the expected data, and which were not feasible due to lack of relevant monitoring. Previous reports (D3.1-D3.6) describe in detail how this approach was apparent in the platform and app's initial development stages, the installation stage and the pre-use stage (baseline period).

The post-installation reports (D3.7, D3.9, D3.11) describe the detailed and multi-dimensional approach we applied and aim to apply for the examination and evaluation of the platform effectiveness. D3.7 focuses on how we approached users and identified gaps between their expectations from the technology and satisfaction. Some of the comments and suggestions have already been incorporated into the second release of the app. Other comments will be fed back to the app developers and if possible be incorporated to the app design. D3.9 focuses on the approach and methodology to identify practices that changed, and to evaluate the extent of change. D3.11 describes the approach and methodology we applied to associate between one type of intervention (from the several we applied) - the notifications - and behavior change.

It is important to note that for users who do not own or use Android phones, and for those who prefer a web interface over a mobile app, such interface was developed. The web platform contains some similar functionalities to the app in terms of engagement, yet a web interface is somewhat more limited and not as handy as a mobile phone. The methodologies described below apply to the interaction with the platform (app or web).

Figure 1 presents an overview of the evaluation approach and the criteria we apply to examine the effectiveness of the app.

In this report we synthesize previous reports, and based on our experience thus far propose guidelines, methodologies and tools for user-oriented evaluation criteria for the InBetween platform effectiveness. While developed for the inBetween project, these user-centric evaluation criteria could be implemented to evaluate other energy saving platforms that rely on user engagement. Full guidelines will be presented in D3.14.



**Figure 1: An overview of the user-centric approach**

## 2 A METHODOLOGY FOR UNDERSTANDING USERS' VIEW ON THE APP USEFULNESS

As described in D3.7, ideally, applying a user-centric approach implies that users have to be consulted throughout the development process about their views on the usefulness and relevance of the app. It also implies a co-design strategy, i.e., that insight from users are incorporated into the app design.

There are several common practices to involve users, including: in-depth structured and semi-structured interviews; focus groups and workshops; surveys; and built-in in-app feedback (e.g., thumb up/down). Each method and practice has its own strengths and weaknesses and provides a unique view and perspective on the product/service and its usefulness. For example, in-depth interviews are time-consuming and the sample usually does not represent the population. Yet interviews provide detailed information about each feature, as well as emphasizing specific user considerations that otherwise could be overlooked. Focus groups and workshops promote discussion and provide opportunities to highlight and elaborate in a group setting different views of the product. Surveys, on the other hand, provide data from large (and hopefully a representative) sample of users, thus provide a good overview of them– including the keen and not so-keen users. In-app feedback allows receiving very simple feedback on specific features at a specific time. The combination of methods is important to gain a rounded and comprehensive idea about user perception.

In the InBetween project, the app was first created by the project Consortium as a whole, with iterative exchanges between technical partners, “demo site owners” (Vilogia, hereafter VIL, and Sonnenplatz, hereafter SON), who are familiar with users' needs in each demosite and therefore were acting as their representatives,

and the energy behavior experts (IDC). The app was then released to the users (i.e. the residents at VIL and the residents and the users of the SON buildings) and based on demo-site owners' interactions and conversations with early app users, some changes were made to the app.

We planned to apply the following methods: several face-to-face semi-structured interviews with the first users of the app, workshop with users, a survey and built-in feedback.

At the time this report is being written, only six interviews took place. A stakeholder engagement workshop is planned for May 2020 and the built-in feedback (thumb up/down) will be implemented soon. We also plan to issue a survey.

## 2.1 Semi-structured interviews

To maximize the benefits of semi-structured interviews, it is important that the questionnaire covers all the aspects of usage as well as all the services and other features offered. In the introduction to the interview, we highlighted that (a) the interview is part of the co-design and user-centric approach of the project, (b) the tool is in its development stage and we are consulting them in order to understand their views, (c) the input they provide will be fed back to developers and used to improve the tool. And finally, we stressed that we need their honest views, and there is no 'right' or 'wrong' questions. We also explained that the information they provide will be used only for these purposes.

The full Questionnaire can be found in Annex 1. For full details on the results see D3.7.

## 2.2 WORKSHOP AND SURVEY

The project builds on the conceptual model of users' agency and capacity as drivers for behavior change: As discussed in previous reports (D3.2, D3.4) the likelihood of an energy user to take action (in our case change behavior) is associated with the user's level of agency and capacity, where 'agency' refers to consumers' willingness and interest in making their own free choices regarding energy consumption and energy related behavior, and 'capacity' refers to users' ability to perform the choices they have made. When levels of both agency and capacity are high – change is more likely to happen.

In previous reports we identified reasons for low levels of agency and capacity. Accordingly, the platform is designed to improve agency and capacity in various ways, ranging from general advice to specific and personalized notifications and individual actuators. For extended explanation on users' agency and capacity and how these concepts are applied in the inBetween project see D3.4.

For consistency, in the workshop (planned for May 2020) and survey (time to be planned) we will use a shorter version of the questionnaire.

In the survey, the questions will be divided into two groups: agency-related and capacity-related (See D3.2, D3.4, D3.7). This, in turn, will allow us to examine:

- (1) if and how the different features of the platform increased users' level of agency and capacity,
- (2) to empirically examine the relations between agency, capacity and action.

### 3 A USER-CENTRIC METHODOLOGY FOR EVALUATING ACTIONS AND BEHAVIOR CHANGE

The goal of the InBetween platform and its different features is to trigger behavior change and reduce energy consumption, taking into consideration user constraints. Hence, we need to be able to correlate between the use of the app and changes in energy consumption (or other behavior, such as security and health).

To evaluate the practices that changed and the extent of change at least four conditions need to be fulfilled:

- (1) Users have been using the app / web interface for at least a month of two, and experienced its different features and options.
- (2) KPIs were developed to measure the usage of the app (first in D1.5)
- (3) KPIs for consumption and other parameters (such as air quality) were developed (first in D1.5)
- (4) Baselines for comparison were set to estimate behaviour changes.

D3.9 discusses in detail how we measure these aspects in the InBetween project.

Figure 2 presents the methodology we applied to connect between the use of the app and the behavioural change.

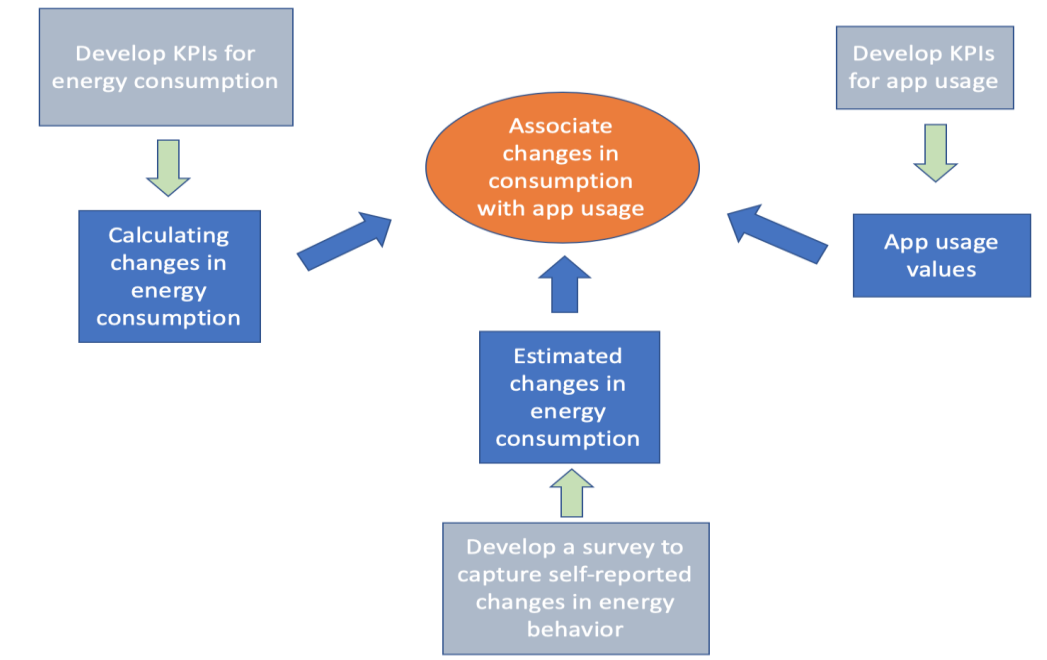


Figure 2 Methodology rationale

#### 3.1 KPIS DEFINITION PROCESS

In D1.5 a first list of KPIs have been defined. These have been further detailed and expanded as part of Task 3.3 , as seen in the Table below.

**Table 1: Summary of selected KPIs**

	<b>Goal</b>	<b>Indicator (from D1.5)</b>	<b>ID</b>
Energy use	Reduce amounts of energy used by occupants	Energy consumption by person – total	01_01
		Energy consumption by floor area – total	01_02
		Energy savings	01_11
	Assess and reduce energy consumption for each end use	Energy consumption by person – disaggregated: space heating, cooling, ventilation, lighting, hot water, others	Da 01_03 a 01_08
	Reduce impacts of energy consumed by the occupants	Source energy consumption	01_09
		CO2 emission <sup>1</sup>	01_10
		CO2 emission savings	01_12
	Reduce costs of energy for occupants	Energy cost savings	01_13
	Assess potential for heating energy savings	Energy consumption % of ideal demand	01_14
	Reduce stress for grid and minimize costs of electricity generation	Peak load indicator	01_15
Optimize self-consumption of renewable energy	Load match index	01_16	
Comfort	Provide thermal environment as comfortable as possible	Temperature discomfort indicator	02_01
	Limit number of occupied hours with uncomfortable thermal environment	% uncomfortable hours	02_02
	Minimize thermal dissatisfaction of occupants with variety of tasks	Thermal discomfort indicator	02_03
	Provide good air quality	Stale air indicator	02_05
		Volatile organic compound levels	02_06
User engagement	Engage users in a sustained way	Recency index	03_01
	Provide users with interesting information	Message opening rate	03_02

<sup>1</sup> Addition from Task 3.3

	Convince users to adopt energy-efficient behaviour	Compliance indicator	03_03
--	--	----------------------	-------

### 3.2 USER ENGAGEMENT KPIS

Due to some limitations, in this report we have been able to focus mainly on:

- Intensity of use,
- Activity,
- Endurability,
- Motivation to use the tool,
- System feedback,
- User-perceived utility of system,
- System accuracy.

According to what reported in Table 1, three KPIs were selected for user engagement with the app. More in details, recency index can be calculated by looking at time passed between individual active app usage sessions, message opening rate is percentage of opened notification (contrasted to ignored ones) and compliance indicator as ratio of followed vs. ignored recommendations. The last one could be a bit tricky to calculate as following the recommendation may be temporally apart from the actual usage.

We are using Google analytics to collect the abovementioned information from five activities (screens):

- Main Activity
- Monitoring Activity
- Login Activity
- About us Activity
- Privacy Policy Activity

For each activity (screen), we collect:

- Views: *The number of app screens or web pages users saw. Repeated views of a single screen or page are counted.*
- Users: *The total number of active users.*
- New users: *The number of users who interacted with the site or launched the app for the first time.*
- Screen views per user
- Engagement time: *The average length of time that the app was in the foreground, or the web site had focus in the browser*
- Event count: *The number of times users triggered an event. An event represent an action made by end-user which entails, but is not limited to, usage of any of the action buttons such as control switch for actuators, scheduling, optimization etc. It can also consider information about app opening, browsing through different activities, i.e. app screen, and information about received notifications and end-user actions, i.e. disable/dismiss.*



### 3.3 ENERGY USAGE KPIS

As reported in D1.5 – Key Performance Indicator for Platform Performance Assessment:

Summarizing the reviewed literature, multiple dimensions can be considered for the quantification of energy use and are discussed in the following paragraphs:

1. quantity of energy,
2. assessment level, system limits and energy boundaries,
3. end use,
4. quality of energy.
5. energy rate and load matching,
6. energy sources (e.g. gas, solar energy) and carriers (e.g. gas, electricity, district heating),
7. financial and environmental costs,
8. life cycle.

Based on the sensors installed, we calculated the following set of KPIs to evaluate energy usage and the use of other services offered by the app. The KPIs include indicators that consider not only the total consumption but also per capita and per floor area consumption, as total consumption does not accurately reflect the user specific characteristics. It is important to note, however, that not all these KPIs could be calculated in both of our demo sites (see D1.5 for details about the KPIs development and calculation and D3.9&D3.11 for details on implementation in demo-sites).

### 3.4 SECURITY AND HEALTH KPI

The sensors installed for energy monitoring and management purposes also enable the platform to provide additional useful information that will benefit users and prompt them to use the app for non-energy purposes, such as improving their security and health purposes. This, in turn, is likely to increase user engagement with the platform and increase the likelihood that they will be exposed to the energy information and options it offers. We have deployed sensors which provide information on temperature, air quality and occupancy. Hence, in addition to advising users that the heating is on and the temperature is higher than 24C, or that the heating is on while a window is open, the platform can also advise them that the house is vacant and a window is open, or that the air quality in a room is low, so that they might want to open a window and ventilate it. This is part of what we refer hereby as “security and health KPIs”.

During the KPIs definition (D1.5) some health-related indicators had been defined, but no security KPI. They have been included in a second stage.

### 3.5 BENCHMARK FOR COMPARISON

To evaluate change, we have also developed three different baselines for comparison. These three comparison options allow users to compare their own past and present consumption, as well as compare themselves to others. Hence, users can see how they improved (or not) compared to how they performed before, but also compare themselves to others, who do not necessarily share their unique characteristics:

- (1) data gathered during the baseline period. This benchmark is relevant to examine practices related to thermal comfort, as thermal comfort is associated with seasons.
- (2) data from previous month. This benchmark is relevant to examine practices that are not tightly linked to thermal comfort or seasons (e.g. use of washing machine, changes in water heater settings, use of actuators).

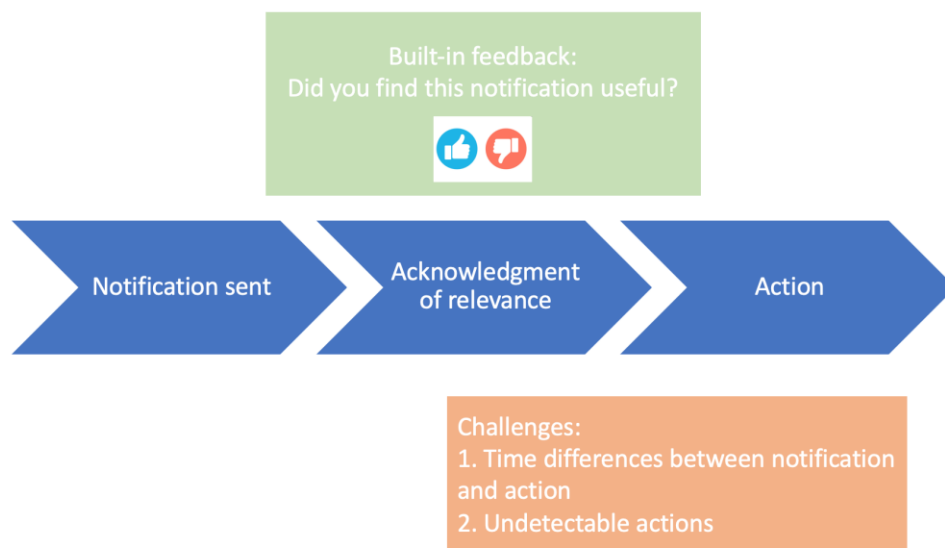
- (3) changes in the ranking compared to previous months and in comparison to other users. This reflects each users improvement (or not) compared to themselves and compared to others.

#### 4 A USER-CENTRIC METHODOLOGY FOR ASSOCIATING NOTIFICATIONS AND USER BEHAVIOR

The InBetween platform provides different types of notifications and advice, not all of them related to energy saving (e.g., security, health). Users have the option to tailor the services the app offers: to select the notifications they want to receive and to disable the ones which they find irrelevant or not helpful. The notifications are central to the InBetween project as they provide tailored and ‘in-time’ recommendations to very specific actions, and thus, are likely to increase both the agency and capacity of the users and lead to energy saving.

Figure 3 presents an overview of our approach for evaluation of the effectiveness of notifications.

Deliverable D3.11 presents in detail the methodology we apply to correlate between notifications and behavior (energy saving actions, as well as health and security-related actions), and discusses the various challenges of relating the two.



**Figure 3: Evaluating the effectiveness of tailored notifications**

##### 4.1 THE NOTIFICATIONS

To avoid an overload of notifications that could annoy users and deter them from using the app, and to learn about the dynamics of consumption vis-à-vis the notifications that we send, we decided to begin with sending limited types of notification (presented in table 1). The types and frequencies need to be re-evaluated later in the development stage. We also decided to use text that is less likely to be perceived as intimidating or invading users’ privacy (i.e., the app as ‘big brother’).

**Table 2: Notifications and text**

<b>Title</b>	<b>Body</b>
Energy Conservation Event	Windows are open while the heating is on. Room: %s Open windows: %s Running devices: %s
Severe Health Issue Alert	Unhealthy air quality has been detected with all windows closed! Consider ventilating if possible or vacating the premises temporarily. Room: %s
Energy Conservation Alert	Temperature in a room is over %s degrees while the heating is on. Consider turning the heating off. Room: %s Running devices: %s
Security Alert	Your home appears to be unoccupied while windows are open! Open windows: %s
Energy Conservation Alert	Your home appears to be unoccupied while the heating is on. Running devices: %s
Severe Health Issue Alert	Unhealthy air quality has been detected! Consider vacating the premises temporarily. Room: %s
Energy conservation and security alert	Your home is unoccupied while a window is open and the heating is turned on. Running devices: %s Open windows: %s

#### 4.2 USER PERCEPTION OF THE NOTIFICATION

Adhering to the user-centric approach, we would like to receive (as close as possible) real time feedback from users about the notification. For this, we are planning to add a built-in feedback feature that provide user accessible opportunity to indicate if they find a notification useful. To keep it simple, we decided to use the ‘thumbs up’ and ‘thumbs down’ symbols for users to reply.



While this is not an accurate or very informative form of feedback, its simplicity increases the likelihood that users will cooperate and give us a ‘quick and dirty’ indication about the usefulness of the notifications to users.

In addition, as described in the following section and in D3.11, some actions are undetectable or take place a long time after a notification was sent, thus this feedback is an additional method to learn about the effectiveness of the notifications.

Together with interviews, the workshop and survey will provide a comprehensive idea about the usefulness of the notifications as perceived by users themselves.

#### 4.3 ASSOCIATING NOTIFICATIONS AND ACTION

The consumption KPIs (described above) will be analysed using three benchmarks, and in association with the type and number of notifications that were sent to each user:

1. Previous year
2. Previous two weeks (or one month)
3. Changes in ranking compared to other users (benchmarking)

The changes in ranking will allow us to have a better idea about the contribution of the notification to energy saving and other actions.

## 5 CONCLUSIONS

The InBetween project applies user-centric approach. This means that throughout the project we developed and applied different methodologies to collect qualitative and quantitative data about the users' views and expectation from the platform / app (D3.1-D3.6), feedback on their experience with platform (D3.7), and KPIs on energy consumption and non-energy related behavior (D1.5). The input we collected during the project is incorporated into the platform.

The guidelines and tools we proposed in this report for comprehensively evaluating the technology effectiveness include:

1. Detailed questionnaire for in depth interviews. A shorter version of the questionnaire (presented above) should be used in the workshop / focus group and in the survey.
2. A survey that includes questions divided into two groups: agency-related and capacity-related. This allows the examinations of:
  - a. if and how the different features of the platform increased users' level of agency and capacity,
  - b. to empirically examine the relations between agency, capacity and action.
3. A set of KPIs for energy consumption.
4. A set of KPIs for non-energy aspect (health and security).
5. Analysis of notifications sent (by type and per user).
6. A set of KPIs for engagement and use of the app.
7. Three optional benchmarks empirically examine changes in consumption.

## 1 ANNEX 1

Below is the full version of the questionnaire:

1. *What interface (tool) do you use?*
  - a. *Web*
  - b. *App*
  - c. *Both*
2. *How long have you been using the tool?*
3. *How did you start using the tool?*
  - a. *Did you need help installing it and setting up?*
  - b. *Do you understand the tool and the different symbols?*
  - c. *If you need help / do not understand, what do you do?*

4. *How intense was your interaction with the tool so far?*
  - a. *open the app to check energy consumption?*
    - i. *How often?*
  - b. *received notifications?*
    - i. *Which ones (energy, temperature, security, health messages)?*
    - ii. *How often?*
  - c. *used actuators to manage/schedule energy use?*
    - i. *Which ones?*
    - ii. *Did they use the optimize option?*
    - iii. *How often?*
  - d. *Received general energy saving tips?*
    - i. *How often?*
  - e. *Check your consumption against other (benchmark)*
    - i. *How often?*
  - f. *Check weather forecast*
    - i. *How often?*
  - g. *Check specific appliances consumption*
    - i. *How often?*
  - h. *Check air quality information?*
    - i. *How often?*
  
5. *Which aspect of the tool you use the most? Can you rank each one (on a scale of 1-10)?*
  - a. *Energy*
  - b. *Weather*
  - c. *Security*
  - d. *Health*
  - e. *Benchmarking*
  
6. *How do you find the following features?*
  - a. *notifications*
    - i. *display clear and easy /unclear and hard to understand (please elaborate on each aspect – energy, security, health)*
    - ii. *content useful and helpful – why?*
    - iii. *content unclear /not useful / annoying – why?*
    - iv. *Intensity of notification – too much/ not enough*
    - v. *how could this feature be improved?*
  - b. *Actuator/energy management*
    - i. *display clear and easy /unclear and hard to understand (please elaborate)*
    - ii. *content useful and helpful – why?*
    - iii. *content unclear / not useful / annoying – why?*
    - iv. *how could this feature be improved?*

- c. *Consumption information*
  - i. *display clear and easy /unclear and hard to understand (please elaborate)*
  - ii. *content useful and helpful – why?*
  - iii. *content unclear / not useful / annoying – why?*
  - iv. *how could this feature be improved?*
- d. *Benchmarking information*
  - i. *display clear and easy /unclear and hard to understand (please elaborate)*
  - ii. *content useful and helpful – why?*
  - iii. *content unclear / not useful / annoying – why?*
  - iv. *how could this feature be improved?*

7. *Did the tool contributed to your energy / energy saving understanding and perception?*

- a. *In what ways?*
- b. *How could it be done better? (what is missing?)*

8. *Which behaviours and practices (if any) have you changed recently:*

- a. *thermal comfort related*
  - i. *which ones and in what way?*
  - ii. *has the tool contributed to this change (if yes – elaborate, is it related to notifications, to general information, etc.)*
- b. *energy management (schedule)*
  - i. *which ones and in what way?*
  - ii. *has the tool contributed to this change (if yes - elaborate, is it related to notifications, to general information, etc.)*
- c. *energy management (remote control via actuators)*
  - i. *which ones and in what way?*
  - ii. *has the tool contributed to this change (if yes - elaborate, is it related to notifications, to general information, etc.)*
- d. *daily routine related*
  - i. *which ones and in what ways?*
  - ii. *has the tool contributed to this change (if yes - elaborate, is it related to notifications, to general information, etc.)*

9. *Do you think that the tool, as it is, helps you to*

- a. *save money*
  - i. *if yes, how, which features contribute to this?*
  - ii. *Could you estimate how much money?*
  - iii. *if no – what is needed to help you save money?*
- b. *save energy*
  - i. *if yes, how, which features contribute to this?*
  - ii. *if no – what is needed to help them save money*

*c. reduce emissions*

- 10. currently, you do not receive notifications between 22:00-08:00 – would you like to receive notification later/ earlier as well?*
- 11. We would like to send a summary report once every 2 weeks / month – would you be interested? what sort of information would you like to receive?*
- 12. Do you have any other comments / suggestions / insights related to the tool?*
- 13. Would you be willing to pay for such tool?*
  - a. If yes – why?*
  - b. If no – why?*

*We would like to ask you now a few questions regarding the sensors and their installation:*

- 14. What did you feel during the installation of the sensors?*
- 15. Do you have feedback specifically about the sensors only? Any event to report? Are you satisfied with them?*
- 16. Do you remember what is which sensor?*
- 17. Are you encountering any problems with them?*

*Any other comments / suggestions / insights that you think we did not cover?*