Ref. Ares(2021)6865744 - 09/11/2021

DELIVERABLE 5.5
Validation Results including Lessons Learned and societal impact

adelphi

Revision: October, 2021

www.eco-bot.eu

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 767625

D5.5: Validation Results including Lessons Learned and societal impact Summary

D5.5 reports on the validation results from the three pilots SEnerCon, EYPESA and DEXMA (April 2020 to May 2021). For this validation, D3.3 serves as a basis, as it defines the Eco-Bot evaluation metrics. The evaluation metrics gives guidance not only on the parameters, but also on the measuring units of each parameter and the measurement method. It is applied in D5.5. The results of D5.5 concern the analysis of the set of five evaluation matrices that consist of 41 parameters in total. The findings show, that Eco-Bot positively influences the energy consumption of users – when compared to a group that only used the pilot's services without making use of Eco-Bot. Moreover, the newly designed NILM module demonstrated that it was consistently effective in achieving the objectives of the project.

DELIVERABLE NUMBER	WORK PACKAGE
D5.5	WP5
LEAD BENEFICIARY	**DELIVERABLE AUTHOR(S)**
adelphi	Kathrin Kohl (adelphi) Lena Domröse (adelphi) Joscha Gretzschel (adelphi) Aggeliki Giakoumaki (RISA) Sylwia Slupik (UEKAT) Lina Stankovic (USTRAT) Claudia Julius (SEC) Oriol Pla (DEXMA) Chloé Coral (EYPESA) Dora Karali (Erra) Stelios Kalogridis (Plegma)
QUALITY ASSURANCE	
Reviewer 1: Stephanos Camarinopoulos Reviewer 2: Aggeliki Giakoumaki Reviewer 3: Sylwia Slupik	RISA RISA UEKAT
PLANNED DELIVERY DATE	**ACTUAL DELIVERY DATE**
30/06/2021	22/10/2021
DISSEMINATION LEVEL	x PU = Public □ PP = Restricted to other programme participants □ CO = Confidential, only for members of the consortium

Table of contents

List of Figures

List of Tables

List of Acronyms and Abbreviations

AA	Any other Appliance
AI	Artificial Intelligence
AC	Air Conditioning
B2B	Business-to-Business
B2B2C	Business-to-Business-to-Customer
B2C	Business-to-Customer
CA	Consortium Agreement
CO	Confidential
DAC	Development Assistance Committee
DMP	Data Management Plan
DoW	Description of Work, referring to the Annex I of the Grant Agreement
EC	European Commission
EMS	Energy Management System
FP	False Positive
GA	Grant Agreement
HF	High Frequency
HVAC	Heating, Ventilation and Air Conditioning
ICT	Information and Communications Technology
iESA	SEnerCon's Energy Savings Account
IPR	Intellectual Property Rights
KPI	Key Performance Indicator
LF	Low Frequency
NILM	Non-intrusive Load Monitoring
NPS	Net Promoter Scale
OA	Open Access
OECD	Organisation for Economic Co-operation and Development
PPR	Project Progress Reports
PSB	Project Steering Board
PU	Public

QA	Quality Assurance
SAB	Security Advisory Board
SEC	SEnerCon
SOA	Service-oriented Architecture
STC	Scientific and Technical Committee
TP	True Positive
UC	Use Case
WP	Work Package

Executive summary

Deliverable D5.5 evaluates the results of the demo phase of Eco-Bot. The demo phase was conducted in three pilots. The residential pilots (SEnerCon and EYPESA) took place from April 2020 to May 2021. While the commercial pilot (DEXMA) took place from April to March 2021. The project team implements the evaluation approach defined in D3.3. The results of this deliverable concern the analysis of the set of five evaluation matrices that consist of 41 parameters in total.

The findings of the behavioural model evaluation show that the recommendations implemented by Eco-Bot – for all three pilots – meet the expectations regarding their usefulness for the users. Moreover, the empirical results (carried out on a geographically extended sample of 4506 respondents) show that the accuracy of the classification model was positively verified. The model classification error is 15% and falls within the assumed threshold not exceeding 20%. Thus, it meets the set target.

With regards to the metrics on the NILM (Non-Intrusive Load Monitoring), the project team developed new algorithms and sampling rates for each of the three pilots. The results show that restaurants and hotels had large change in operating procedures and unusual patterns of use during the Covid-19 pandemic lockdowns in 2020/2021. For supermarkets, the target of 80 % on accuracy in estimated energy consumption was met. In terms of NILM performance detection accuracy (for the residential pilots), the project team obtained an average classification accuracy of 72%. The NILM was consistently effective in reaching the project goals.

The subjective experience of participants using Eco-Bot is reflected in the results of the user surveys. Each pilot sent two surveys to the participating users during the pilot periods. The results show that the user satisfaction for most of the aspects of the bot was above average with a slight increase in the second survey at the end of the pilots. This can be attributed to the improvements of the bot's intelligence and the user interface. By the end of the pilots most of the survey respondents indicated that they would recommend Eco-Bot to a colleague or a friend.

The findings on user engagement and user retention for the residential pilots show the challenge to reach a consistent number of regular users, i.e. participants that use Eco-Bot regularly over the whole pilot period. The pilot partners conducted extensive efforts for recruitment and engagement of participants over the whole demo phase. As the number of total users changed over the period of the pilots, the project team performed a cohort analysis by grouping users into cohorts based on the month they started using Eco-Bot.

The lockdowns due to the Covid-19 pandemic affected – and partly increased – the energy consumption of most of the participants' homes (residential pilots) and of the supermarkets, (commercial pilot). With regard to energy savings, the target (15 % energy savings) was achieved by the Dexma pilot commercial buildings (restaurants and hotels). For the SEnerCon and the Eypesa pilot the total target – considering all Eco-Bot users – was missed. However, the target was achieved (and overreached) by several individual users participating in the residential two pilots. As a whole, the project team reasons that the use of Eco-Bot prevented a further increase

of energy consumption among its users. For the SEnerCon pilot, Eco-Bot participants achieved 1.9 percent electricity savings and 0.6 percent energy savings for space heating[1]. Thus, the targets of 15 percent electricity and 5 percent heating energy savings have not been achieved but the use of Eco-Bot has presumably prevented an increase of energy consumption among its users. Similarly, for the Eypesa pilot, Eco-Bot mitigated the increase of energy consumption. With regards to the commercial buildings (Dexma pilot), the energy saving targets were reached for the whole group of locations. An alternative methodology for the calculation of energy savings that considers other variables like degree days and the day of the week is presented.

The project team summarises all findings of the demo phase in 12 key takeaways (chapter 4):

1) Eco-Bot is rated to be a useful energy efficiency tool, offering diverse functionalities and covering use cases the users are interested in.

2) The NILM module was consistently effective in achieving the project's objectives.

3) Eco-Bot positively influences the energy consumption of users. Users showed a reduction or a lower increase in energy consumption.

4) Eco-Bot motivated both residential and commercial users to improve their energy consumption behaviour.

5) The application of smart meters is advantageous for Eco-Bot users.

6) Eco-Bot newcomers benefit from personal onboarding and guidance at first use.

7) The Covid-19 lockdowns affected the energy consumption of all Eco-Bot participants (residential and commercial).

8) The users' evaluation of Eco-Bot 's conversational intelligence increased across the pilot phase.

9) The mechanisms implemented for conversational intelligence of Eco-Bot allowed the improvement of its performance during the pilots and remain valuable for future use.

10) The behavioural model identifies the user's motivations, educates and encourages energy consumers to manage their energy efficiently and sustainably and successfully meets the project objectives.

11) When integrated in an Energy Management System, Eco-Bot increases its value.

12) Thanks to important feedback mechanisms in the bot, users made suggestions for additional features. These features were constantly implemented and improved Eco-Bot.

---

[1] Compared to iESA users, who had in general a 5 percent increase in electricity consumption and a 2 percent increase in energy consumption for space heating during the Covid-19 pandemic lockdowns, with a higher occupation of people in their homes.

The project team concludes, that the NILM module created for the project was effective. Moreover, the recommendations offered by Eco-Bot meet the expectations in terms of usefulness for the users. Despite challenges encountered due to the Covid-19 pandemic and partly missed targets, Eco-Bot proves that it can positively influence the energy consumption of both residential and commercial Eco-Bot participants. Additionally, the improvements implemented during the pilot phase regarding Eco-Bot's performance were not only efficient, but remain valuable for future use of the bot and its commercialisation.

1. Introduction

Deliverable D5.5 is based on Task T5.5 (Evaluation of the demonstration results). The aim of this deliverable is to evaluate the three pilots that were executed from April 2020 until the end of May 2021 (residential pilots SEnerCon and EYPESA) and from April 2020 to March 2021 (commercial pilot DEXMA). The basis for the evaluation and its approach was set in D3.3. It defines all relevant parameters that evaluate the effectiveness and acceptance of Eco-Bot as a personalized virtual energy assistant. The present deliverable is a continuation and implementation of D3.3 and validates the set targets. It presents and analyses the results of each of the 41 parameters (as defined in D3.3).

This deliverable provides aggregated insights into the results of the three pilots and shows the results of relevant key parameters.

The deliverable consists of seven chapters. Chapter 1 presents the introduction and the aim of the deliverable. Chapter 2 explains the validation methodology, which was previously defined in the DoW in line with the Key Performance Indicators (KPIs) and the targets for the Eco-Bot project. The proceeding chapter 3 depicts the main challenges occurred during the pilot phase and addresses mitigation actions. Chapter 4 presents and analyses the results of four of the five Eco-Bot evaluation matrices, while chapter 5 presents insights into qualitative results. Chapter 6 picks up the main outcomes of the evaluation and presents 12 key takeaways. The closing chapter 7 presents a conclusion of the deliverable.

2. Validation Methodology

All Key Performance Indicators (KPIs) and targets were defined in the proposal (Section 1.1 and Section 2.1 of the DoW). The respective evaluation metrics for Eco-Bot (defined in D3.3) use these KPIs as a basis and define all parameters relevant for an evaluation of the effectiveness and acceptance of Eco-Bot. Hence, the validation method for this deliverable is based on the approach defined in D3.3.

The evaluation of the Eco-Bot performance also considers the **five Development Assistance Committee (DAC) criteria** defined by the Organisation for Economic Co-operation and Development (OECD). Applied to the Eco-Bot case, these criteria and the respective questions to be considered for the evaluation are the following:

1. Relevance (To what extent does Eco-Bot meet the end-user's needs? And does it remain relevant to the customer?)
2. Effectiveness (To what extent were the objectives of the Eco-Bot project achieved?)
3. Efficiency (Were all objectives achieved on time?)
4. Impact (What happened as a result of the project? Did it have an impact on the energy savings behaviour of the Eco-Bot users? Have the energy saving targets been reached?)
5. Sustainability (To what extent can the benefits of the use of Eco-Bot - regarding energy efficient behaviour - continue after the end of the project?)

The criteria are considered for the analysis of the results (chapter 3). In chapter 4, the criteria are reflected upon once again in the key takeaways.

The evaluation of the three pilots was conducted in two phases: the first evaluation period and the end evaluation period, illustrated in the following timelines.

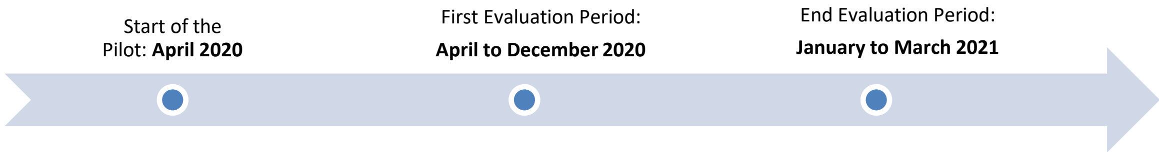

Figure 1: DEXMA Evaluation Periods

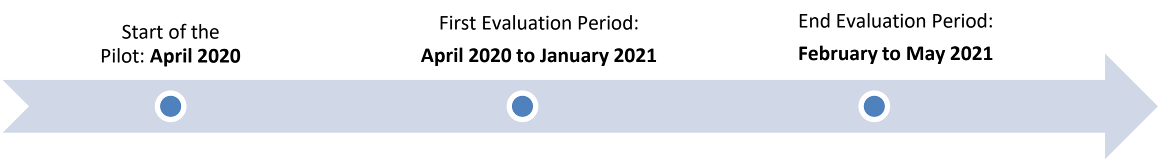

Figure 2: SEnerCon and EYPESA Evaluation Periods

The conduct of the evaluation of Eco-Bot's performance required a close and regular collaboration between all nine project partners – in particular towards the last months of the project. The figure below depicts the interaction of the partners and their respective roles.

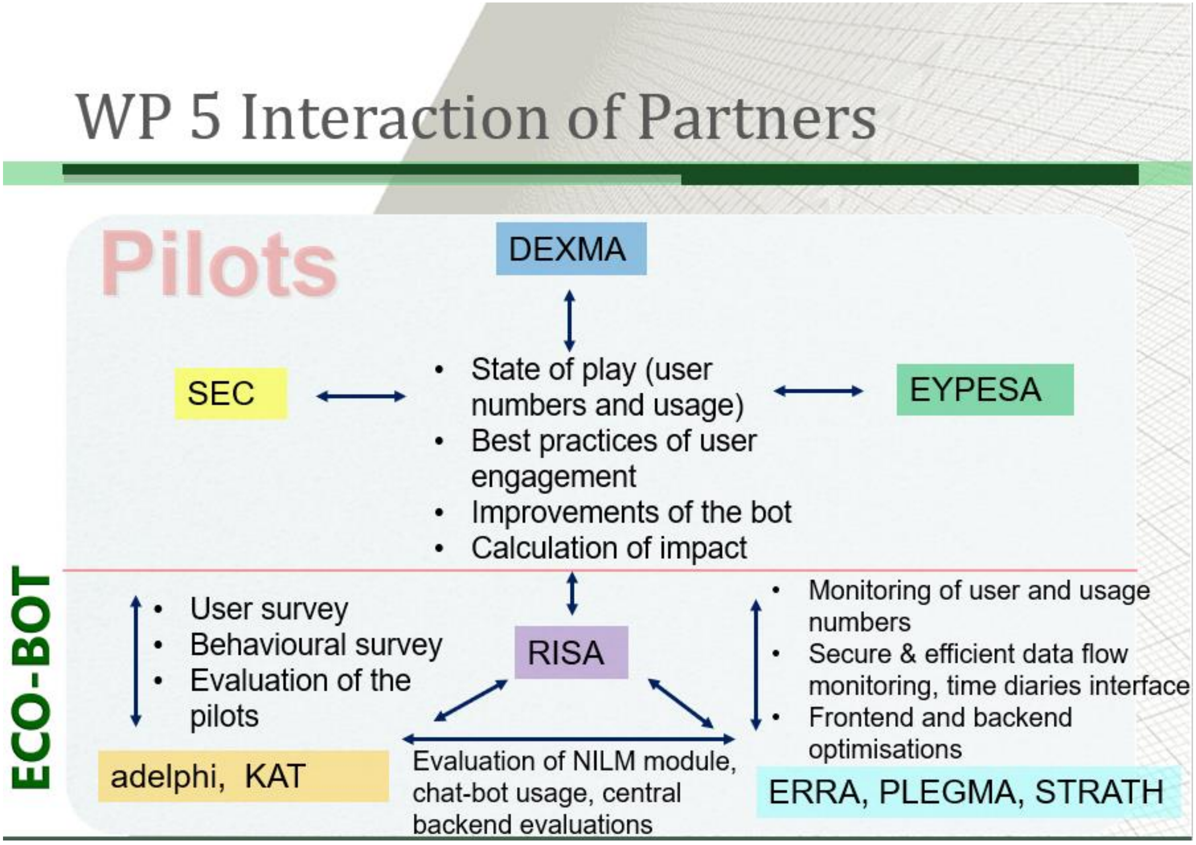

Figure 3: Interaction of Partners

3. Barriers and challenges during the pilot phase

For the residential pilots (SEnerCon and EYPESA) it proved to be challenging to reach a consistent number of regular users – despite all extensive efforts for recruitment and engagement of participants. This affected some of the parameter analysis and impeded to quantify the benefits brought by the use of Eco-Bot. In the case of the Spanish pilot, where only 15 regular users are considered in the evaluation of the metrics, the results are much more affected by the unpredictability and variability of each singular participant consumption. Therefore, it proved to be challenging to validate to which extent Eco-Bot influenced the electricity consumptions of the users and the advantages brought to them, based on the parameter calculation.

For the EYPESA pilot the findings prove that users entered the bot with more variability than expected, having months of intense activity alternated with several months of inactivity. However, this was the trend in the majority of participants and it does not correspond to the initial target set. Yet, the users were able to make good use of the bot, test it, implement recommendations and change behaviour. In addition, it was not possible to finalise the installation of most of the advanced meters (due to Covid-19 lockdowns). Therefore, the majority of users recruited during the pandemic were attributed to the basic cohort group.

Due to the exceptional situation of the Covid-19 pandemic, most people stayed at home working, while children were taught in online lessons. In some of the participants' homes this led to a higher energy consumption (for both residential pilots). Yet, in the case of the SEnerCon pilot, the Eco-Bot users had a slight reduction in energy consumption compared to the group of normal iESA users (i.e. households not using Eco-Bot but iESA only). Especially the results of the group of smart meter users are promising (10% savings, higher user satisfaction). But due to delayed smart meter roll-out in Germany, it was difficult to convince users to contract smart metering service at own cost[2]. The restricted number of features available for non-smart meter users and the precondition to enter energy data frequently exacerbated the issue. The findings revealed that for non-smart meter users, it can be of added value to have more linguistic features (e.g. information on specific energy topics such as solar panels or shared renewable projects). As a result of the non-continuous usage of Eco-Bot in the German pilot, it was challenging to evaluate the energy savings. This especially holds true for the energy savings for space heating, as it requires a sufficient number of months of the heating period. Thus, some users needed to be excluded from this evaluation. Despite of these challenges, successes could be achieved and especially some of SEnerCon's power users liked Eco-Bot, used it frequently and would be interested in a continuation of this additional service to the iESA.

---

[2] This might change in future, latest in 2032 when the smart meter application in all households will be obligatory

The commercial pilot (Dexma) achieved the savings goals set in D3.3 (despite the unusual operating hours of the recruited locations – supermarkets, restaurants and one hotel).

4. Quantitative results of the evaluation

This section presents the results of four of the five categories: 1) The Behavioural Model Evaluation Metrics, 2) The NILM Evaluation Metrics, 3) The Eco-Bot Impact Evaluation Metrics and 4) The Pilot Specific Metrics. Chapter 5 depicts a summary of the remaining category The Chatbot Evaluation Metrics – while the detailed findings of these metrics are presented in Annex B.

4.1. The Behavioural Model Evaluation Metrics

The behavioural model evaluation metrics includes the usefulness of the tailored recommendations by the Eco-Bot and consist of two parameters (P1 and P2). P1 deals with the usefulness of the recommendations, while P2 assesses the accuracy of the classification model.

P1 (Recommendation usefulness) evaluates the percentage of recommendations that were found useful by both the individual/household and commercial energy users. The recommendations, prepared and presented in D3.2, were evaluated during the WP5. The detailed results of the evaluation are reported in D3.4 with a summary of the most important points presented here.

For the individual/household (SEnerCon and EYPESA) energy users, 88.42% of the overall recommendations were seen as useful in accordance with the previous assumption. The overall value of the measure falls into the category of 80-100%, thus not requiring any changes to the recommendation set. Values of P1 for different segments vary (for some segments the value is as high as 100%, for some it is 70%). Appropriate action presented in D3.4 was taken accordingly to these values. The evaluated recommendations were analysed in order to discern which type of recommendations were most often evaluated as useful and not useful, i.e. diminishing comfort of users as one of the prominent reasons why some of the recommendations were found as not useful. In addition, changes to the recommendations set were proposed and implemented (see the annex to D3.4). Overall, it can be said that the prepared recommendation set was well aligned to the needs of individual/household users.

In the case of the commercial energy users (DEXMA), the overall value of P1 was lower than 75%. Yet it exceeded the 61% threshold set in D3.3 and is well aligned to the needs of energy/facility managers interested in using an energy management application. The detailed analysis (presented in D3.4) shows that the situation in the case of different segments of commercial energy users is more diverse than in the case of the individual/household energy users. In the case of the more homogeneous segments (like hotels or supermarkets), the values of P1 exceeded the 80% threshold. This indicates a very good alignment. However, the alignment was lower in the case of the more diverse segments that include different types of facilities. It was expected that with diverse segments the need for specific types of recommendations would be greater. Detailed analysis of which recommendations were evaluated as useful or not by different users allowed for an interesting observation: there are similarities between particular user segments that depend on the size of the facility more than on the type – e.g. larger facilities more often evaluated recommendations considering automation of the processes or the development of internal policies concerning energy management to be more useful than smaller facilities.

The recommendation usefulness threshold of 61-80%, assumed in D3.3, indicates the need to reformulate and/or better specify selected recommendations from the set. This action was completed. Suggestions for actions concerning recommendations' sets for different user segments are reported in D3.4 and all new or adapted recommendations in case that such action was needed can be found in the annexes to the D3.4.

Recommendations for both the individual/household and commercial energy according to the assumption presented in D3.3 meet the expectations concerning their usefulness for the users.

The procedure of evaluating the accuracy of the classification model should be carried out on the basis of a detailed behavioural survey of Eco-Bot application users, obtained during the pilot phase.

Due to the small size of this group (mainly due to the Covid-19 pandemic and likely resulting from unsystematic use of the application as well as recruiting new eco-bot users at different time intervals), the estimation of the model accuracy may be erroneous and unreliable. There is a strong belief that such an assessment would be based on randomly treated data and not on methodological assumptions and consequently would not reflect the actual quality of the model's functioning. Therefore, the project team decided to increase the size of the research sample and expand it geographically considering individual energy consumers, who are not using Eco-Bot. These energy consumers reside in the following countries: France, Romania, Czech Republic, Poland, Greece, Great Britain, Spain and Germany.

Although the remodelling procedure was carried out on a different research sample than planned during the construction phase of the first version of the model, it should be noted that the theoretical research assumptions remained unchanged. This means that all the elements of a) the segmentation approach (described in D3.1 report), b) the research tool (survey questionnaire) and c) the statistical methods (the discriminant model described in D 3.2 report) used in the procedure are the same as if the model that was built and used in the Eco-Bot application. Accordingly, it is possible to assess the accuracy of the classification model.

The model was built on the training set, which was formed using the data provided beforehand in the extended survey. In the survey, the most important elements were the segmenting questions. The answers to these questions allow to recognize the true behavioural type of the user, according to the segmentation procedure (described in D3.2) identified as a ground truth. The next step was to examine whether the classification model assigns the given user to the same segment as was derived from the extended survey. This method allowed to evaluate the percentage of the correct classification model predictions (comparing the user's behavioural type obtained from the classification model with the ground truth).

The model classification error estimated on the entire research sample – considering all questions from the survey – is 10%. However, asking several questions during the registration process may result in discouragement or weariness of users and their resignation from using the bot. Considering the needs of Eco-Bot, the project team decided to limit the number of questions, allowing the classification of users to the appropriate segments only to the most important 8, which at the same time increases the model error to 15%, but is within the acceptable error range.

Other versions of the model with different parameters were also checked and described in detail in the D3.4 report. It should be noted that all analysed variants of the model meet the metric P2 (as described in D3.3). This is caused by the fact that in each case the classification agreement between the true behavioural type (determined using clustering methods) and the segments to which the users were assigned by the classification model is above or equals 80%.

In conclusion, P2 evaluation metric was positively verified.

4.2. The NILM Evaluation Metrics

The NILM Evaluation Metrics evaluate not only the accuracy of the NILM algorithm, but also the effectiveness and performance of the itemized billing service. It is composed of five parameters (P3 to P7).

Parameter 3[3] (related to KPI_4.1) and P4[4] (related to KPI_4.2) were both completed and reported in D4.2. In summary, the findings show that – due to the limits of sampling rate of smart meter deployments at scale and the lack of submetering data – we would tackle the challenge of disaggregation of active power at very low sampling rates (15 mins, 1 hour) and low sampling rates (10 sec) using transfer learning. A detailed evaluation of the state-of-the-art proved that there were no NILM algorithms suitable for any of the three Eco-Bot pilots. Therefore, the project team developed novel algorithms for each of the pilots and sampling rates. It demonstrated that it was able to disaggregate at performances close to those reported in the literature where transfer learning is not used. Transfer learning inherently poses a risk since there is no way of telling how individuals use their appliances in their home and we can only learn based on the available data.

Metrics P5, P6 and P7 measure performance of transfer learning during demonstration. The following sections present the findings.

4.2.1. Metrics for commercial pilot (DEXMA)

Parameter 5[5] (related to KPI_4.4) was evaluated during Task 4.2 and reported in D4.2 on public datasets, as well as during Task 4.5 and reported in D4.5 for the small-scale validation. The final P5 evaluation was carried out as part of the final demonstration phase in WP5. This metric can be calculated only if submetering data is available. This is only relevant for the DEXMA commercial pilots for the final demonstration.

Here, the evaluation results of P5 are presented for the commercial buildings (DEXMA) for the first evaluation phase 01 April 2020 to 31 December 2020, and the second evaluation phase 01 January 2021 to 31 March 2021. It should be noted that pilots had access to daily consumption prediction of disaggregated loads, and the algorithm did not make use of square footage of the building nor geographical conditions such as temperature or daylight hours to minimise dependency on contextual data availability and maintain scalability of the model.

There were three categories of commercial buildings, namely supermarkets, hotels, and restaurants. Whilst supermarkets operated more often during the Covid-19 pandemic lockdown in the period 06 April 2020 to 05 April 2021 which included both evaluation phases, this was not the case for hotels and restaurants. Furthermore, all three categories did not

---

[3] Accuracy and limitations of state-of-the-art NILM algorithms at low smart-meter data sampling rates on real datasets at scale
[4] Practicality, scalability and near real-time suite of NILM algorithms that yield accurate appliance-specific disaggregation, with little to no training, robust to appliance heterogeneity
[5] Accuracy in estimated energy consumption

operate under normal conditions, with reduced capacity and partial operating hours of loads of interest especially concerning Lighting and Heating, Ventilation and Air Conditioning (HVAC).

As Figure 4 shows, consumption estimation accuracy of refrigeration performed the best out of the sub-metered appliances with 9 out of 10 supermarkets achieving above 80% accuracy. For the first evaluation phase, the average P5 metric across all 10 supermarkets and load categories is 77%. Per sub-meter, we achieved an average performance of 83% for Refrigeration, 73% for HVAC, 82% for Lighting, and 70% for AA across all supermarkets. Here AA represents any other appliance, which included machinery and large appliances.

Figure 5 shows that for the second evaluation phase, the average P5 metric across all 10 supermarkets and disaggregated loads is 79%. This includes 81% for Refrigeration, 74% for HVAC, 81% for Lighting, and 81% for AA, respectively per submeter.

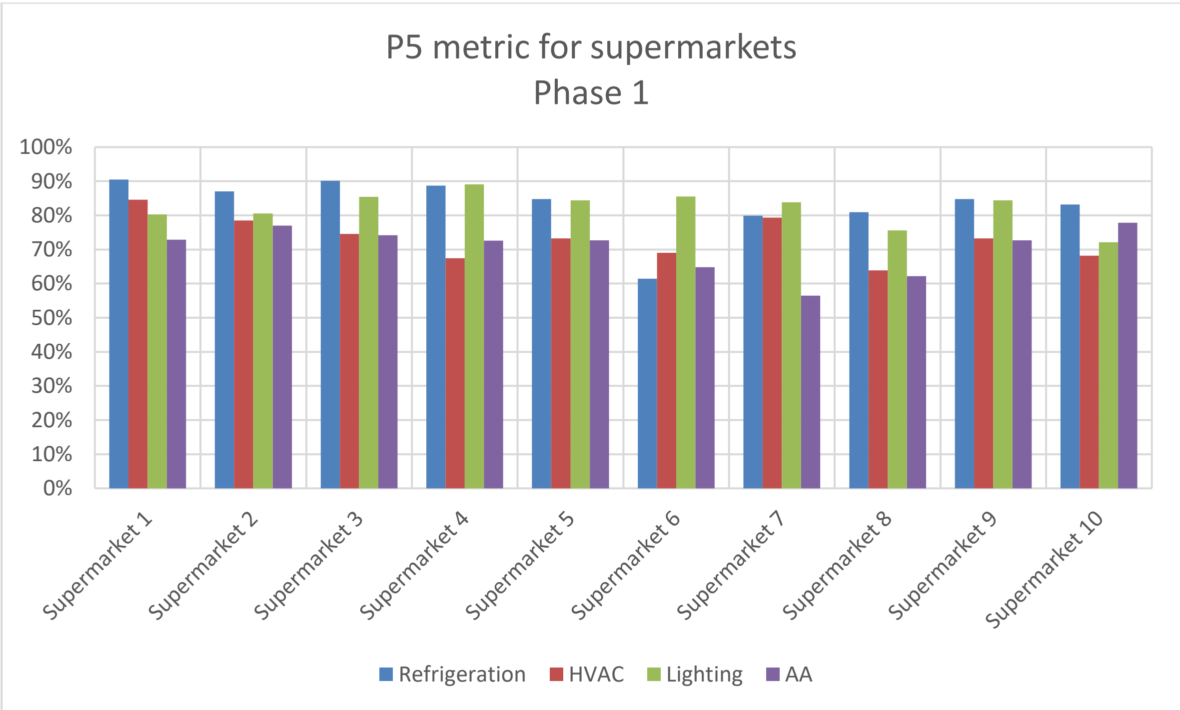

Figure 4: Energy consumption accuracy for first evaluation phase 06 April 2020 to 31 December 2020

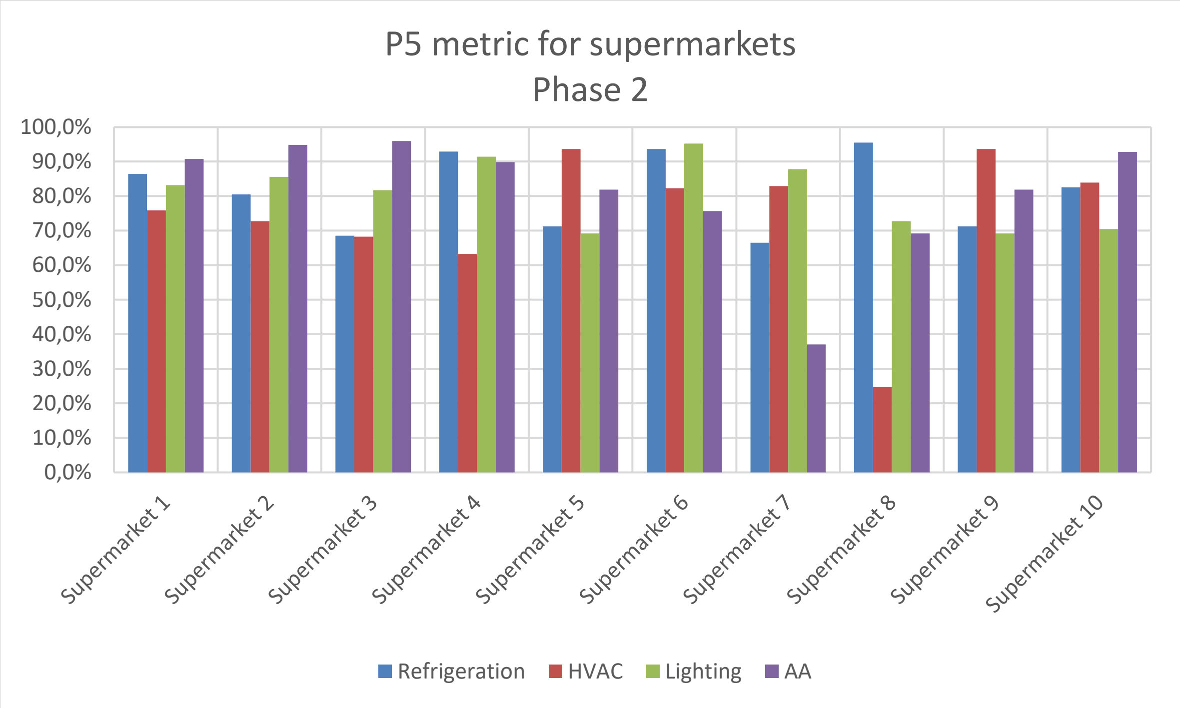

Figure 5: Energy consumption accuracy for second evaluation phase 01 January 2021 to 31 March 2021

After the first lockdown, we analysed the effect the pandemic had on our pilots. The observations prove (Figure 6) that patterns of consumption – which we had expected were either no longer present or were reduced – show aggregate consumption pre-Covid-19 and at the peak of Covid-19. The difference in consumption is coherent with the unmetered appliances "AA". For the same days, one year apart, there has been a steep change in the normal consumption pattern of the supermarket. This dip has also been observed in other supermarkets which previously did not show this pattern. The other sub-metered appliances were very similar in their consumption and patterns of use. Specifically, lighting during opening times was reduced by 5%, HVAC was increased by 4% (this can be attributed to heating during the winter period), and AA (e.g., machinery and large appliances, not specifically grouped by the sub-metering) was decreased by 20% when open on a time of year ratio. As a result, the implementation of how store openings were handled was adjusted as opening times in some cases had shifted, and weights for lighting, HVAC and AA were adjusted accordingly, resulting in keeping close to our original target of 80% accuracy for P5.

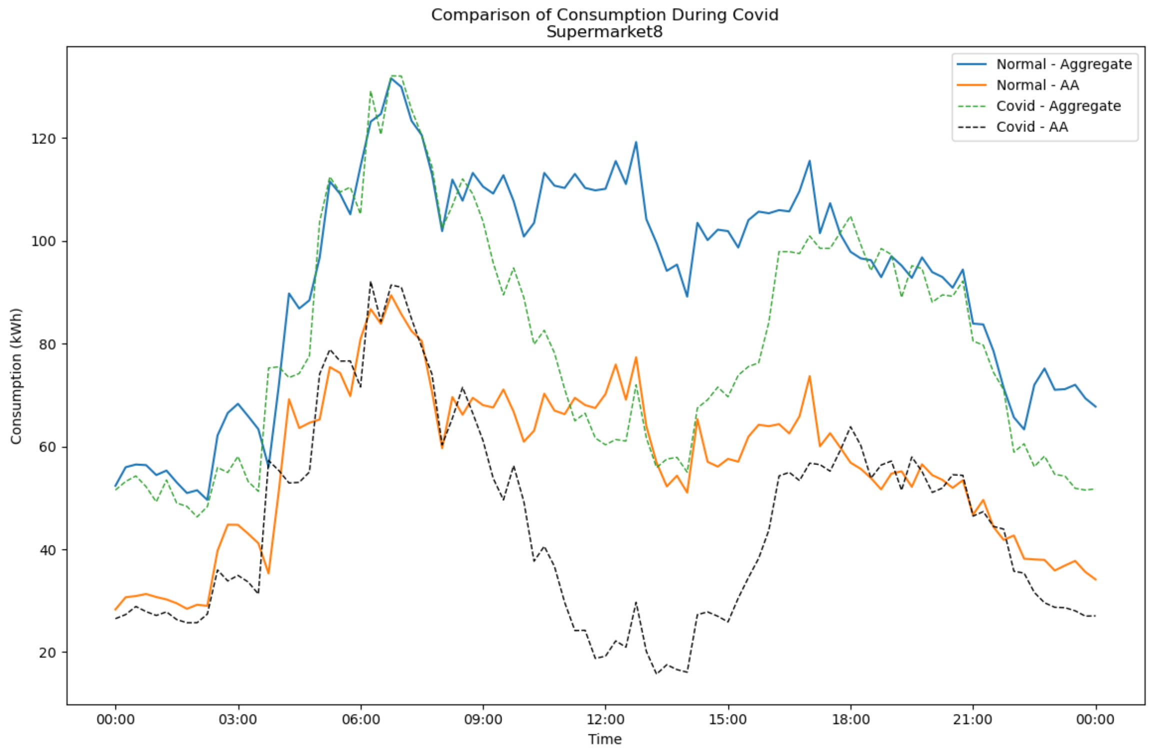

Figure 6: Actual energy consumption of Supermarket 8
02 March 2020 vs 02 March 2021

Next, we discuss P5 performance of restaurants. The data that we had for restaurants did not include all the possible appliances and in the evaluation phases, the monitored restaurants have a wide variety of circuits monitored which were not directly linked to the loads the commercial NILM had initially been trained on and reported in D4.2. Since sub-metered loads (used purely as ground truth for evaluation) vary from restaurant to restaurant, we grouped the sub-metered loads into the categories the commercial NILM algorithm had originally been trained on for comparison with the small-scale study reported in D4.5. These are summarized in Table 1.

Table 1: Grouping of sub-metered loads for the purposes of evaluation

HVAC	Machinery	Refrigeration
AC	Coffee	Freezer
Aircon	Oven	Cold
Central	Dishwasher	
AC1	Electric	
AC2	Extraction	
	Deep	

Restaurants also had the largest amount of unavailable data via submetering, likely due to closure of these buildings during the pandemic. There was also a service gap in aggregate consumption data for during 07 to 18 of January 2021 for restaurants 2, 3, 4, 5, and 6. All these restaurants stopped sending aggregate data from the 15 February 2021, during Phase 2. Lack of submetering data meant we could not evaluate NILM performance for the affected periods since submetering data provides the ground truth of consumption and enables evaluation of NILM accuracy. Table 2 summarises the dates for which evaluation for both phases was performed, considering submetering and aggregate consumption data availability discussed previously.

Table 2: Available sub-metering and aggregate data dates used for evaluation

	Phase 1 evaluation (Apr – Dec 2020)	Phase 2 evaluation (Jan - Mar 2021)
Restaurant 1	01 Apr to 24 Apr, 05 Jun to 31 Dec	No submetering data
Restaurant 2	11 Sept to 31 Dec	01 Jan to 07 Jan; 19 Jan to 15 Feb
Restaurant 3	26 May to 31 Dec	01 Jan to 07 Jan; 19 Jan to 15 Feb
Restaurant 4	01 Apr to 11 May; 19 May to 31 Dec	01 Jan to 07 Jan; 19 Jan to 15 Feb
Restaurant 5	18 Sept to 29 Nov	No submetering data
Restaurant 6	09 Jun to 31 Dec	01 Jan to 07 Jan; 19 Jan to 15 Feb
Restaurant 7	01 Apr to 31 Dec	01 Jan to 31 Mar

Restaurants 2, 4, and 6 contained several similar sub-meters for the Machinery category, which returned results that were close to our target P5 accuracy of 80%, reporting 74%, 74%, and 73% respectively for phase 1. Phase 2 had slightly less disruption that Phase 1 in terms or total shutdowns, however, opening times were still different. Phase 2 Machinery performance was 80%, 75%, and 85% accuracy for restaurants 2, 4 and 6 respectively. Accuracy was calculated on days where sub-metering data was available across both phases.

Restaurants had the largest change in operating procedure during the pandemic, with many eateries closing for prolonged periods of time, or altering operating hours significantly. In Figure 7, we show the change in consumption between pre-Covid-19/normal and during Covid-19.

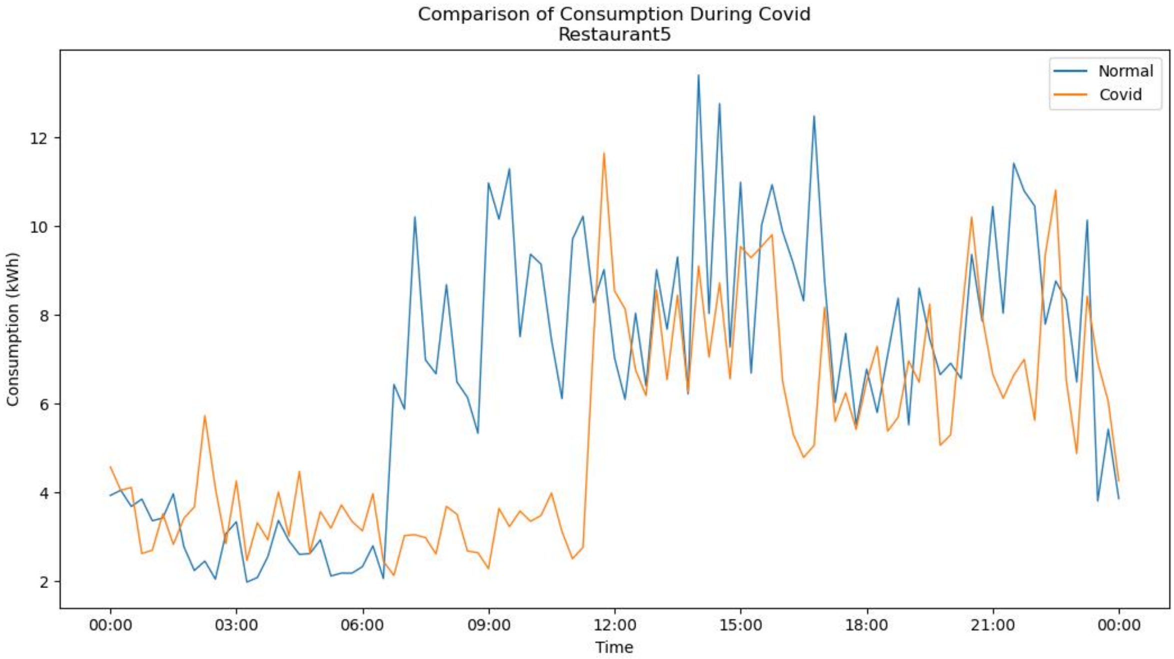

Figure 7: Actual aggregate energy consumption for Restaurant 5 pre- and during Covid-19 on 02 March 2020 and 02 October 2020, respectively

The hotel that was monitored likely had to shut down for guests during the pandemic. In fact, total consumption for the period of June-July-August 2020 was down by around 30% from the 2019 season, as can be seen in Figure 8. The floor sensor data from DEXMA also showed a discrepancy between the main and sub-meters, with the floor sensors having negative values as well as values higher than designated “main” aggregate sensor.

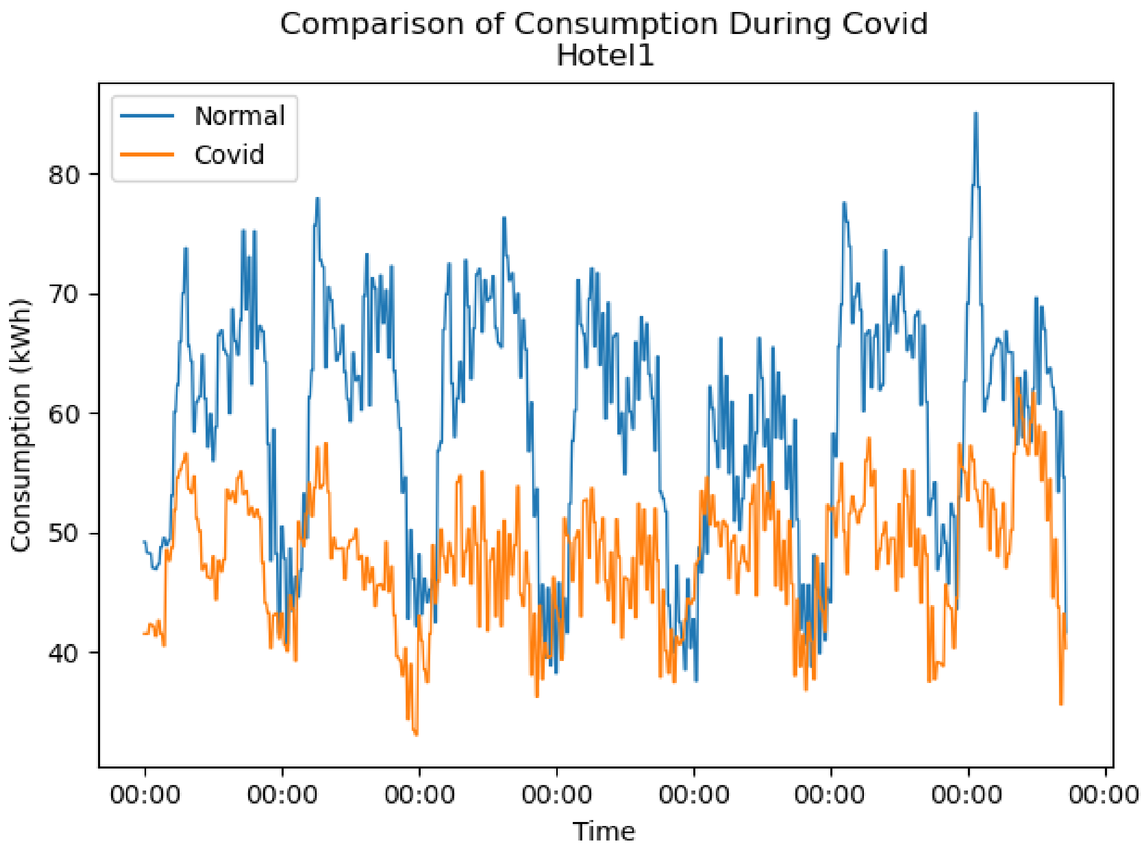

Figure 8: Actual aggregate energy consumption for Hotel 1 pre- and during Covid-19 on 19/02/2020 to 26/02/2020 and the same dates in 2021

Similarly, to the restaurant there are many sub-meters with different names from what were available to train on. The grouping used for evaluation is summarised in Table 3.

Table 3: Grouping of Hotel 1 sub-metered labels for evaluation into Refrigeration, HVAC and Lighting load categories

Hotel	Sub-metered load labels	Category
1	Chiller1	Refrigeration
	Chiller2	
	MCCP1	HVAC
	MCCP2	
	MCCP3	
	MCCP4	
	MCCP5	
	Conference	Lighting
	FirstFloorNorth	

	FirstFloorSouth	
	GroundFloor	
	SecondFloorNorth	
	SecondFloorSouth	
	ThirdFloorNorth	
	ThirdFloorSouth	

Evaluation performance P5 for the hotel for phases 1 and 2 combined (01 April 2020 to 31 March 2021) were as follows: Refrigeration 71%, Lighting 60%, and HVAC 64%.

4.2.2. Metrics for residential pilots (SenerCon and EYPESA)

Evaluation of NILM performance in terms of detection accuracy (P6 metric) and reliability of time diary (P7 metric) for the residential pilots was only available through time diaries due to unavailability of submetering data (as discussed in D4.5). During Phases 1 and 2, a total of 42 days of time diaries were filled and returned by 13 different households: 3 SEnerCon pilot houses returning smart meter data at 10sec resolution denoted by S in Table 4; 9 EYPESA pilot houses returned data at hourly resolution denoted by E-LF in the table; and 1 EYPESA pilot house returned data at 1min resolution denoted by E-HF. In Table 4, we list all 42 time-diary day entries (date column) corresponding to the House IDs and pilots. The date of the time diary entry is recorded together with the appliances for which we returned the NILM result. The last two columns indicate, for the corresponding day, the diary entry (ON if there is a non-zero entry and OFF if there is no entry) and the NILM result returned (ON if a non-zero consumption was returned, OFF if a zero consumption is returned for that day).There are three time diary entries that we could not verify because EYPESA did not receive data from the smart meter that day and the day after. This could be due to communication problems or malfunctioning of the meter.

Table 4: Cross-check of appliance usage time diary and NILM return (refer to Annex C for plots)

UID	PILOT	DATE	APPLIANCE	DIARY ENTRY AS	DETECTED AS	REASON	SEE PLOT
6TW9NWMQYFCVBDP5QQAD23	E-LF	22/09/2020	Dishwasher	OFF	OFF	TN	
		-	Washing Machine	OFF	OFF	TN	
			Tumble Dryer	OFF	OFF	TN	
8BD815PHSQQAEUJGAFDNNU	E-LF	30/07/2020	Dishwasher	OFF	-	Missing Data	
		-	Tumbler Dryer	OFF	-	Missing Data	
		-	Washing Machine	OFF	-	Missing Data	
9AT8S8QVHJKY625ZPDAW2S	E-LF	20/01/2021	Dishwasher	OFF	ON	FP Model	Figure 54
		-	Washing Machine	ON	ON	TP	
		-	Tumble Dryer	ON	ON	TP	
		18/02/2021	Tumble Dryer	ON	ON	TP	
		-	Washing Machine	ON	OFF	FN Time diary	Figure 42
		-	Dishwasher	ON	ON	TP	
FZWBBJUYZ6N9QY6TJFG5WQ	E-LF	03/02/2021	Dishwasher	ON	ON	TP	
		-	Tumble Dryer	OFF	ON	FP Model	Figure 55
		-	Washing Machine	OFF	ON	FP Model	Figure 55
		23/03/2021	Washing Machine	ON	ON	TP	
		-	Tumble Dryer	OFF	ON	FP Model	Figure 56
		-	Dishwasher	OFF	ON	FP Model	Figure 56

		28/03/2021	Washing Machine	ON	ON	TP	
		-	Tumble Dryer	OFF	OFF	TN	
		-	Dishwasher	OFF	OFF	TN	
K2SJAJF93VTBCT7KCQSESD	S	03/06/2020	Washing Machine	OFF	ON	FP Model / Super High Consumption	Figure 52
			Tumble Dryer	OFF	OFF	TN	
			Dishwasher	OFF	ON	FP Model / Super High Consumption	Figure 52
		15/06/2020	Washing Machine	OFF	ON	FP Model / Super High Consumption	Figure 53
			Tumble Dryer	OFF	OFF	TN	
			Dishwasher	OFF	ON	FP Model / Super High Consumption	Figure 53
NPQSUXTRMWMNVI6D61GWXB	E-LF	14/05/2020	Washing Machine	ON	ON	TP	
		-	Dishwasher	ON	ON	TP	
			Tumble Dryer	ON	ON	TP	
		17/05/2020	Washing Machine	ON	ON	TP	
		-	Dishwasher	ON	ON	TP	
			Tumble Dryer	ON	ON	TP	
NQFPRNPGZPHCY1YGKQBQWK	E-LF	31/01/2021	Dishwasher	OFF	-	Data Missing	

			Washing Machine	ON	-	Data Missing	
		01/02/ 2021	Dishwashe r	OFF	OFF	TN	
			Washing Machine	OFF	OFF	TN	
		02/02/ 2021	Dishwashe r	ON	ON	TP	
			Washing Machine	OFF	ON	FP Model	Fig ure 57
		03/02/ 2021	Washing Machine	OFF	ON	FP Model	Fig ure 58
			Dishwashe r	OFF	ON	FP Model	Fig ure 58
		04/02/ 2021	Dishwashe r	ON	ON	TP	
			Washing Machine	OFF	ON	FP Model	Fig ure 59
		07/02/ 2021	Washing Machine	ON	ON	TP	
			Dishwashe r	OFF	OFF	TN	
		12/02/ 2021	Washing Machine	ON	-	Data missing	
			Dishwashe r	OFF	-	Data missing	
NWTUKANECUR64M SJGWELZK	E- LF	21/01/ 2021	Dishwashe r	ON	ON	TP	
			Washing Machine	ON	ON	TP	
		02/02/ 2021	Dishwashe r	OFF	ON	FP Model	Fig ure 60
			Washing Machine	OFF	ON	FP Model	Fig ure 60
PYLS83FNNZZ1F8HI BVUKYU	E- HF	30/05/ 2020	Dishwashe r	ON	ON	TP	
			Washing Machine	OFF	OFF	TN	
		31/05/ 2020	Dishwashe r	ON	ON	TP	
			Washing Machine	ON	OFF	FN Model	Fig ure 48

		01/06/2020	Dishwasher	ON	ON	TP	
			Washing Machine	OFF	OFF	TN	
		02/06/2020	Dishwasher	ON	ON	TP	
			Washing Machine	ON	OFF	FN Model	Figure 49
		03/06/2020	Dishwasher	ON	ON	TP	
			Washing Machine	OFF	ON	FP Model	Figure 61
		04/06/2020	Dishwasher	ON	OFF	FN Model/App Type	Figure 43
			Washing Machine	ON	OFF	FN Time diary	Figure 43
		06/06/2020	Dishwasher	ON	ON	TP	
			Washing Machine	ON	OFF	FN Time Diary	Figure 44
		15/06/2020	Washing Machine	ON	ON	TP	
			Dishwasher	OFF	OFF	TN	
		16/06/2020	Dishwasher	ON	OFF	FN Model	Figure 65
			Washing Machine	OFF	OFF	TN	
		17/06/2020	Dishwasher	ON	ON	TP	
			Washing Machine	ON	ON	TP	
		20/06/2020	Washing Machine	OFF	ON	FP Model	Figure 62
		-	Dishwasher	OFF	ON	FP Model	Figure 62
SE76BW67DKS8XA2GFMFWZS	S	25/05/2020	Dishwasher	ON	ON	TP	
			Tumble Dryer	OFF	OFF	TN	

			Washing Machine	ON	ON	TP	
		27/05/ 2020	Washing Machine	ON	ON	TP	
		-	Dishwasher	ON	ON	TP	
			Tumble Dryer	ON	OFF	FN Missing Data	Figure 9
VTUS8R35ARK4UQC MAGPDMM	E-LF	22/09/ 2020	Dishwasher	ON	OFF	FN Model	Figure 50
			Washing Machine	OFF	OFF	TN	
W1MRKTKCOYKFUC D9EKIPTG	S	08/05/ 2020	Washing Machine	ON	OFF	FN Model	Figure 45
		-	Dishwasher	ON	OFF	FN Time Diary	Figure 45
			Tumble Dryer	OFF	OFF	TN	
		09/05/ 2020	Dishwasher	OFF	OFF	TN	
		-	Washing Machine	OFF	OFF	TN	
			Tumble Dryer	OFF	OFF	TN	
		10/05/ 2020	Washing Machine	ON	OFF	FN Time Diary	Figure 47
			Dishwasher	OFF	OFF	TN	
			Tumble Dryer	OFF	OFF	TN	
		11/05/ 2020	Dishwasher	ON	ON	TP	
			Washing Machine	OFF	ON	FP Model	Figure 63
			Tumble Dryer	OFF	OFF	TN	
		12/05/ 2020	Dishwasher	OFF	OFF	TN	
			Washing Machine	OFF	OFF	TN	
			Tumble Dryer	OFF	OFF	TN	

		13/05/2020	Washing Machine	ON	OFF	FN Model / App Type	Figure 51 & Figure 64
			Dishwasher	OFF	OFF	TN	
			Tumble Dryer	OFF	OFF	TN	
		14/05/2020	Dishwasher	ON	OFF	FN Time Diary	Figure 47
			Washing Machine	OFF	ON	FP Model	Figure 47
			Tumble Dryer	OFF	OFF	TN	
XN8OVTZNHPNBQX5MEOVDXQ	E-LF	18/05/2020	Washing Machine	ON	ON	TP	
			Dishwasher	ON	ON	TP	

The fifth column of Table 4 shows the ground truth as reported by users in time diaries versus the NILM result returned to pilots (6th column) for all 42-time diary entries received for the residential pilots of EYPESA (E-LF + E-HF) and SEnerCon (S). Column 7 explains the results briefly (in terms of TP, TN, FP, and FN) and column 8 is a cross-reference where incorrect detection is explained visually later in this section.

As shown in Table 4, we correctly detected most of the appliance usages (as indicated by TP and TN in the 7th column), within the scope of the returned time diaries. A true positive (TP) is when an appliance is correctly detected as on when it was recorded as on in the time diary. A true negative (TN) is when an appliance is correctly detected as OFF when it was recorded as OFF in the time diary. A false negative (FN) is when an appliance is incorrectly detected as off when it was recorded as on in the time diary. A false positive (FP) is when an appliance is incorrectly detected as on when it was not recorded as on in the time diary. In summary, there were 34 true positives, 29 true negatives, 13 false negatives and 19 false positives.

Of the 13 false negatives, 1 is due to missing data at exactly the time when the appliance was run (Figure 9), 6 are because of inaccurate time diaries (Figure 42, Figure 43, Figure 44, Figure 45, Figure 46, Figure 47), 3 (dishwasher in Figure 43, washing machine in Figure 45, Figure 51) are due to appliance cycles that differ significantly from typical appliance signatures, 3 (Figure 48, Figure 49, Figure 50) are due to other large loads operating at the same time as the missed appliance.

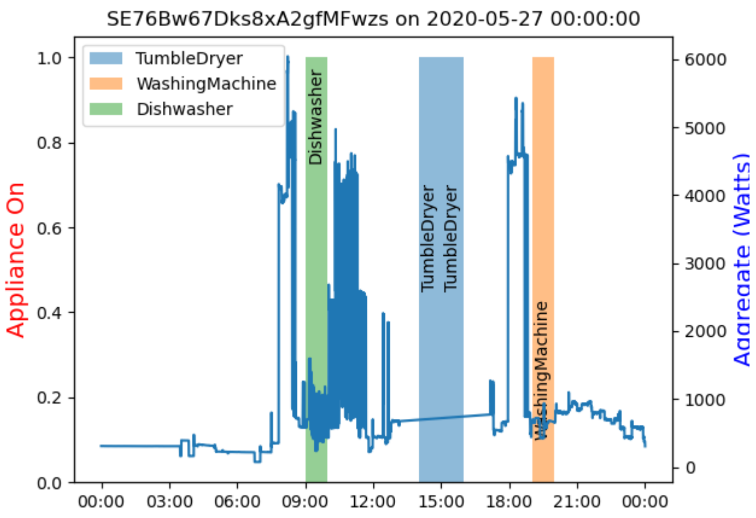

Figure 9: Missing data explaining NILM OFF result when time diary recorded ON for tumble dryer - SEnerCon user

Please refer to Annex C for a detailed analysis of metrics P6 and P7

As reported in deliverable D4.5, for the small-scale evaluation of P7, time diaries are somewhat unreliable, as they require people to record what appliances were operational during a specific day after the event. Time diaries from low frequency households, e.g., Figure 42, are easier to validate since an appliance run appears as a distinct spike in consumption when it has been used. For the high frequency time diaries, it becomes more obvious when an event has been misremembered (e.g. Figure 46 and Figure 47) or perhaps entered for a different day by accident.

To summarize the findings of P5, P6 and P7: For commercial NILM (DEXMA Pilot), we report the P5 metric, which is the most accurate NILM metric due to the availability of submetering. As discussed in section 4.2.1, we obtained 83% and 79% accuracy for supermarkets for 1st and 2nd evaluation periods respectively, meeting our target of 80% for P5 metric. Due to Covid-19 lockdown, restaurants were closed on many days, or submetering was unavailable, or were only partially operational. This resulted in 74% average consumption accuracy for Phase 1, that encountered the most disruption, and 80% average accuracy for Phase 2 when disruption was lower. The hotels were most affected by Covid-19 lockdown in terms of their energy consumption. Very unusual patterns of use emerged due to partial opening hours and load usage, but despite this, an average consumption accuracy of 65% was obtained over both evaluation periods.

For residential households discussed in section 4.2.2, we obtained an average classification accuracy of 72% (P6 metric) over both LF and HF households. Considering the unreliability of time diaries (P7 metric), this is in line with performance obtained during the small-scale validation period. This is acceptable for large scale deployments of heterogeneous households running many appliances, some of which have slightly different energy signatures to what is commonly expected in public datasets used for training. This indicates that there is always room for improvement by including additional appliance signatures, through submetering or crowdsourcing.

4.3. The Eco-Bot Impact Evaluation Metrics

This subchapter portrays the parameters that evaluate the different dimensions Eco-Bot has an impact on. It compares users' energy consumption before using Eco-Bot and after. These dimensions include energy saving actions, as well as the environmental and economic impact. Furthermore, this section presents parameters concerned with measuring Eco-Bot's impact on user's awareness for the rebound effect.

Despite the low number of users participating in the residential pilots, the project team considers the impact-related results of all three pilots as relevant for the project and beyond. While the residential pilots did not reach all of the energy and monetary saving targets, the commercial pilot did reach the targets. Moreover, a number of individual users of the residential pilots reached the energy savings targets (despite the lockdowns due to the Covid-19 pandemic). Additionally, the targets of the parameter "implemented energy savings measures" were reached and the outcomes of the parameters "users, who made a behavioural change" and "users, who made an investment" are positive – this shows the effectiveness of Eco-Bot as an energy saving assistant.

For the evaluation of the energy savings, only regular users were considered as only for these users Eco-Bot can be considered to influence their energy consumption. For SEC's pilot regular users were defined as users that used the bot at least three times and at least every two months except for the summer months June to September 2020. For EYPESA's pilot regular users were defied as users that did not have more than 3 months of inactivity in a row and logged in at least for a number of times equal to the total number of participation months.

Comparing the energy consumption of users before and after employing Eco-Bot, P25 indicates how much energy savings Eco-Bot users achieved as a percentage of their baseline consumption. This metric is calculated individually by each pilot, using the energy consumption data history of their users.

4.3.1. Results of energy saving actions related parameters

Parameters 25 to 41 monitor the behavioural changes and investments facilitated by the usage of Eco-Bot. The targets presented here were either provided in the DoW or have been derived from the assessment of possible scenarios, partly adapted to the context of the respective pilots.

SEnerCon (B2B2C)

The calculation of energy savings considers only the energy consumption data of regular users[6]. For the evaluation, only the months were considered in which the users registered and used Eco-Bot, including one month after their last interaction, as the effect of the interaction might be delayed. This consumption data was then compared to the data of similar months of the last year before the start of the pilot in April 2020. For some regular participants historical data as well as recent energy data for space heating was missing (hence, not considered for the evaluation). For the majority of users however, the data base was sound, consisting of a sufficient number of meter readings and energy bills. The iESA calculates monthly energy consumptions from the data base and generates diagrams with consumption figures indicated by mouse-over function. For the evaluation of energy savings for space heating the annual consumption of 2020 was compared to the consumption of 2019, both values were climate-adjusted to be comparable.

Due to lockdown measures caused by the Covid-19 pandemic, a lot of people worked from home and children were taught in online lessons. This led to a higher occupation rate in most of the participants' homes. As a result, more energy was spent for cooking (restaurants and canteens were closed), ICT devices and space heating. Evaluations of the energy consumption of iESA users revealed an 5% increase in electricity consumption and a 2 % increase in heating energy consumption7. Thus, the savings achieved by pilot participants can be most likely related to the usage of Eco-Bot.

To conclude, it can be noted that 6 of the participants were able to reach the target of at least 15% electricity savings and 3 participants achieved energy savings for space heating of at least 5 %.

[6] Regular users are users, who used the bot at least three times and at least every two months (excluding the summer months from June to September 2020).

[7] https://www.co2online.de/corona-energiebilanz/?mobile=1&utm_source=regular&utm_medium=email&utm_campaign=co2online+news&utm_content=co2&utm_term=354837976236+2021-05+Newsletter+co2online+05.07.2021

Table 5: Total increase of energy savings by participating user - SEnerCon

P25 - Total increase of energy savings by participating user: Percentage increase of energy savings after Eco-Bot use		Target reached
1.9 % total savings for electricity and 0.6 % total savings for space heating after Eco-Bot use		
Electricity savings in the user group with smart meters	10 % total savings after Eco-Bot use	NO
Electricity savings in the user group with conventional mechanical meters	0.01 % total savings after Eco-Bot use	No
Electricity savings (both groups)	1.9 % total savings after Eco-Bot use	No
Energy savings for space heating (both groups)	0.6 % total savings after Eco-Bot use	No
Target	Adapted target for the duration of the project: 15% Note: This target was adapted. This adapted target refers to the energy savings to be achieved by the end of the pilot. 20% is the expected target after the end of the Eco-Bot project.	No

The target for Parameter 38 of 5 percent difference of energy saving between the groups of smart meter users and the group of users with conventional meters was achieved.

The assumption that due to Eco-Bot iESA users would enter more energy saving events into the iESA has not been achieved.

Table 6: SEnerCon pilot-specific parameters (P38-40)

SEnerCon-specific parameters (P38-40)		Target reached
P38 – Energy savings achieved by users with smart meters compared to users without smart meters Difference in energy savings achieved in percent	10 % savings of smart meter users compared to 0.01 % savings of users without smart meters (for the duration: April 2020 to May 21)	Yes
Target	5% more saving	
P39 - Increase of energy saving events entered into the iESA system Increase in percent related to Eco-Bot usage	0 %	No
Target	15 % increase	
P40 - Evaluation of Eco-Bot as channel of communication Percentage of users preferring Eco-Bot as communication channel	Not determined[8]	
Target	30% share of younger (< 45) and/or female users preferred Eco-Bot as a channel of communication over the other options	not determined

[8] Channel of communication (use female user) could unfortunately not be determined as the user survey where the question whether Eco-Bot is the preferred channel of communication was anonymous and the gender and age of the respondents was mistakenly not inquired. In general, in only one third of the regular Eco-Bot users of the German pilot were women. In comparison: The share of female user of all iESA users is 25 percent, so possibly female users were attracted to the dialogue-oriented concept of Eco-Bot.

EYPESA (B2C)

For the calculation of energy savings only the energy consumption data of regular users were considered, also only considering the months of the pilot period when they were registered and actively using Eco-Bot. These consumption data were then compared to the data of corresponding months of the previous year to each registration. (Example: If user Y participated in the pilot from June 2020 to May 2021, this period is compared with June 2019 to May 2020).

For some participants historical consumption data presented several hourly missing values. Therefore, the consumption difference between the pilot period and the baseline period was calculated excluding the time frames that presented NA values in one or both datasets. Consequently, for each participant the hourly mean of the consumption difference and the total consumption of the baseline year was calculated and considered as the new value for filling the holes in the datasets. Once the datasets were cleaned, the percentage of savings was calculated for each user by dividing the total savings for the total consumption of the baseline year.

However, especially depending on the outside temperature, consumption can vary from one year to another. It is therefore hard to understand till which extent Eco-Bot made an impact or not in the energy consumption of the users.

Furthermore, it also needs to be considered that in Spain, the lockdown measures caused by the Covid-19 pandemic lead to a higher occupation in most of the participants homes. Focusing on the case of Estabanell clients, the increase of electricity consumption for the residential sector was calculated to be around 9% for each client in the first month of lockdown (April 2020), with an average of 5,6% increase throughout all the year 2020.

Thus, even if the total savings of the Eco-Bot pilot participants are negative, it can be noticed that the number is below the average increase of the whole portfolio of residential clients, especially during full lockdown times.

Due to this peculiar situation provoked by the pandemic, it is hard to quantify the benefits related to the use of Eco-Bot, but it could be possible that Eco-Bot helped to mitigate the increase in electricity consumption.

To conclude it can be noticed that 5 of the participants were able to reach the target and more than a 15% increase in energy savings.

Table 7: Total increase of energy savings by participating user - EYPESA

P25 - Total increase of energy savings by participating user: Percentage increase of energy savings after Eco-Bot use		Target reached
-2,6 % total savings after Eco-Bot use		
Target	Adapted target for the duration of the project: 15%	No

P25 - Total increase of energy savings by participating user: Percentage increase of energy savings after Eco-Bot use		Target reached
	Note: This target was adapted. This adapted target refers to the energy savings to be achieved by the end of the pilot. 20% is the expected target after the end of the Eco-Bot project.	

P35 aims to understand if users with more detailed feedback (minute meter) are more engaged and saves more energy than basic ones with hourly granularity measurements.

Comparing the control group of non-Eco-Bot users with the Eco-Bot users, in P36, will allow to deduce which percentage of the energy savings must be related to external factors that affected all pilot participants across both groups (such as weather conditions and economic situation, or in this actual case the change of habits due to the pandemic).

Finally, as one of the objectives of Eco-Bot for EYPESA was to relieve the customer service, P37 aims to understand if the bot was able to not refer the users to the customer service in more than 50% of their Eco-Bot sessions.

Table 8: EYPESA pilot-specific parameters (P35-P37)

EYPESA-specific parameters (P35-37)		Target reached
P35 – Energy savings achieved by users with basic smart meters compared to users with advanced smart meters Difference in energy savings achieved in percent	No relevant difference was seen among the two groups. Basic users saved 1,2% more than advanced users, for such a small sample cannot be considered a significant result.	
Target	No specific target value is defined. The aim of this parameter is to provide empiric proof for the underlying hypothesis that users with advanced smart meters have more detailed insights in their energy consumption on an appliance level and thus achieve greater energy savings then the basic smart meter users.	n/a
P36 - Energy savings achieved by Eco-Bot users compared to the control group of non-Eco-Bot users Difference in energy savings achieved in percent	16 participants that filled the participation survey but never completed the registration or completed it but never logged were taken into consideration for this metric. Savings resulted in a -45% comparing the pilot period (from April 2020) to the previous year (from April 2019)	
Target	No specific target value was defined; however, the underlying hypothesis was that users of eco-bot could achieve greater savings than the control group.	n/a

EYPESA-specific parameters (P35-37)		Target reached
P37 – Self Service Rate **% of sessions and % of inquiries that the bot was able to go through without directing the user to the customer service office.**	The self-service rate is 99.6%. 0.2% of inquiries could not be matched to an intent, given two previous failed attempts. In this case, the chatbot replies with the contact details of the company so that a human can help. On top of that, we include in the calculation the percentage of cases (0.2%) that the user asked to talk to a human, assuming that the chatbot could not help.	
Target	More than 50%	**YES**

DEXMA (B2B)

As historical data was not available for the whole previous year in the pilots, the daily average savings were calculated for the pilot period for which there was a historical reference consumption. Once the daily savings per location were obtained, an average was calculated to obtain the pilot savings after the pilot launch. One restaurant (Restaurant 1) was not considered in the calculation of this parameter because it went out of business during the first months of the pilot.

Due to lockdown measures caused by the Covid-19 pandemic, most European countries in the first months of the pilot closed restaurants and hotels and they had restrictions during most of the duration of the pilot. This caused consumption savings greater in restaurants and hotels than in supermarkets, which in some cases even increased their consumption as supermarkets remained open during the lockdown because they were considered essential businesses and increased their operating hours.

Table 9: Total increase of energy savings by participating user - DEXMA

P25 - Total increase of energy savings by participating user: Percentage increase of energy savings after Eco-Bot use		Target reached
14% total savings after Eco-Bot use		
Energy savings in supermarkets	2% total savings after Eco-Bot use	No
Energy savings in restaurants	35% total savings after Eco-Bot use	**YES**

P25 - Total increase of energy savings by participating user: Percentage increase of energy savings after Eco-Bot use		Target reached
Energy savings in hotels	21% total savings after Eco-Bot use	YES
Target	Adapted target for the duration of the project: 15% Note: This target was adapted. This adapted target refers to the energy savings to be achieved by the end of the pilot. 20% is the expected target after the end of the Eco-Bot project.	YES

The pilot-specific parameter for DEXMA refers to the Net Promoter Score (NPS) and compares the user's NPS rating of DEXCell EM with and without Eco-Bot.[9] The first parameter is considered as achieved, as the average NPS result given by the pilot participants is greater (8.4 vs 6.8) than the average NPS given by DEXMA's partners, who receive a survey to evaluate DEXMA's EMS, DEXMA Analyse, formerly known as DexCell EM. On the other hand, 60% of the users that answered the final evaluation survey were promoters of DEXMA Analyse with Eco-Bot, that is, they gave a score of 9 or greater. The results are depicted in the table below.

Table 10: DEX pilot-specific parameters (P41)

P41 – NPS comparison of "DEXCell EM with Eco-Bot" and "DEXCell EM without Eco-Bot"		Target reached
Average NPS result at final user evaluation: 8.4 Average NPS result gathered internally from DEXMA's partners (April 2020 – March 2021): 6.8 50% of users are promoters		
Target	➔ More than 50% of users should give higher NPS rating for 'DEXCell EM with Eco-Bot' than for 'DEXCell EM without Eco-Bot' ➔ 50% are promoters of Eco-Bot (NPS 9 to 10)	YES

[9] The Facility Manager (user) is regularly asked to give the NPS rating via DexCell EM Service. At the end of the pilot the FM/user will be asked to give the NPS rating for "DexCell EM with Eco-Bot" through the following question in a user survey: How likely is it that you would recommend "DEXCell EM with Eco-Bot" to a friend or colleague? This rating is compared with his previous NPS rating of "DEXCell EM without Eco-Bot" to see the improvement or deterioration on the score.

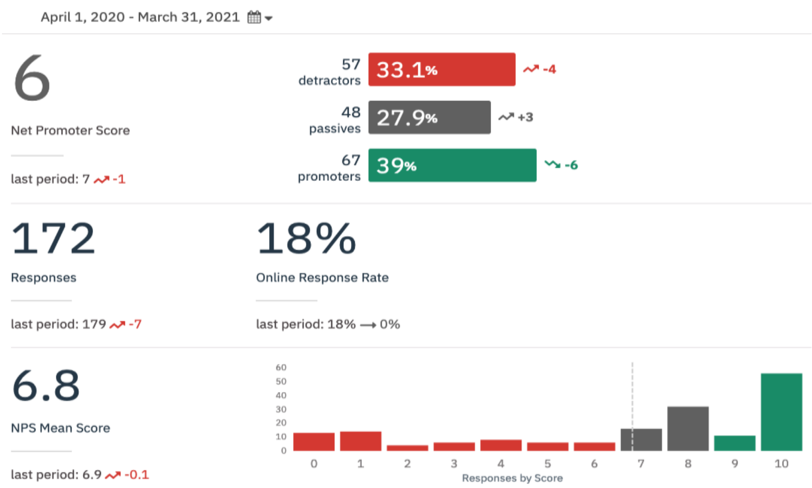

Figure 10: Average NPS score for DEXMA EMS (without Eco-Bot) according to replies from DEXMA's customers during the pilot phase[10]

The NPS score of DEXMA's EMS, DEXMA Analyse, is constantly being asked to DEXMA's customers. The average NPS score given during the pilot phase (April 2020 to March 2021) was 6.8 (figure 10). The average NPS score gathered in the final user satisfaction survey sent to Eco-Bot users was 8.4, notably higher than the previous NPS score. Even though there was no way to simultaneously ask the users to give an NPS score for DEXMA EMS without Eco-Bot and with Eco-Bot, comparing both averages it can be stated that users give a higher NPS score to DEXMA EMS with Eco-Bot than without Eco-bot, so the first of the goals in this parameter has been achieved. Furthermore, out of 4 respondents, 2 gave an NPS score of 9 or greater, which indicates that the second initial goal was achieved.

Users that made a change to save energy (behavioural change or investment)

The P26 metric assesses the percentage of users that actively changed their energy consumption patterns, either by behavioural changes or investments made, i.e. in new appliances. P26 accounts for both changes recommended to the users by Eco-Bot as well as other energy saving actions by the user after the starting to engage with Eco-Bot.

For the evaluation of the behavioural changes and investments only regular users were considered. As Eco-Bot's recommendations are sent automatically to all users, it is important to define users that had the chance to read them as basis for the evaluation. For the residential pilots SEnerCon and EYPESA, regular users were defined as users that used the bot at least three times and at least every 2 months except for the summer months June to September 2020.

[10] Information provided by Wootric: https://inmoment.com/wootric/

Table 11: Users that made a change to save energy (all pilots)

P26 - Users that made a change to save energy (behavioural change or investments)		Target reached
Target	More than 30% of users will have changed their behaviour towards energy efficiency in the selected measurable sample.	
Results achieved	**EYPESA: 11 of 15 regular users changed behaviour (73.3 %)** **SEnerCon: 24 of 30 regular users changed behaviour (80 %)** **Resulting in a total of 35 regular users out of 41 regular users changed behaviour (77.8 %)**	YES

Consumers making monetary investments to save energy

Parameter 27 indicates the share of Eco-Bot users who made a monetary investment that improved the energy efficiency of their at-home power consumption. This includes, for example, the purchase of more energy-efficient appliances. As this relates to only residential users, the evaluation concerns EYPESA and SEnerCon pilots only.

Table 12: Consumers making monetary investments to save energy (residential pilots)

P27 – Consumers making monetary investments to save energy		Target reached
Target	10% of users made energy efficiency investments	
Results	**SEnerCon: 16 out of 30 regular users made investments (53.3 %)** **EYPESA: 4 out of 15 regular users made investments (26.7 %)** **In total: 20 out of 45 regular users made investments to save energy (44.4 %)**	YES

Commercial buildings (facilities) that were affected by a change to save energy (behavioural change or investments)

Parameter 28 relates to the commercial buildings that underwent a change in their energy consumptions patterns, either due to investment or behavioural changes related to their energy consumption. Both changes recommended by Eco-Bot and adaptations undertaken independent for the chat bot are considered. This metric is assessed exclusively for the DEXMA pilot, as only DEXMA served commercial clients in their B2B model.

DEXMA (B2B)

18 buildings participated during the whole pilot period from April 2020 until March 2021 (12 months). 10 additional buildings (bank branches - deployment 6145: locations 211380 to 211390) participated for 1 month (March 2021)

The first stage evaluation involved DEXMA pilot findings for the period April 2020 – December 2020 (9 months).

- Total buildings: 18[11]

During the first stage evaluation, all buildings (18 out of 18) have implemented at least one energy saving change after the use of Eco-Bot. Thus, the result amounts to 100 %.

The second stage evaluation involved DEXMA pilot findings for the period: January 2021-March 2021 (3 months).

- Total buildings: 17 (plus 10 bank branches included only in the last month of the pilot)

During the second stage evaluation, all of the buildings (27 out of 27) have implemented at least one energy saving change after the use of Eco-Bot. Thus, the result amounts to 100 %.

The final evaluation, performed after the completion of the pilot, involves the whole pilot period: April 2020 – March 2021 (12 months)

- Total buildings: 18 (plus 10 bank branches included only in the last month of the pilot)

During the final evaluation, all of the buildings (18 out of 18) have implemented at least one energy saving change after the use of Eco-Bot. Thus, the result amounts to 100 %.

Implemented energy saving measures recommended by Eco-Bot

Based on the two metrics above, this parameter exclusively indicates the number of energy-saving measures that were recommended by Eco-Bot and later implemented by the respective user. This indicator is evaluated for all three pilots.

SEnerCon (B2B2C)

232 recommendations were implemented by regular users[12]

11 recommendations were implemented by non-regular users

[11] One of them – a restaurant – went out of business in August 2020
[12] Regular users were defined as users that used the bot at least three times and at least every two months except for the summer months June to September 2020

The 30 regular users of the SEnerCon pilot indicated that they implemented in total 232 recommendations suggested by Eco-Bot. Thus, on average each regular user implemented 7.4 recommendations. Following the logic of the definition and calculation of the parameter that 30 percent of the user implement at least 10 recommendations, 30 percent (10 in total) of the regular iESA users implemented 23 recommendations on average. Or, if the 10 recommendations are assumed to be the fixed number of recommendations, 74 percent of the regular Eco-Bot users implemented Eco-Bot's recommendations. Thus, the target of 30 percent was reached.

EYPESA (B2C)

The 15 regular users of EYPESA pilot implemented in total 52 recommendations. If the target was that 30% of the users implemented10 recommendations it means that 30% of 15 (5) should have implemented 10 recommendation each. This is a total number of 50. Since 52 recommendations were implemented, the target was reached.

DEXMA (B2B)

Total buildings: 18

- 18 buildings participated during the whole pilot period from April 2020 until March 2021 (12 months)
- 10 additional buildings (bank branches - deployment 6145: locations 211380 to 211390) participated for 1 month (March 2021)

Adapted target: Since the buildings participating were finally 18, using the assumptions of P29 in D3.3, the targeted implemented recommendations are 140.

First stage evaluation: The first stage evaluation involved DEXMA pilot findings for the period: April 2020 – December 2020 (9 months).

Total buildings: 18 (one of them, a restaurant, went out of business in August 2020)

During the first stage evaluation, we have measured **41** recommendations implemented in the pilot commercial buildings. It should be noted that the restaurant that went out of business in August 2020 had only performed 1 implemented recommendation out of the 41 total.

Second stage evaluation: The second stage evaluation involved DEXMA pilot findings for the period: January 2021-March 2021 (3 months).

Total buildings: 17 (plus 10 bank branches included only in the last month of the pilot)

During the second stage evaluation, we have measured **158** recommendations implemented in the pilot commercial buildings. It should be noted that these **158** implemented recommendations do not consider the 10 bank branches that participated for only 1 month.

If we also consider the implemented recommendations in the bank branches, we have an additional **80**, resulting to a total of **238** implemented recommendations.

Final evaluation: The final evaluation – performed after the completion of the pilot – involves the whole pilot period: April 2020 – March 2021 (12 months).

Total buildings: 18 (plus 10 bank branches included only in the last month of the pilot)

During the final evaluation, we have measured **199** recommendations implemented in the pilot commercial buildings, taking into account the whole demonstration period. It should be noted that these **199** implemented recommendations do not consider the 10 bank branches that participated for only 1 month. If we also consider the implemented recommendation in the bank branches, we have an additional **80** implemented recommendations, resulting to a total of **279** implemented recommendations.

4.3.2. Results of green-impact related parameters

The following parameters P30 to P33 evaluate the environmental impact of Eco-Bot, especially the avoidance of CO_2 emissions, as well as the economic gain (measured in € saved). The assess the energy and associated emissions saved, as well as the number of users that became interested in turning to renewable energy sources. As the pilots operate in different countries and target different use cases, each with distinct emission baselines, different coefficients of carbon intensity have to be applied. While EYPESA and SEnerCon used the country specific coefficient for residential users, DEX used a European average to accommodate its EU-wide customers.

Parameter 30 indicates the energy savings (in MWh) achieved by the Eco-Bot pilot participants during the period when they regularly used Eco-Bot. It is calculated by comparing the baseline consumption of the respective user with the consumption during the pilot phase of Eco-Bot. Parameter 31 indicates the amount of carbon dioxide emissions that were saved per user during the usage of Eco-Bot. This metric relates directly to the energy saving achieved, as it is calculated by multiplying the Parameter 30 metric by the carbon intensity coefficient of the electricity consumed, in the respective country. Parameter 32 indicates the number of users of Eco-Bot, who indicated in the user surveys that they were interested in either switching to renewable energy plan/provider or producing/supporting renewable energy themselves, for example by making financial investments in solar panels or wind farms. It is only relevant for the residential pilots. Parameter 33 describes the economic gain Eco-Bot's users where able to obtain, measured in € saved per household or facility. The prices for one kW/h are set individually for each pilot to account for price differences for energy in the pilot countries. The financial savings were calculated by multiplying the energy savings calculated for each household / facility with the respective energy prices.

SEnerCon (B2B2C)

For the SEnerCon pilot, regular usage meant such users logged in at least 3 times without long monthly gaps (except for the summer months June to August, when a lot of German users are on vacation and not using the iESA). Unfortunately, only 30 of the 92 registered users used the bot on a regular basis. Thus, only these “regular” users were considered for the calculation of the energy savings as their saving behaviour could be related to the usage of Eco-Bot. In total energy savings of around 2 percent have been achieved for electricity and around 1 percent for space heating. Thus, the targets of 15 percent electricity savings and 5 percent energy savings for space heating set in the DoW have both not been achieved. However, it has to be mentioned that due to the Covid-19 pandemic a lot of people worked at home office, children were taught in online lessons and in general the occupation at home was higher than in the baseline months before the pandemic leading to higher energy consumption. Evaluations of co2online on the development of energy consumption of iESA users (1300 users) during the Covid-19 pandemic have revealed that iESA users had on average a 5 percent increase on electricity consumption and a 2 percent increase on energy consumption for space heating. Thus, an effect of Eco-Bot on the energy behaviour of Eco-Bot pilot participants is most likely.

As mentioned in P25, the lockdown measures caused by the Covid-19 pandemic leaded to a higher occupation in most of the participants homes. In the case of iESA users, the average increase of electricity consumption in the residential sector was calculated to be around 5 % for electricity consumption and 2 % for energy consumption for space heating.

It is therefore understandable that the total amount of savings is negative.

However, considering singularly every participant, 4 of them were able to reach the target and save more than 3487 kWh/year*15% = 521 kWh/year.

As for the energy savings, also for the CO2 emissions avoided the Covid-19 pandemic and its impact on private households has resulted in increased energy consumption in most of the households. However Eco-Bot users showed low amounts of savings. Nevertheless, the targets have not been achieved.

Table 13: Green impact related parameters - SEnerCon

P30 – Overall energy savings achieved (in MWh)		Target reached
Result	**0,618 MWh electricity** **0,413 MWh energy for space heating**	
Target	90 MWh/year for electricity and 112.5 MWh/year for space heating	No
P31 - Average amount of avoided CO_2 emissions of each user (in kg)		
Result	**3,1 kg/user for electricity and 40,2 kg/user for space heating**	
Target	**84.4 kg/year per user for electricity and 150kg/year user for space heating**	No
P32 - Turn to sustainable energy: Number of users interested in turning to renewable/sustainable energy		
Result	**80% of the participants of the German pilot were interested in investing in renewable energy sources.**	n/a
P33 - Money saved per household		
Result	**18,20 € per household** **0,50 € for electricity and 17,30€ for space heating**	
Target	SEnerCon: 180 €/year per household for electricity (and 45 €/year per household for space heating *Calculation for electricity:* ➔ *90 000 kWh/year (overall energy savings achieved) / 150 (participating households) = 600 kWh/year per household* ➔ *600 kWh (energy savings per household) x 0,3 € (average price of one kWh in Germany) = 180 €/year per household* *Calculation for space heating:* ➔ *112 500 MWh/year for space heating/ 150 (participating households) = 750 kWh/year per participating household* *750 kWh (energy savings per household) x 0,06 € (average price of one kWh in Germany for gas) = 45 €/year per household*	No
Comments	The money saved result from the energy savings. For the energy cost, the real costs for each participant were considered, which are available in the participant's iESA accounts.	

EYPESA (B2C)

As mentioned in P25, the lockdown measures caused by the Covid-19 pandemic lead to a higher occupation in most of the participants homes. In the case of Estabanell clients, the increase of electricity consumption in the residential sector was calculated to be around 9% for each client in the first month of lockdown (April 2020), and stabilized to an increase of 2,3% in the subsequent months of partial lockdown. It is therefore understandable that the total amount of savings is negative[13]. However, considering singularly every participant, 4 of them were able to reach the target and save more than 3487 kWh/year*15% = 521 kWh/year. As, most likely due to pandemic, the energy savings resulted negative, also the CO2 and economic savings follow the same trend.

Table 14: Green impact related parameters - EYPESA

P30 – Overall energy savings achieved (in MWh)		Target reached
Result	**- 1,58 MWh saved throughout the pilot phase**	
Target	**Adapted target (to participation months and number of users):** **4,67 MWh/year** 15 users (instead of 66) and an average duration of 7,5 months (instead of 12) per pilot participant were considered. Hence, the adapted target for the actual pilot should be: 3487 kWh/year*15 users*15%*(7,5/12) = 4674 kWh/year → 4,67 MWh/year	No
P31 - Average amount of avoided CO_2 emissions of each user (in kg)		
Result	**-35,96 kg/user**	
Target	**178,3 kg/year per user for electricity**	No
P32 - Turn to sustainable energy: Number of users interested in turning to renewable/sustainable energy		
Result	**According with the answers received in the final survey, 81% of the participants defined themselves interested in investing money into renewable energy sources such as self-consumption or by participating in crowdfunding projects.**	n/a

[13] The total amount of energy savings calculation was based on 15 users that are considered regular users during the pilot phase. However, not all the 15 users participated throughout the 12 months of the demo phase and the result was calculated based on the number of participation months for each user. This, plus the previously explained situation of the increase of activities at household level due to the lockdown, lead to a total negative amount of savings of – 1,58 MWh.

P30 – Overall energy savings achieved (in MWh)		Target reached
P33 - Money saved per household		
Result	**- 15,80 € per household on average**	
Target	**EYPESA: 78,45 €/year per household** *Calculation for electricity savings:* ➔ *34 521 kWh/year (overall energy savings achieved) / 66 (participating households) = approx. 523 kWh/year per household* *523 kWh (energy savings per household) x 0,15 € (average price of one kWh in Spain) = 78, 45 €/year per household*	No
Comments	As, most likely due to pandemic, the energy savings resulted negative, also the economic savings follow the same trend.	

DEXMA (B2B)

Parameter 30 quantifies the amount of energy (in MWh) that the facilities saved throughout the Eco-Bot pilot. All results are illustrated in table 13, showing the overall savings and the savings per facility (i.e. supermarkets, restaurants and hotels). The hotel achieved more absolute savings, as it was greatly affected by the Covid-19 lockdown and ensuing travel restrictions, which caused a great reduction in its electrical consumption. On the other hand, supermarkets achieved savings in some cases, but they represented a small fraction of their consumption, because they remained open during the Covid-19 lockdown in Italy (in some cases even opening more hours). Restaurants reduced their consumption due to their closure during more than 2 months and their indoor capacity was restricted since. But their electrical consumption is smaller than supermarkets and hotels, so the absolute energy savings are also smaller.

The savings presented in table 15 were obtained by directly comparing the available historical consumption from the year prior to the start of the pilot phase, as this was how the parameter was defined in D3.3. However, an alternative calculation is presented after the table. DEXMA's EMS has a tool to automatically generate baselines, either introducing the variables or using a formula to calculate the baseline. This tool is called Automatic Baseline Calculator (ABC) and its results are displayed as M&V projects, where the savings are compared to a goal. In the case of the commercial buildings of the pilot, meteorological data from nearby meteorological stations was available, so heating and cooling degree days (HDD and CDD) were considered in the baseline calculation, as well as the historical main consumption available and the day of the week.

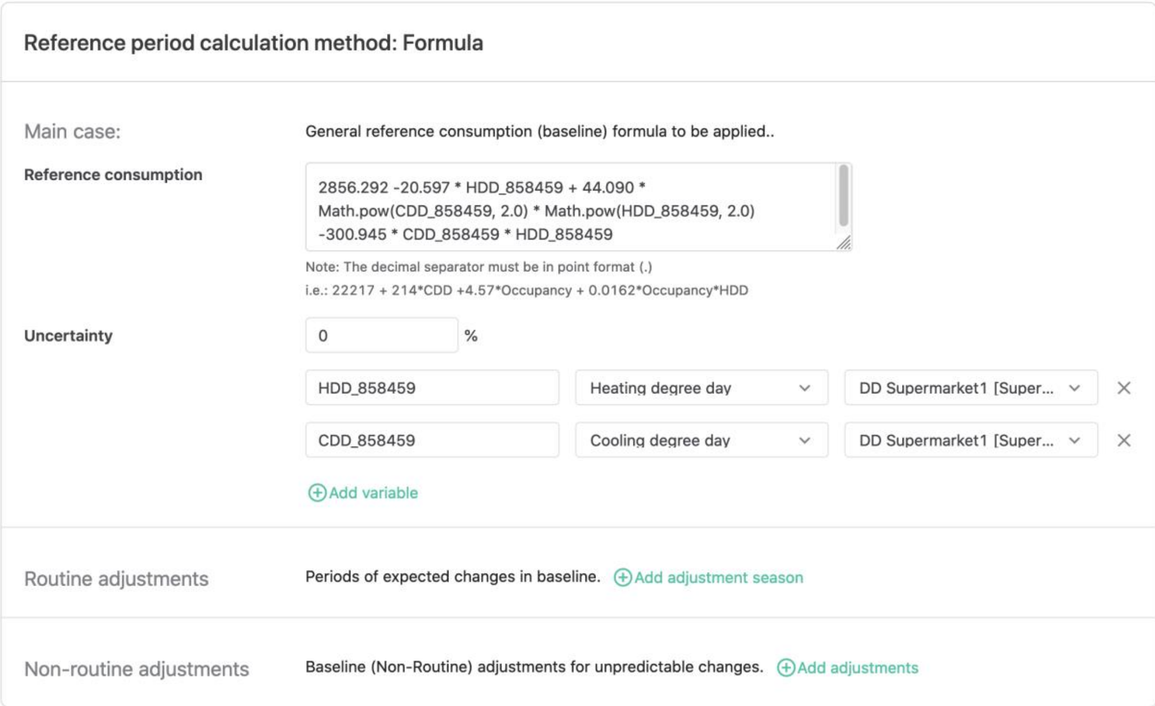

Figure 11: Baseline parameters in a M&V project in DEXMA Analyse

In the figure above, the automatically generated formula is displayed – it can be edited, variables can be added, and other adjustments can be made. A baseline was created for all the locations with available historical and meteorological data and the savings were calculated as the difference between the baseline and the real consumption.

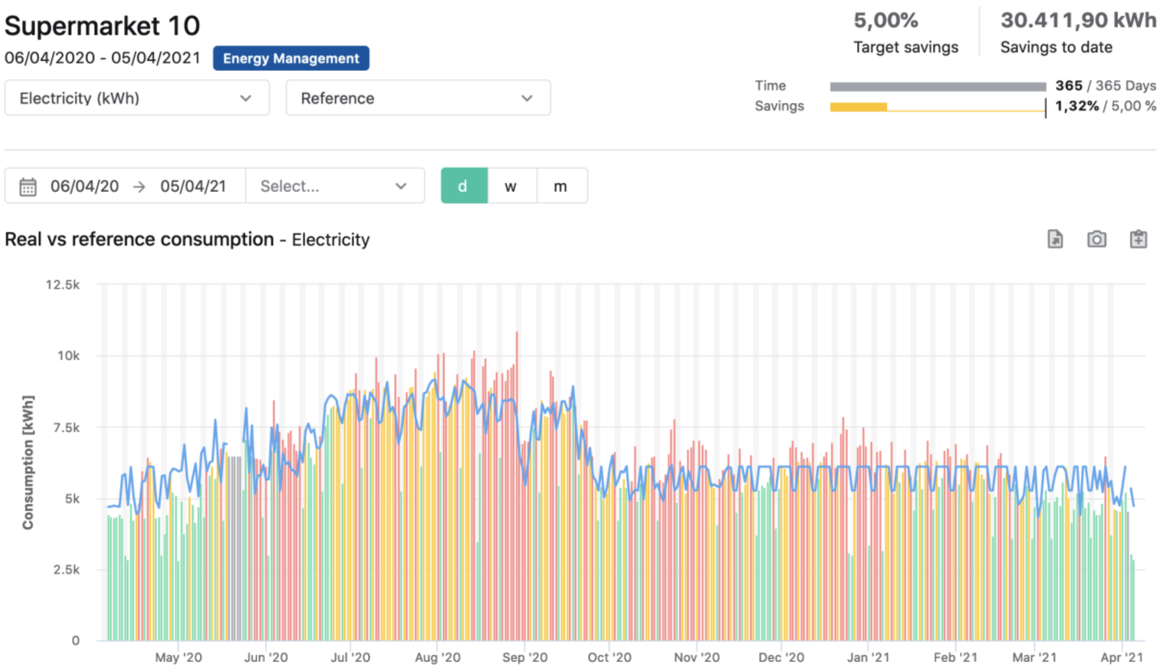

Figure 12: Baseline (blue line) compared to real consumption for Supermarket 10

Table 15: Overall energy savings achieved - DEXMA

P30 – Overall energy savings achieved (in MWh) **Definition: Amount of energy that users saved throughout the pilot phase**		**Target reached**
Overall energy savings	792 MWh saved throughout the pilot phase	**YES**
Energy savings in supermarkets	244 MWh saved throughout the pilot phase	-
Energy savings in restaurants	122 MWh saved throughout the pilot phase	-
Energy savings in hotels	426 MWh saved throughout the pilot phase	-
Target	DEXMA: 1500 MWh/year (assuming 20 participating facilities) *Calculation and assumptions for energy saving:* → *each participating facility will achieve 7500 kWh of energy savings (18*7500 kWh= 135MWh/year)* → *(Assuming that 37500 kWh/year is the annual energy consumption for an average building.)*	

P30 – Overall energy savings achieved (in MWh) Definition: Amount of energy that users saved throughout the pilot phase		Target reached
	Note: This target was adapted in accordance with a realistic scenario applicable to the scope of the pilot. Taking into account the actual number of Eco-Bot users in the pilot (with the DEXMA pilot of 7 Facility managers and around 20 participating facilities). Furthermore, contrary to the projected impact that the DoW indicates, the adapted target will measure the actual achieved energy savings in the course of the pilot.	

According to the savings calculation using the ABC tool in DEXMA Analyse, the summary of savings for the restaurants is the following:

Name	Type	Target	Saved %	Saved	
Supermarket 1	Energy Management	15,00%	-3,13	-29.850,21 kWh	365 / 365 Days -3,13% / 15,00 %
Supermarket 2	Energy Management	5,00%	0,00	-1,48 kWh	365 / 365 Days 0,00% / 5,00 %
Supermarket 4	Energy Management	5,00%	-0,03	-610,56 kWh	365 / 365 Days -0,03% / 5,00 %
Supermarket 6	Energy Management	5,00%	0,04	1.094,50 kWh	365 / 365 Days 0,04% / 5,00 %
Supermarket 7	Energy Management	5,00%	0,21	5.973,83 kWh	365 / 365 Days 0,21% / 5,00 %
Supermarket 8	Energy Management	5,00%	0,87	24.241,26 kWh	365 / 365 Days 0,87% / 5,00 %
Supermarket 9	Energy Management	5,00%	-0,14	-2.004,26 kWh	365 / 365 Days -0,14% / 5,00 %
Supermarket 10	Energy Management	5,00%	1,32	30.411,90 kWh	365 / 365 Days 1,32% / 5,00 %
Supermarket 5	Energy Management	5,00%	0,00	-2,28 kWh	320 / 320 Days 0,00% / 5,00 %
Supermarket 3	Energy Management	5,00%	0,24	1.120,45 kWh	315 / 315 Days 0,24% / 5,00 %

Figure 13: Savings of supermarkets during the pilot phase calculated by comparing with the baseline generated with ABC

The total savings in supermarkets is of 30372 kWh or 30.372 MWh, a bit more than 10% of the savings presented in Table 81 (244 MWh). This amount is more logical because it is difficult for supermarkets in Italy to have saved so much energy in a year where they faced a bigger influx of customers and much more workload. It must be noted that some supermarkets did not have all 12 months of consumption data and had some periods of missing data (supermarkets 5 and 3).

With regards to the restaurants, the savings calculated with the ABC-created baseline are displayed below:

Name	Type	Target	Saved %	Saved	
Restaurant 2	Energy Management	5,00%	10,21	3.799,18 kWh	**365** / 365 Days **10,21%** / 5,00 %
Restaurant 3	Energy Management	5,00%	1,88	360,44 kWh	**365** / 365 Days **1,88%** / 5,00 %
Restaurant 4	Energy Management	5,00%	2,97	1.478,17 kWh	**365** / 365 Days **2,97%** / 5,00 %
Restaurant 5	Energy Management	5,00%	11,22	12.795,04 kWh	**365** / 365 Days **11,22%** / 5,00 %
Restaurant 6	Energy Management	5,00%	3,48	1.548,43 kWh	**365** / 365 Days **3,48%** / 5,00 %
Restaurant 7	Energy Management	5,00%	14,15	9.252,48 kWh	**365** / 365 Days **14,15%** / 5,00 %
Restaurant 1	Energy Management	5,00%	14,56	993,63 kWh	**143** / 143 Days **14,56%** / 5,00 %

Figure 14: Savings of restaurants during the pilot phase calculated by comparing with the baseline generated with ABC

In the case of restaurants, the total savings achieved (using this methodology) was around 30225 kWh or 30.225 MWh. This is almost the same as the supermarkets, which have more locations and have a much higher consumption than restaurants. About 25% of the savings were obtained with the initial methodology, which did not consider meteorological data, for instance. However, in comparison with supermarkets the decrease is much lower because restaurants in Spain were affected by restrictions during the whole pilot phase, so it seems logical to save energy under these circumstances. It must be noted that a restaurant closed down in August 2020 and could not complete the pilot phase, so the baseline only compared the consumption for 43 days.

The hotel managed to save the following amount of energy with this methodology:

Name	Type	Target	Saved %	Saved	
Hotel 1	Energy Management	5,00%	1,08	17.050,76 kWh	**365** / 365 Days **1,08%** / 5,00 %

Figure 15: Savings of the hotel during the pilot phase calculated by comparing with the baseline generated with ABC

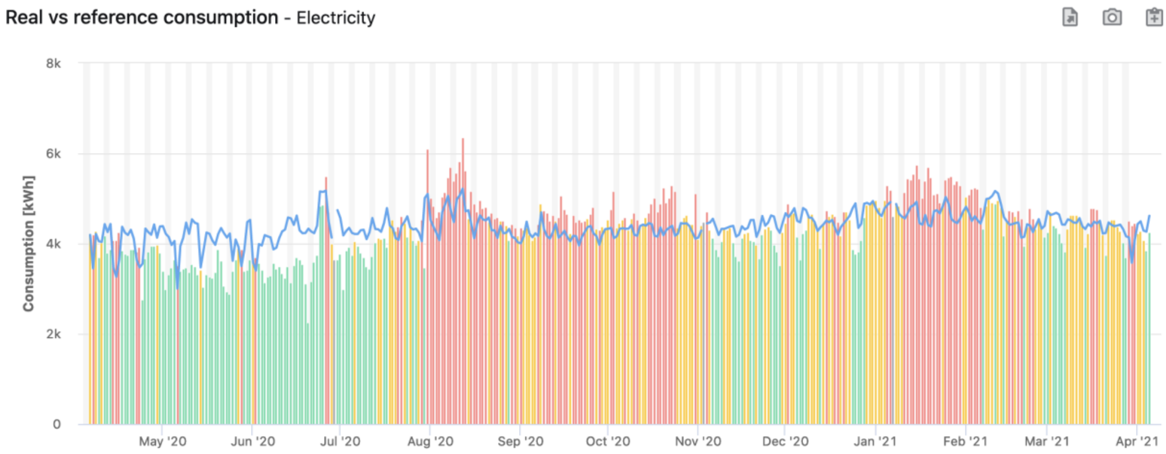

Figure 16: Baseline (blue line) compared to real consumption for Hotel 1

In the case of the hotel, savings were achieved during the period from April to July 2020, when the heaviest restrictions were applied, and travel was very restricted. During the rest of the pilot, some periods had greater consumption than the baseline. This resulted in savings of 17050 kWh or 17.05 MWh, less than 5% of the savings obtained by simply comparing historical and pilot phase consumptions. This difference can have several causes: differences in meteorological data from one to another or a period from April 2019 to March 2020 with more activity than usual, for instance.

In summary, the savings achieved with this method amount to approx. 77.647 MWh, slightly more than 60% of the goal set (135 MWh). This amount represents 0,46% total savings (Parameter 25). This is a considerable amount considering that fewer retrofits or energy-saving actions could be implemented in the pilot sites compared to other periods (not affected by the Covid-19 pandemic).

Table 14 shows the results of parameter 31 and describes how much CO_2 emissions have been reduced by each user (i.e. facility) through the use of Eco-Bot. The overall energy savings were divided by 18 locations. The CO_2 avoided is directly related to the electrical savings obtained in P30.

Table 16: Average amount of avoided CO_2 emissions of each user - DEXMA

P31 - Average amount of avoided CO_2 emissions of each user (in kg)		Target reached
13023 kg CO_2 avoided per location		
Energy savings in supermarkets	7217 kg CO_2 avoided per supermarket	YES

P31 - Average amount of avoided CO_2 emissions of each user (in kg)		Target reached
Energy savings in restaurants	5153 kg CO_2 avoided per restaurant	YES
Energy savings in hotels	126175 kg CO_2 avoided per hotel	YES
Target	*CO_2 per kWh coefficient = 0.296 kg CO_2 / kWh* Created target pilot based on energy savings achieved: DEX: kg/kWh = 2 220 CO2 kg/year per participating facility	YES

Considering the savings obtained using the alternative methodology (using the ABC and M&V functionalities in DEXMA Analyse) and the same CO_2 per kWh coefficient (0.296), the results are the following:

- Supermarkets: 899 kg CO_2 saved per supermarket on average
- Restaurants: 1278 kg CO_2 saved per supermarket on average
- Hotel: 5046.8 kg CO_2 saved by the hotel

Table 16 shows the results of economic parameter 33, i.e. how much money has been saved per participating facility by the use of Eco-Bot. The target of 1.125 €/year per facility was overachieved. The overall energy savings were divided by 18 locations (i.e. supermarkets, restaurants and hotels). Even though one restaurant had to close, and no savings were registered. The economic savings are directly related to the electrical savings.

Table 17: Economic parameter - DEXMA

Economic Parameter (P33)		Target reached
P33 – Amount of money saved per facility	5.280 € saved per location / year	
Money saved in supermarkets	2.926 €saved per supermarket / year	YES
Money saved in restaurants	2.089 € saved per location/ year	YES
Money saved in hotels	51.152 € saved per location/ year	YES

Economic Parameter (P33)		Target reached
Target	DEX: 1125 €/year per facility	YES

Considering the savings obtained using the alternative methodology (using the ABC and M&V functionalities in DEXMA Analyse), and the same price (0.12 €/kWh) the results are the following:

- Supermarkets: 364€ saved per supermarket on average
- Restaurants: 518€ saved per supermarket on average
- Hotel: 2046€ saved by the hotel

4.3.3. Results of rebound effect related parameter

This subchapter addresses the assessment of the sensibilization of users for the rebound effect, considering both direct and indirect rebound effects. While measuring the magnitude of the rebound effect for three different pilots exceeds the project's scope, raising user awareness of it is a realistic goal for Eco-Bot. Through the final user survey sent in June 2021, participants of SEnerCon's and EYPESA's pilot were asked to judge the usefulness of the information provided by Eco-Bot.

SEnerCon (B2B2C)

According to the final survey, 50% of the users that received information on the rebound effect found it useful. Thus, the target of 50% was reached.

EYPESA (B2C)

According to the final survey, 50% of the users that received information on the rebound effect found it useful. Therefore, the target (>50% of users who received information on the rebound effect found it useful) was reached**.**

5. Insights into qualitative results

This section presents an explanation of the Chatbot Evaluation Metrics and gives insights into the results. The Chatbot Evaluation Metrics relates to the user experience, the user engagement and the chatbot usability.

A detailed analysis of all findings can be retrieved from Annex B.

The results of the user experience survey show that the user satisfaction for most of the aspects of the Eco-Bot was above average with a slight increase in the second survey at the end of the pilots. This can be attributed to the improvements of the bot's intelligence and the user interface. By the end of the pilots most of the survey respondents indicated that they would recommend Eco-Bot to a colleague or a friend.

The findings on user engagement and user retention for the residential pilots[14] show the challenge to reach a consistent number of regular users, i.e. participants that use Eco-Bot regularly over the whole pilot period. Moreover, given that the pilots actively performed recruitment and engagement activities over the whole pilot period, the number of participants was changing over time and there was no constant number of users throughout the pilot. Therefore, the project team performed a cohort analysis by grouping users into cohorts based on the month they started using Eco-Bot.

The conclusions that can be drawn from the results of Chatbot usability related parameters show, that all measures implemented to improve Eco-Bot's performance during the pilots were efficient and remain valuable for future use of the bot.

[14] See "Results of Engagement and Retention Related Parameters" in Annex B.

6. The evaluation of the Eco-Bot performance: Key takeaways

The evaluation of the Eco-Bot performance considers the **five Development Assistance Committee (DAC) criteria** defined by the Organisation for Economic Co-operation and Development (OECD). **Relevance** (Eco-Bot's usefulness); **Effectiveness** (achievement of the objectives of the project); **Efficiency** (onetime achievement of the objectives); **Impact** (Eco-Bot's impact on energy savings behaviour of its users and achievement of the related energy saving targets); **Sustainability** (benefits of the use of Eco-Bot regarding energy efficient behaviour and to what extent the benefits can continue after the end of the project).

These criteria are reflected in the following **12 key takeaways**.

Key takeaway 1

On average, all Eco-Bot users agree (3,86 from the 5-point Likert Scale) that a) **Eco-Bot is a useful energy efficiency tool** and b) that it offers diverse and / or enhanced functionalities compared to alternatives (e.g. websites or search engine) and c) that it covers use cases they are interested in.

Key takeaway 2

The novel **AI-based NILM module** developed specifically for Eco-Bot is relevant towards achieving energy efficient behaviour and is based not on estimates, but the household's actual smart meter aggregate readings. It did not require manual intervention, but instead worked seamlessly with the backend and pilots to provide timely disaggregated results to multiple buildings every day while still maintaining privacy (the module had no knowledge of the building owners). This goes way beyond the state-of-the-art. The implementation is sustainable long after the project ends. Despite Covid-19 lockdowns, the **NILM module** demonstrated that it was consistently **effective in achieving the objectives** of the project, by providing daily appliance level energy consumption that is reliable and accurate.

Key takeaway 3:

Eco-Bot positively influences the energy consumption of users compared to the group of non-Eco-Bot users (EYPESA customers and iESA users). **Eco-Bot users** showed a reduction or a **lower increase in energy consumption**.

Key takeaway 4:

Eco-Bot motivated both residential and commercial users **to improve their energy consumption behaviour** both by implementing the Eco-Bot recommendations and making use of the suggestions to invest in more energy-efficient appliances, throughout the whole demonstration period.

Key takeaway 5:

The **application of smart meters is advantageous for the Eco-Bot** usage to profit of its full range of features and to exploit its full potential to save energy.

Key takeaway 6:

The introduction of an energy service chatbot like Eco-Bot needs **support from customer-relation staff** to guide users at first use. **Personal onboarding** of new participants in the second half of the pilot avoided user dropouts. This is especially true for smart meter users, where the smart meter had to be connected to the iESA and energy data had to be transferred prior to registering to Eco-Bot.

Key takeaway 7:

The **Covid-19 pandemic affected the energy consumption** of the users in all pilots, while the targets for the commercial pilot (Dexma) were reached. In spite of facing a shift of priorities in the target group of users of the commercial pilot (energy managers), savings were achieved even by supermarkets, which had more workload during the pilot period.

Key takeaway 8:

The evaluation of **the conversational intelligence of Eco-Bot increased** across the pilot phase. Visual elements in the chatbot helped users understand Eco-Bot's features.

Key takeaway 9:

The analysis of the results on the use of the conversational agents allowed the **constant monitoring of the pilot's needs** and offered valuable insight on the offered features; revealing the need to train/refine the bot, to recruit and guide users on how interact with the bot. The mechanisms implemented for these purposes allowed the **improvement of the performance during the pilots** and remain valuable for future use.

Key takeaway 10:

The behavioural model developed for the project – segmenting consumers according to their motivation to save energy – is **efficient**. It is of particular importance for the current and future consumer behaviour towards sustainable energy management at home and in enterprises. The classification model provides fully personalised recommendations. It serves as an **effective and universal tool** for identifying types of consumers, their attitudes and motivation to save energy. For this reason, the model and recommendations can be used for **educational purposes**, also **after the end of the project**. The **recommendations** addressed to the Eco-Bot users were assessed as **useful and relevant** (also in the context of the Covid-19 pandemic), and **well adapted to the needs** of individual and commercial users. Its educational value may **impact the change of behaviour** into more **ecological and sustainable** now and in the future, as results of the model evaluation indicated.

Key takeaway 11:

When **integrated in an EMS**, **Eco-Bot increases its value** and makes more users become promoters of the tool. NPS responses by Eco-Bot users improved the score (P41) that DEXMA obtained from its customers during the pilot phase.

Key takeaway 12:

Users were interested in **additional functionalities** (that originally were not designed), that were **constantly implemented and improved the bot**. Some of these features suggested by the Eco-Bot users are: comparison of different periods and highest/lowest comparison days or hours, itemized billing information and goals progress monitoring, as well as additional settings and alerts for facility managers. This shows the **importance of the feedback mechanisms** and the need to focus on keeping the **communication channels** alive and further increasing customer engagement.

7. Conclusion

D5.5 performed the evaluation of the Eco-Bot system during the pilot phase (April 2020 to May 2021) based on the five matrices (and consisting of 41 parameters) identified in D3.3.

The effectiveness of Eco-Bot – as an energy saving tool in the business and residential sectors – was positively verified. The results show that the recommendations for both individual and commercial energy meet the expectations regarding their usability for users. For households, 88.42% of the overall recommendations were found to be useful, while for commercial energy consumers the overall P1 value was around 75%. Such a result can be considered as very good and requires only minor changes in the prepared sets of recommendations. However, an assessment was also made in individual segments and therefore the necessary corrections, changes and new recommendations were added. Moreover, by carrying out the remodelling procedure, the accuracy of the classification model was assessed, where the model error was 15% and did not exceed the 20% threshold set in D3.3. Therefore, it can be concluded that the behavioural model developed for Eco-Bot works well and meets the project assumptions.

The newly created NILM module provides a detailed and informed personalised feedback to each household about the daily energy consumption for large white goods, namely refrigeration devices, washing machine, dishwasher and tumble dryer (SEnerCon and EYPESA). For commercial buildings, it provides feedback on daily consumption of refrigeration, HVAC and lighting, from the aggregate smart meter readings (DEXMA). The NILM module was able to seamlessly return NILM data to all users of all pilots every day for all demonstrations meeting the target. The deliverable also reports that the targets for both phases of demonstration were met (80% for supermarkets and restaurants). An average consumption accuracy of 65% was obtained over both evaluation periods. With regard to the AI algorithm's learning performance, or classification accuracy with regard to appliance signatures from user feedbacks the project team obtained an average classification accuracy of 72% with smart meter recordings of 10 seconds, 1 minute and 1-hour resolutions. This is in line with performance obtained during the small-scale validation period. The findings show that human error in recording appliance time of use is always an issue – albeit small – since we cannot always remember exact day or time an appliance was run. Hence, this needs to be cross-referenced with the smart meter recordings when calculating P6 metric.

Compared to a study by co2online[15], German Eco-Bot users achieved energy savings of 1.9% for electricity and 0.6% for heating energy. Consequently, cost savings and CO2 reductions were achieved, while the targets of 15% electricity savings and 5% heating energy savings could not be reached (SenerCon pilot). The electricity savings achieved by the smart meter user group were higher (10%), most likely because they benefited from more

[15] A study of co2online based on iESA user data revealed that electricity consumption increased by 5 % and energy consumption for space heating by 2 %.

interesting features of Eco-Bot. This group was rather small as smart meters could not be installed (due to Covid-19 pandemic). In addition, the interest of iESA users in smart meter installation was rather low (due to costs for the smart metering service after project lifetime and/or costs for the re-installation of a conventional meter).

For the EYPESA pilot, the results show that Eco-Bot users have an increase of 2.6% in consumption and the control group has a 45% increase among the pilot phase and the previous year. Moreover, the average increase in consumption of all the residential client portfolio of Estabanell (therefore without considering a specific area and type of clients) was of 5.6% throughout the year 2020. Part of this mitigation in the increased consumption can be attributed to Eco-Bot. The commercial pilot (Dexma) achieved the savings goals set in D3.3 (despite the unusual operating hours of the recruited locations – supermarkets, restaurants and one hotel). The consumption in restaurants decreased by 41% (122 MWh) due to the closure and limited capacity during the pilot phase caused by Covid-19-related restrictions. The hotel, in a period of tight restriction on travel, decreased its consumption by 21% (426 MWh). Surprisingly, supermarkets also decreased their consumption by 2% (244 MWh) compared to the same period of the previous year. The alternative methodology presented to calculate energy savings gave much lower energy saving results, as it considered meteorological and other variables. The total savings achieved with this method are around 77 MWh.

The usage of Eco-Bot by energy managers was consistent along the pilot phase period, despite the fact that many ESCOs had to shift their priorities to customer retention, except for the summer period, where there was a logical decrease in Eco-Bot usage. Towards the end of the pilot there was an increase of the usage due to the additional recruitment of 10 bank branches.

The users' evaluation of the conversational intelligence of the bot increased across the pilot phase (user survey). This is due to the continuous work in updating the bot and analysing the wrong intents, in order to deliver a better user experience. The user satisfaction for Eco-Bot was slightly above average and it increased from the first to the second survey. This can be attributed to various aspects: to the improvement of the bot's intelligence, to the user interface, and to the improved support of users while onboarding during the second period of the pilot. In addition, the chatbot's learning curve may also explain the improved user experience in the second half of the pilots.

One of the main goals of the Eco-Bot project was to improve the users' consumption behaviour by assisting them to make changes in their energy consumption behaviour or by motivating them to make investments in more energy-efficient appliances. The percentage of Eco-Bot users, who actively improved their energy consumption patterns exceeded the initial expectations. Additionally, the results of the parameter "percentage of Eco-Bot users who made a monetary investment that improved the energy efficiency of their at-home power consumption" also exceeded the set target. Parameter measuring "the percentage of commercial buildings that were positively affected by an energy saving change recommended by Eco-Bot" reached 100% in both evaluation periods – showing that facility managers made the most out of Eco-Bot's suggestions. Moreover, in terms of absolute numbers, the results are also promising, as the assumption that 30% of the users would implement at least 10 recommendations was validated for all three pilots.

All measures implemented to improve Eco-Bot's performance during the pilots were efficient and remain valuable for future use of the bot.

Table 18: User Experience Survey

1. Efficiency:
• Is it possible to use Eco-Bot fast and efficiently? 0

	1	2	3	4	5
slow (1) / fast (5)	○	○	○	○	○
inefficient (1) / efficient (5)	○	○	○	○	○
impractical (1) / practical (5)	○	○	○	○	○

2. Perspicuity:
• Is it easy to get familiar with Eco-Bot? Is it easy to learn how to use it? 0

	1	2	3	4	5
confusing (1) / clear (5)	○	○	○	○	○
difficult to learn (1) / easy to learn (5)	○	○	○	○	○
complicated (1) / easy (5)	○	○	○	○	○

3. Dependability:
• Do you feel that Eco-Bot is reliable and trustworthy? Is the interaction secure and predictable? 0

	1	2	3	4	5
unpredictable (1) / predictable (5)	○	○	○	○	○
not secure (1) / secure(5)	○	○	○	○	○
unreliable (1) / reliable (5)	○	○	○	○	○

4. Stimulation:
• Is it interesting and exciting to use Eco-Bot? Do you feel motivated to use it further? 0

	1	2	3	4	5
boring (1) / exiting (5)	○	○	○	○	○
not interesting (1) / interesting (5)	○	○	○	○	○
demotivating (1) / motivating (5)	○	○	○	○	○

5. Novelty:
• Is Eco-Bot creative and inventive? 0

	1	2	3	4	5
dull (1) / creative (5)	○	○	○	○	○
uninventive (1) / inventive (5)	○	○	○	○	○

6. Chat Interface:
• How do you rate Eco-Bot's visual look? 0

	1	2	3	4	5
very poor (1) / very good (5)	○	○	○	○	○

7. Chat Interface:
• Does the user interface looks organised? 0

	1	2	3	4	5
cluttered (1) / organised (5)	○	○	○	○	○

8. Usefulness: 0

	Strongly Disagree	Somewhat Disagree	Neither Agree nor Disagree	Somewhat Agree	Strongly Agree
A chatbot in the energy efficiency domain is useful	○	○	○	○	○
Eco-Bot outperforms its existing alternatives (website, app, customer service line, search engines) by offering diverse and/or enhanced functionalities	○	○	○	○	○
Eco-Bot covers the use cases I am interested in	○	○	○	○	○

9. Usefulness:
• Which are the Eco-Bot features you are more interested in: 0

- ◯ Energy saving recommendations
- ◯ Information on total energy consumption
- ◯ Information on appliance level consumption
- ◯ Comparison of consumption for different periods
- ◯ Customer-service information (contract, guidelines on how to change contact details etc.)
- ◯ Other (please specify) ...
- ◯ Monitoring of energy efficiency goals
- ◯ High/low consumption days
- ◯ Alerts on high consumption
- ◯ Monitoring of energy saving events

10. Usefulness:
• What other functionalities would you like Eco-Bot to offer? 0

11. Conversational Intelligence: 0

	Strongly Disagree	Somewhat Disagree	Neither Agree nor Disagree	Somewhat Agree	Strongly Agree
Eco-Bot understands the input text	○	○	○	○	○
Eco-Bot interprets commands accurately	○	○	○	○	○
Eco-Bot executes requested tasks	○	○	○	○	○
Eco-Bot is able to retain conversational context (maintain themed discussion) and follow up on a query	○	○	○	○	○

12. Overall Satisfaction:
• How would you rate your overall satisfaction with Eco-Bot 0

1 Very poor	2	3	4	5 Very good
★	★	★	★	★

13. Future Use:
• How likely is it that you will use Eco-Bot in the future? 0

	1	2	3	4	5
Not at all likely (1) / Extremely likely (5)	○	○	○	○	○

14. Net Promoter Scale (NPS):
• How likely is it that you would recommend Eco-Bot to a friend or colleague? 0

	0	1	2	3	4	5	6	7	8	9	10
Not at all likely (0) / Extremely likely (10)	○	○	○	○	○	○	○	○	○	○	○

ANNEX B: Detailed analysis of Metric: Chatbot Evaluation

The metrics presented in this section focus on the evaluation of the chatbot itself. It assesses the quality of the product the Eco-Bot consortium has developed during the project phase. The first parameters relate to the user experience of Eco-Bot, followed by evaluation parameters depicting the user engagement and retention and the assessment of the chatbot usability parameters.

Results of the User experience related parameters

These parameters refer to the experience users had with the use of Eco-Bot within the pilot phase. It covers pragmatic and hedonic quality aspects of the chat bot. Users were asked to subjectively evaluate Eco-Bot based on their personal experience. The parameter was evaluated on the basis of a user experience questionnaire (see Annex A). In this questionnaire users were asked to give feedback on perceived quality in parameters related to: efficiency, perspicuity, dependability, stimulation, novelty, chat interface, conversational intelligence, and generic user experience of Eco-Bot.

The results of the survey responses are evaluated according to a 5-point Likert Scale. This means that all results indicate the average value of the users' score (from 1 to 5) for each question.

SEnerCon (B2B2C)

SEnerCon's B2B2C pilot sent the questionnaire to all registered users (anonymously) in the end of January 2021 and collected answers until February 2021. This is part of the first phase evaluation. For the end evaluation period, SEnerCon sent the questionnaire in the beginning of June 2021 to all registered users including two more questions on the rebound effect and the willingness to invest in renewable energy.

29 users in total took part in both surveys, 19 in the first one and 10 in the second survey.

The following tables show the results of the two surveys (mid-term and final), including their respective analyses. The results of the first questions show that users rated the pragmatic quality aspects of Eco-Bot above average in most cases (above 3). However, the target of 4,25 was not reached. For the efficient use of Eco-Bot, users of the first survey rather did not agree (2,71). However, in the second survey this opinion had changed to a neutral opinion on the efficiency of Eco-Bot (3,2). After the first pilot phase, SEnerCon put additional effort in the onboarding of new users, assisting them in personal on-line session in the registration and explaining Eco-Bot's features. This proved to be effective, as the ratings resulting from the second survey are higher, as shown in the table below.

Table 19: User Experience Survey: Pragmatic Quality Aspects - SEnerCon

PRAGMATIC QUALITY ASPECTS				
Efficiency	**Is it possible to use Eco-Bot quickly and efficiently?**			
		Slow (1) / Fast (5)	Inefficient (1) / Efficient (5)	impractical (1) / practical (5)
First user survey (mid-term)	Median Response	4	3	3
	Average Response	**3,47**	**2,71**	**2,88**
Second user survey (end of pilot)	Median Response	4	3	3,5
	Average Response	**3,6**	**3,2**	**3,8**
Perspicuity	**Is it easy to get familiar with Eco-Bot? Is it easy to learn how to use it?**			
		confusing (1) / clear (5)	difficult to learn (1) / easy to learn (5)	complicated (1) / easy (5)
First user survey (mid-term)	Median Response	3	3	3
	Average Response	**3,12**	**3,41**	**3,24**
Second user survey (end of pilot)	Median Response	4	4	4
	Average Response	**3,9**	**3,9**	**3,8**
Dependability	**Do you feel that Eco-Bot is reliable and trustworthy? Is the interaction secure and predictable?**			
		unpredictable (1) / predictable (5)	not secure (1) / secure (5)	unreliable (1) / reliable (5)
First user survey (mid-term)	Median Response	3	4	4
	Average Response	**2,94**	**3,53**	**3,53**
Second user survey (end of pilot)	Median Response	4	4	3,5
	Average Response	**3,4**	**4,1**	**3,7**

The hedonic quality aspects of Eco-Bot got average ratings or ratings slightly below average in the first survey (see table 6). Whereas in the second survey Eco-Bot was regarded as inventive and creative. Eco-Bot has more inventive features for smart meter users. And in the second survey, more users with smart meters participated in the survey.

Table 20: User Experience Survey: Hedonic Quality Aspects - SEnerCon

HEDONIC QUALITY ASPECTS				
Stimulation	**Is it interesting and exciting to use Eco-Bot? Do you feel motivated to use it further?**			
		boring (1) / exiting (5)	not interesting (1) / interesting (5)	demotivating (1) / motivating (5)
First user survey (mid-term)	Median Response	3	3	3
	Average Response	**2,82**	**2,76**	**3,00**

Second user survey (end of pilot)	Median Response	3	3,5	4
	Average Response	**3,2**	**3,9**	**3,8**
Novelty		**Is Eco-Bot creative and inventive?**		
		dull (1) / creative (5)	uninventive (1) / inventive (5)	
First user survey (mid-term)	Median Response	3	4	
	Average Response	**2,88**	**3,18**	
Second user survey (end of pilot)	Median Response	4,5	4	
	Average Response	**4,1**	**4,1**	

For the chatbot specific responses (Table 21), most users rather agreed that the visual user interface is good and organised with a slight improvement of the rating in the second survey. The improvement of Eco-Bot's interface after the first survey (upon suggestions of some users) proved to be beneficial. More users who participated in the second survey voted Eco-Bot to be a useful tool in the energy efficiency domain and outperforming existing tools. Again, more smart meter users responded to the second survey than to the first one. Non-smart meter users have less data evaluation features and mainly profit from the energy recommendations that are also provided in co2online's energy specials and tools on the co2onlines energy portal. An improvement of the conversional quality of Eco-Bot proved to be beneficial for the users - the parameters were rated higher in the second survey than in the first survey.

Table 21: User Experience Survey: Chatbot-Specific Aspects - SEnerCon

CHATBOT-SPECIFIC ASPECTS		
Chat Interface		**How do you rate Eco-Bot's visual look?**
		Very Poor (1) / Very Good (5)
First survey (mid-term)	Median Response	4
	Average Response	**3,76**
Second survey (end of the pilot)	Median Response	3,33
	Average Response	**4**
		Does the user interface look organised?
		cluttered (1) / organised (5)
First survey (mid-term)	Median Response	4
	Average Response	**3,76**
Second survey (end of the pilot)	Median Response	4
	Average Response	**4**

Usefulness	**A chatbot in the energy efficiency domain is useful.**	
	Strongly Disagree (1) / Strongly Agree (5)	
First survey (mid-term)	Median Response	3
	Average Response	**3,18**
Second survey (end of the pilot)	Median Response	4,5
	Average Response	**4,2**
	Eco-Bot outperforms its existing alternatives (website, app, customer service line, search engines) by offering diverse and/or enhanced functionalities.	
	Strongly Disagree (1) to Strongly Agree (5)	
First survey (mid-term)	Median Response	3
	Average Response	**3,06**
Second survey (end of the pilot)	Median Response	4
	Average Response	**3,8**
	Eco-Bot covers the use cases I am interested in.	
	Strongly Disagree (1) to Strongly Agree (5)	
First survey (mid-term)	Median Response	4
	Average Response	**3,41**
Second survey (end of the pilot)	Median Response	4
	Average Response	**3,9**
Conversation-al Intelligence	**Eco-Bot understands the input text.**	
	Strongly Disagree (1) to Strongly Agree (5)	
First survey (mid-term)	Median Response	3
	Average Response	**2,71**
Second survey (end of the pilot)	Median Response	4,5
	Average Response	**3,6**
	Eco-Bot interprets commands accurately.	
	Strongly Disagree (1) to Strongly Agree (5)	
First survey (mid-term)	Median Response	3
	Average Response	**2,94**
Second survey (end of the pilot)	Median Response	3,5
	Average Response	**3,4**
	Eco-Bot executes requested tasks.	

	Strongly Disagree (1) to Strongly Agree (5)	
First survey (mid-term)	Median Response	3
	Average Response	**2,88**
Second survey (end of the pilot)	Median Response	4
	Average Response	**3,9**
	Eco-Bot is able to retain conversational context (maintain themed discussion) and follow up on a query.	
	Strongly Disagree (1) to Strongly Agree (5)	
First survey (mid-term)	Median Response	3
	Average Response	**2,94**
Second survey (end of the pilot)	Median Response	5
	Average Response	**4**

While users of the first survey were rather not satisfied with Eco-Bot and rather not decided to use it in future, users of the second survey were very satisfied (5 out of 10 users replied with a 5 to this question). Most of the users of the second survey responded that they were likely to recommend Eco-Bot to a friend. See all results in the table below.

Table 22: User Experience Survey: Generic User Experience - SEnerCon

GENERIC USER EXPERIENCE		
Overall Satisfaction	**How would you rate your overall satisfaction with Eco-Bot?**	
	Very Poor (1) to Very Good (5)?	
First survey (mid-term)	Median Response	2
	Average Response	**2,53**
Second survey (end of the pilot)	Median Response	5 (only 5 responses)
	Average Response	**5 (only 5 responses)**
Future Use	**How likely is it that you will use Eco-Bot in the future?**	
	Not at all likely (1) to Extremely likely (5)	
First survey (mid-term)	Median Response	3
	Average Response	**2,88**
Second survey (end of the pilot)	Median Response	3,5
	Average Response	**3,8**
	How likely is it that you would recommend Eco-Bot to a friend or colleague?	

Net Promotor Scale (NPS)		Not at all likely (0) / Extremely likely (10)
First survey (mid-term)	Median Response	3
	Average Response	**3,24**
Second survey (end of the pilot)	Median Response	7,5
	Average Response	**7**

EYPESA (B2C)

EYPESA sent out the questionnaire to all registered users (anonymously) in late January 2021 and collected answers until February 2021. For the second evaluation period, EYPESA sent out the questionnaire in early June 2021 to all registered users. This second questionnaire included two more questions on the rebound effect and the willingness to invest in renewable energy. 15 users took part in the first phase survey and 16 in the final survey.

Regarding the pragmatism of the chatbot (Table 23), results do not reach the target. However, even if the results are not outstanding, the metrics score between 3 and 4 on the 5-point Likert-scale in both evaluations. In the case of efficiency and perspicuity, the first phase evaluation presents slightly better results than the final one, while dependability presents improvements in the results of the final evaluation.

Table 23: User Experience Survey: Pragmatic Quality Aspects - EYPESA

PRAGMATIC QUALITY ASPECTS				
Efficiency	Is it possible to use Eco-Bot quickly and efficiently?			
		Slow (1) / Fast (5)	Inefficient (1) / Efficient (5)	impractical (1) / practical (5)
First user survey (mid-term)	Median Response	4	3	3
	Average Response	**3,91**	**3,55**	**3,15**
Second user survey (end of pilot)	Median Response	4	4	3
	Average Response	**3,47**	**3,54**	**3,19**
Perspicuity	Is it easy to get familiar with Eco-Bot? Is it easy to learn how to use it?			
		confusing (1) / clear (5)	difficult to learn (1) / easy to learn (5)	complicated (1) / easy (5)
First user survey (mid-term)	Median Response	5	4	4
	Average Response	**4,09**	**3,47**	**3,57**
Second user survey (end of pilot)	Median Response	3	4	4
	Average Response	**3,06**	**3,79**	**3,31**

Dependability	Do you feel that Eco-Bot is reliable and trustworthy? Is the interaction secure and predictable?"			
		unpredictable (1) / predictable (5)	not secure (1) / secure (5)	unreliable (1) / reliable (5)
First user survey (mid-term)	Median Response	4	3	3
	Average Response	**3,36**	**3,14**	**3,33**
Second user survey (end of pilot)	Median Response	4	4	4
	Average Response	**3,5**	**3,93**	**3,88**

For the hedonic aspects, Eco-Bot scores mostly between 3 and 4 (Table 24). There is not much difference among the two evaluation periods. However, the target regarding creativity was exceeded in the first phase evaluation and achieved a score of 4,46 out of 5. The same metric in the final survey is somewhat lower. This is probably due to the fact that the "surprise" effect and novelty faded with the months of usage.

Table 24: User Experience Survey: Hedonic Quality Aspects - EYPESA

HEDONIC QUALITY ASPECTS				
Stimulation	Is it interesting and exciting to use Eco-Bot? Do you feel motivated to use it further?			
		boring (1) / exiting (5)	not interesting (1) / interesting (5)	demotivating (1) / motivating (5)
First user survey (mid-term)	Median Response	3	4	3
	Average Response	**3,38**	**3,64**	**3,29**
Second user survey (end of pilot)	Median Response	4	4	3
	Average Response	**3,25**	**3,60**	**3,50**

Novelty	Is Eco-Bot creative and inventive?		
		dull (1) / creative (5)	uninventive (1) / inventive (5)
First user survey (mid-term)	Median Response	5	4
	Average Response	**4,46**	**3,60**
Second user survey (end of pilot)	Median Response	4	4
	Average Response	**3,81**	**3,80**

Even if for the specific aspects the target was not reached, the rate received for the conversational intelligence drastically improved during the second evaluation phase (Table 25). While in the first part of the pilot it seems that there were difficulties in the communication with the chatbot, the final results show that users appreciate the ability of the bot to understand their questions and give proper answers. This is due to the continuous efforts during the pilot of analysing the wrong intents and updating the bots' intelligence.

Table 25: User Experience Survey: Chatbot specific Aspects - EYPESA

CHATBOT-SPECIFIC ASPECTS		
Chat Interface	**How do you rate Eco-Bot's visual look?**	
		Very Poor (1) / Very Good (5)
First survey (mid-term)	Median Response	4
	Average Response	**3,50**
Second survey (end of the pilot)	Median Response	4
	Average Response	**3,38**
	Does the user interface look organised?	
		cluttered (1) / organised (5)
First survey (mid-term)	Median Response	4
	Average Response	**3,64**
Second survey (end of the pilot)	Median Response	4
	Average Response	**3,69**
Usefulness	**A chatbot in the energy efficiency domain is useful.**	
	Strongly Disagree (1) / Strongly Agree (5)	
First survey (mid-term)	Median Response	4
	Average Response	**4**
Second survey (end of the pilot)	Median Response	4
	Average Response	**3,81**
	Eco-Bot outperforms its existing alternatives (website, app, customer service line, search engines) by offering diverse and/or enhanced functionalities.	
	Strongly Disagree (1) to Strongly Agree (5)	
First survey (mid-term)	Median Response	3
	Average Response	**3,27**
Second survey (end of the pilot)	Median Response	3
	Average Response	**3,64**
	Eco-Bot covers the use cases I am interested in.	
	Strongly Disagree (1) to Strongly Agree (5)	
First survey (mid-term)	Median Response	3
	Average Response	**3,09**

Second survey (end of the pilot)	Median Response	4
	Average Response	**3,67**
Conversational Intelligence	**Eco-Bot understands the input text.**	
	Strongly Disagree (1) to Strongly Agree (5)	
First survey (mid-term)	Median Response	2
	Average Response	**2,85**
Second survey (end of the pilot)	Median Response	4
	Average Response	**3,8**
	Eco-Bot interprets commands accurately.	
	Strongly Disagree (1) to Strongly Agree (5)	
First survey (mid-term)	Median Response	3
	Average Response	**2,73**
Second survey (end of the pilot)	Median Response	4
	Average Response	**3,73**
	Eco-Bot executes requested tasks.	
	Strongly Disagree (1) to Strongly Agree (5)	
First survey (mid-term)	Median Response	4
	Average Response	**3,45**
Second survey (end of the pilot)	Median Response	4
	Average Response	**3,93**
	Eco-Bot is able to retain conversational context (maintain themed discussion) and follow up on a query.	
	Strongly Disagree (1) to Strongly Agree (5)	
First survey (mid-term)	Median Response	3
	Average Response	**3,27**
Second survey (end of the pilot)	Median Response	4
	Average Response	**3,73**

Table 26 shows the results that refer to the generic user experience. It can be seen that the metrics are exceeding their evaluation targets in all cases. Not much difference can be seen between the two evaluation periods, with only slight improvements regarding overall satisfaction and NPS at the end of the pilot. On the other end, while in the first phase evaluation the willingness to keep using the bot was high and surpassed the target, it decreased by the end of the pilot. The reason might be that once the users understand their behaviour and are able to change it, the information given by Eco-Bot is less attractive than at the beginning.

Table 26: User Experience Survey: Generic User Experience - EYPESA

GENERIC USER EXPERIENCE		
Overall Satisfaction	**How would you rate your overall satisfaction with Eco-Bot?**	
	Very Poor (1) to Very Good (5)?	
First survey (mid-term)	Median Response	3
	Average Response	**3,08**
Second survey (end of the pilot)	Median Response	4
	Average Response	**3,4**
Future Use	**How likely is it that you will use Eco-Bot in the future?**	
	Not at all likely (1) to Extremely likely (5)	
First survey (mid-term)	Median Response	4
	Average Response	**4,33**
Second survey (end of the pilot)	Median Response	3
	Average Response	**3,33**
Net Promotor Scale (NPS)	**How likely is it that you would recommend Eco-Bot to a friend or colleague?**	
	Not at all likely (0) / Extremely likely (10)	
First survey (mid-term)	Median Response	7
	Average Response	**5,86**
Second survey (end of the pilot)	Median Response	5
	Average Response	**6,06**

DEXMA (B2B)

For DEXMA's pilot (B2B), the questionnaire was sent to all registered users (anonymously) in late December 2020 and answers were collected until January 2021. For the end evaluation period, the questionnaire was sent to all registered users in mid-April 2021. 9 users took part in the survey.

The results of the first questions show that users think Eco-Bot is an efficient and practical tool. Comparing the results from both surveys it can be noted that the perception of the tool's speed has been greatly improved, probably due to the technical improvements carried out during the pilot phase. It must also be noted that, according to the users, the efficiency of the chatbot has decreased, but still has a high grade.

Regarding the perspicuity of the chatbot, the chatbot has lost some clarity during the pilot, which may be caused by the new push notifications due to the Covid-19 pandemic and for the evaluation of recommendations. These new features may have also caused a slight increase in the chatbot's learning curve slope and difficulty to use.

According to the user's replies, the chatbot is a secure tool; it achieves a slightly above average score regarding its reliability. It is probable that users have had occasional issues communicating orders to the chatbot towards the end of the pilot, when they have more knowledge of the tool and may have introduced more complex orders. This affected their perception of the chatbot's reliability. Eco-Bot is viewed by users as an above-average predictable tool, with a score that increases slightly towards the end of the pilot (when users know more what they can expect from it).

Table 27: User Experience Survey: Pragmatic Quality Aspects – DEXMA

PRAGMATIC QUALITY ASPECTS				
Efficiency	Is it possible to use Eco-Bot quickly and efficiently?			
		Slow (1) / Fast (5)	Inefficient (1) / Efficient (5)	impractical (1) / practical (5)
First user survey (mid-term)	Median Response	3	5	5
	Average Response	**2,5**	**5**	**4,5**
Second user survey (end of pilot)	Median Response	4	4	4
	Average Response	**4**	**4**	**4,2**
Perspicuity	Is it easy to get familiar with Eco-Bot? Is it easy to learn how to use it?			
		confusing (1) / clear (5)	difficult to learn (1) / easy to learn (5)	complicated (1) / easy (5)
First user survey (mid-term)	Median Response	4	5	4
	Average Response	**4**	**4,5**	**4**
Second user survey	Median Response	3	4	4

(end of pilot)	**Average Response**	**3**	**3,4**	**3,8**
Dependability	**Do you feel that Eco-Bot is reliable and trustworthy? Is the interaction secure and predictable?**			
		unpredictable (1) / predictable (5)	not secure (1) / secure (5)	unreliable (1) / reliable (5)
First user survey (mid-term)	Median Response	3	4	5
	Average Response	**3**	**4**	**4,5**
Second user survey (end of pilot)	Median Response	4	5	3
	Average Response	**3,4**	**4,8**	**3,2**

Users clearly think that Eco-Bot is a motivating and interesting tool for energy managers, as they can check a great variety of functionalities in one tool. They can obtain energy insights from their whole portfolio, create goals, and league tables and check the disaggregated consumption, among other functionalities. However, users find the chatbot very interesting and quite motivating in the first survey. In the second one they rate it interesting above average and consider it very motivating. It is probable that different people answered in the first and second surveys, which led to this difference in perception. It could also be that users find it more interesting during the initial part of the pilot and this perception decreases towards the end.

On the aspect of novelty, users mostly see Eco-Bot as a creative and inventive tool. However, the score indicates that many users don't see it as neither creative nor dull. This is because energy managers have most functionalities in the chatbot also available in DEXMA's EMS.

Table 28: User Experience Survey: Hedonic Quality Aspects – DEXMA

HEDONIC QUALITY ASPECTS				
Stimulation	**Is it interesting and exciting to use Eco-Bot? Do you feel motivated to use it further?**			
		boring (1) / exiting (5)	not interesting (1) / interesting (5)	demotivating (1) / motivating (5)
First user survey (mid-term)	Median Response	3	5	4
	Average Response	**3,5**	**5**	**4**
Second user survey (end of pilot)	Median Response	4	3	5
	Average Response	**3,6**	**3,6**	**4,8**
Novelty	**Is Eco-Bot creative and inventive?**			
		dull (1) / creative (5)		uninventive (1) / inventive (5)
First user survey (mid-term)	Median Response	3		5
	Average Response	3,5		4

Second user survey (end of pilot)	Median Response	3	2
	Average Response	**3,4**	**2,6**

The opinion of users regarding the chatbot's chat interface is very positive, both for its visual look and its organisation. This score even increases slightly in the final survey, especially for the interface organisation.

The respondents see Eco-Bot as a useful tool, which covers most of the use cases that users demand. The lowest score, while still above average and having improved in the final survey, goes for the opinion about the outperformance of Eco-Bot with respect to its alternatives. Users may give this score because they are used to the EMS functionalities in DEXMA Analyse. Additionally, they may feel that these functionalities are not as complete in the chatbot – whose objective is to give a quick overview of the main energy KPIs ESCO's portfolio and follow-up on previously set goals and energy-saving events.

The answers of the users to the questions related to the chatbot's conversational intelligence have mixed results. On one hand, users consider the chatbot's ability to understand input text below average and its ability to maintain a themed discussion is just perceived as average. This was probably triggered by several bugs that were detected. Some related to the translations to Spanish of the notifications sent periodically by the chatbot. Others related to the fact that the user cannot scroll back to the previous step, but must return to the main menu, which may have caused some confusion. Regarding the interpretation and execution of commands, the score given by users is clearly above average.

Table 29: User Experience Survey: Chatbot Specific Aspects – DEXMA

CHATBOT-SPECIFIC ASPECTS		
Chat Interface	**How do you rate Eco-Bot's visual look?**	
		Very Poor (1) / Very Good (5)
First survey (mid-term)	Median Response	4
	Average Response	4
Second survey (end of the pilot)	Median Response	4
	Average Response	**4,2**
	Does the user interface look organised?	
		cluttered (1) / organised (5)
First survey (mid-term)	Median Response	4
	Average Response	4
Second survey (end of the pilot)	Median Response	5

	Average Response	**4,6**
Usefulness	**A chatbot in the energy efficiency domain is useful.**	
	Strongly Disagree (1) / Strongly Agree (5)	
First survey (mid-term)	Median Response	5
	Average Response	4,50
Second survey (end of the pilot)	Median Response	5
	Average Response	**4,4**
	Eco-Bot outperforms its existing alternatives (website, app, customer service line, search engines) by offering diverse and/or enhanced functionalities.	
	Strongly Disagree (1) to Strongly Agree (5)	
First survey (mid-term)	Median Response	3
	Average Response	3,25
Second survey (end of the pilot)	Median Response	4
	Average Response	**3,6**
	Eco-Bot covers the use cases I am interested in.	
	Strongly Disagree (1) to Strongly Agree (5)	
First survey (mid-term)	Median Response	4
	Average Response	4
Second survey (end of the pilot)	Median Response	4
	Average Response	**3,8**
Conversational Intelligence	**Eco-Bot understands the input text.**	
	Strongly Disagree (1) to Strongly Agree (5)	
First survey (mid-term)	Median Response	4
	Average Response	4
Second survey (end of the pilot)	Median Response	3
	Average Response	**2,8**
	Eco-Bot interprets commands accurately.	
	Strongly Disagree (1) to Strongly Agree (5)	
First survey (mid-term)	Median Response	3
	Average Response	3
Second survey (end of the pilot)	Median Response	4

	Average Response	**3,8**
	Eco-Bot executes requested tasks.	
	Strongly Disagree (1) to Strongly Agree (5)	
First survey (mid-term)	Median Response	4
	Average Response	3,75
Second survey (end of the pilot)	Median Response	4
	Average Response	**3,8**
	Eco-Bot is able to retain conversational context (maintain themed discussion) and follow up on a query.	
	Strongly Disagree (1) to Strongly Agree (5)	
First survey (mid-term)	Median Response	3
	Average Response	3
Second survey (end of the pilot)	Median Response	3
	Average Response	**3**

The user's overall satisfaction with the chatbot is high, with a slight decrease in the second survey. This can be explained by the fact that when users have more experience with the chatbot, they tend to think that it could be improved so it is probable that it is less attractive for them. However, the likeliness of users continuing to use Eco-Bot and recommending it to other users is high, even higher in the second survey. This reflects that users are satisfied with it and are (on average) almost NPS promoters of the chatbot.

Table 30: User Experience Survey: Generic User Experience - DEXMA

GENERIC USER EXPERIENCE		
Overall Satisfaction	**How would you rate your overall satisfaction with Eco-Bot?** Very Poor (1) to Very Good (5)?	
First survey (mid-term)	Median Response	4
	Average Response	**4,25**
Second survey (end of the pilot)	Median Response	4
	Average Response	**4**
Future Use	**How likely is it that you will use Eco-Bot in the future?** Not at all likely (1) to Extremely likely (5)	
First survey (mid-term)	Median Response	4
	Average Response	**4,25**
Second survey (end of the pilot)	Median Response	5
	Average Response	**4,6**
Net Promotor Scale (NPS)	**How likely is it that you would recommend Eco-Bot to a friend or colleague?** Not at all likely (0) / Extremely likely (10)	
First survey (mid-term)	Median Response	8
	Average Response	**7,5**
Second survey (end of the pilot)	Median Response	8
	Average Response	**8**

Results of Engagement and Retention Related Parameters

The parameters refer to the engagement and retention of Eco-Bot users. These parameters are especially important as Eco-Bot will be used as a tool to improve communication with and engagement of customers. By analysing user log-in and activity data, P10 to P19 show how Eco-Bot engages its users. The parameters Total Users, Active Users, and Engaged Users show the number and respective activity degrees of Eco-Bot users, while the retention rate shows how many users return to the platform in a given time frame. How actively Eco-Bot is used is shown through the parameters Sessions Per Day and Per User, as well as Messages Per Session, which is the sum of Bot and User Messages Per Session.

Total Users

Parameter 10 describes the total number of Eco-Bot users registered during the 12 months pilot period. The table below shows details relevant for the evaluation of this parameter.

Given that at the beginning of the pilot phase the targeted number of total users had not been reached by either of the three pilots, recruitment activities were ongoing during the whole pilot phase. This resulted in a growing number of total users (participants) over time. In order to enable analysis and extraction of meaningful conclusions despite the lack of a constant number of users for the whole pilot period, we performed a cohort analysis. That means, we evaluated the chatbot-related metrics by grouping users into cohorts based on the month they started using Eco-Bot.

SEnerCon (B2B2C)

Total users at the end of the pilot: 87 (7 smart meters users and 80 non-smart meter users)

Target: 150 users (50 smart meter users and 100 non-smart meter users) – *Not achieved*

Table 31 shows the total number of recruited participants achieved over time for the SEnerCon pilot. For instance, in December 2020, the SEnerCon pilot had 38 participants, 7 of them owning a smart meter and 31 without a smart meter. While in the last month of the pilot (May 2021), the total number of participants reached 87, as the number of non-smart meter participants increased to 80.

Table 31: Total recruited participants over time – SEnerCon

	April 2020	May 2020	June 2020	July 2020	Aug. 2020	Sept. 2020	Oct. 2020	Nov. 2020	Dec. 2020	Jan. 2021	Feb. 2021	March 2021	April 2021	May 2021
Smart meter users	3	3	3	3	5	5	5	5	7	7	7	7	7	7
Non-smart meter users	3	9	17	17	17	18	26	30	31	35	49	71	76	80
Total users	6	12	20	20	22	23	31	35	38	42	56	78	83	87

Figure 17 depicts the number of total users (recruited participants) over time for the SEnerCon pilot, and the distribution of these users into the two subgroups.

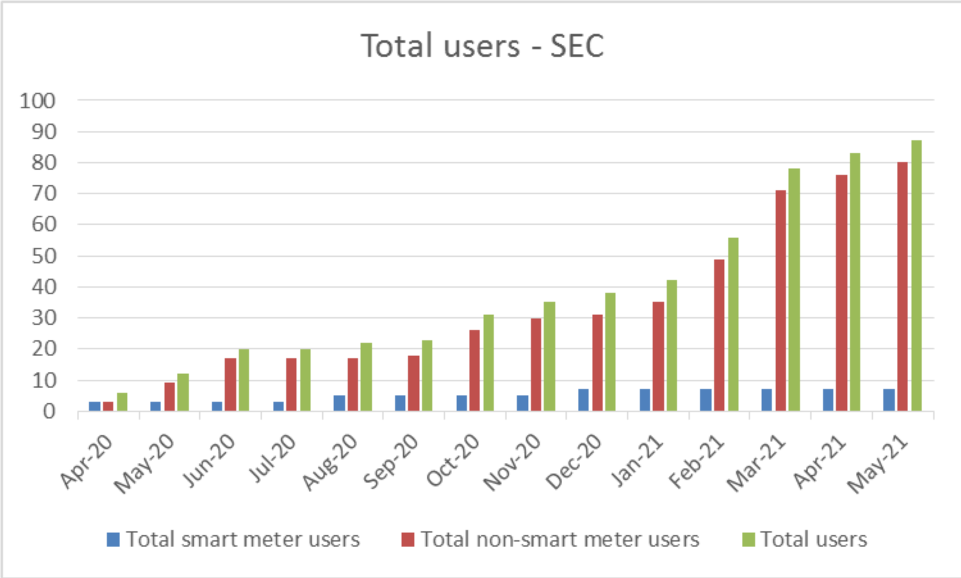

Figure 17: Total Users – SEnerCon

Table 32 shows the number of users per cohort for the SEnerCon pilot, i.e. how many users started using Eco-Bot during each specific month. For each of these cohorts, the table also shows the number of users belonging to each of the two subgroups, i.e. smart meter owners and non-smart meter owners. For instance, in December 2020, 3 users started using Eco-Bot, 2 owning a smart meter and 1 without a smart meter.

Table 32: Number of users per cohort – SEnerCon

	April 2020	May 2020	June 2020	July 2020	Aug. 2020	Sept. 2020	Oct. 2020	Nov. 2020	Dec. 2020	Jan. 2021	Feb. 2021	March 2021	April 2021	May 2021
Total users per cohort	6	6	8	0	2	1	8	4	3	4	14	22	5	4
Smart meter users per cohort	3	0	0	0	2	0	0	0	2	0	0	0	0	0
Non-smart meter users per cohort	3	6	8	0	0	1	8	4	1	4	14	22	5	4

Figure 18 illustrates the grouping of SEnerCon users into cohorts based on the month they started participating in the pilot – considering the subgroup they belonged to. Indicatively, in July 2020 there were no new participants from either of the two subgroups, while in October 2020 there were 8 new participants, all belonging to the non-smart meter group.

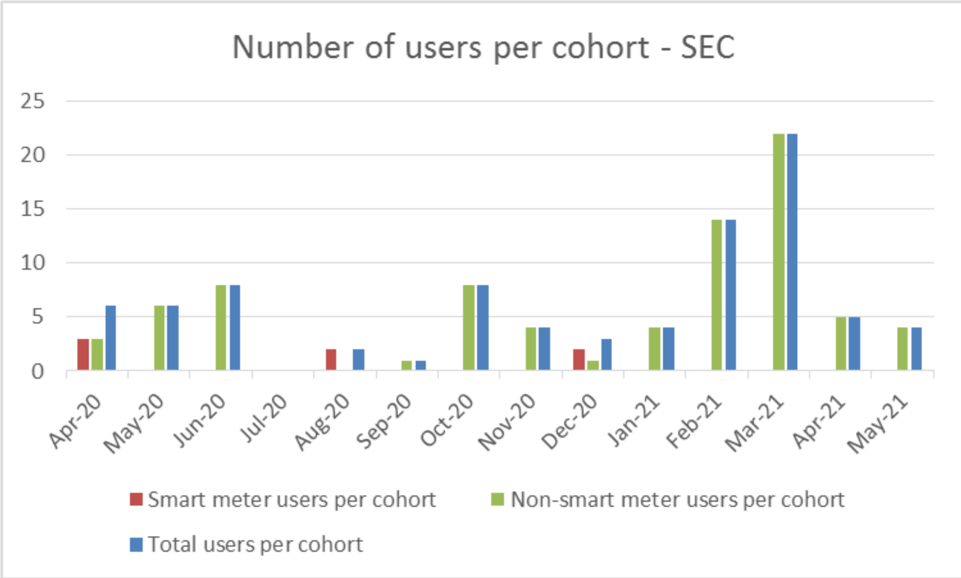

Figure 18: Number of users per cohort – SEnerCon

EYPESA (B2C)

Total users at the end of the pilot: 32[16] (6 advanced users and 24 basic users)

Target: 100 users[17] – *Not achieved*

Table 33 shows the total number of recruited participants achieved over time for the EYPESA pilot. For instance, in June 2020, the EYPESA pilot had 17 participants, 7 of them owning an advanced smart meter and 10 owning the basic smart meter (advanced users and basic users, respectively), while in February 2021 the total number of participants reached 32, 8 of them being advanced users and 24 basic users.

Table 33: Total recruited participants over time – EYPESA

	April 2020	May 2020	June 2020	July 2020	Aug. 2020	Sept. 2020	Oct. 2020	Nov. 2020	Dec. 2020	Jan. 2021	Feb. 2021	March 2021	April 2021	May 2021
Advanced users	6	6	7	7	7	7	7	7	7	7	8	8	8	8

[16] According with the definition of total users in D3.3: "The total number of users registered to the Eco-Bot system will be retrieved from the Eco-Bot database." The total number of users at the end of the pilot was 43. However, 11 of them – even if completed the registration and assigned to a segment – never logged in. It was decided therefore to calculate all the metrics considering only 32 users (the ones that logged in at least once) as the total number of participants.

[17] The involvement of advanced users, i.e. users owning an advanced smart meter in addition to the basic meter, was decided at a later stage of the project. Therefore, there is no specific target defined in the Grant Agreement. While in D3.3 the target for advanced users was 50% of the total participants, i.e. 50 advanced users out of 100 participants in total.

	April 2020	May 2020	June 2020	July 2020	Aug. 2020	Sept. 2020	Oct. 2020	Nov. 2020	Dec. 2020	Jan. 2021	Feb. 2021	March 2021	April 2021	May 2021
Basic users	10	10	10	16	16	17	17	18	18	24	24	24	24	24
Total users	16	16	17	23	23	24	24	25	25	31	32	32	32	32

Figure 19 depicts the number of total users (recruited participants) over time for the EYPESA pilot, and the distribution of these users into the two subgroups.

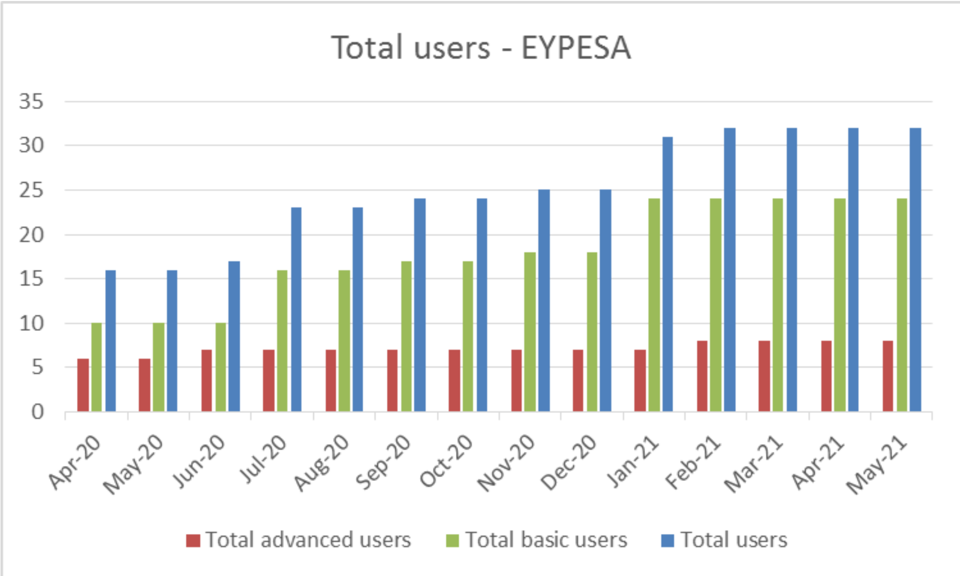

Figure 19: Total Users – EYPESA

Table 34 shows the number of users per cohort for the EYPESA pilot, i.e. how many users started using Eco-Bot during each specific month. For each of these cohorts, the table also shows the number of users belonging to each of the two subgroups, i.e. advanced users and basic users. For instance, in April 2020, 16 users started using Eco-Bot, 6 advanced and 10 basic users.

Table 34: Number of users per cohort – EYPESA

	April 2020	May 2020	June 2020	July 2020	Aug. 2020	Sept. 2020	Oct. 2020	Nov. 2020	Dec. 2020	Jan. 2021	Feb. 2021	March 2021	April 2021	May 2021
Total users per cohort	16	0	1	6	0	1	0	1	0	6	1	0	0	0

	April 2020	May 2020	June 2020	July 2020	Aug. 2020	Sept. 2020	Oct. 2020	Nov. 2020	Dec. 2020	Jan. 2021	Feb. 2021	March 2021	April 2021	May 2021
Advanced users per cohort	6	0	1	0	0	0	0	0	0	0	1	0	0	0
Basic users per cohort	10	0	0	6	0	1	0	1	0	6	0	0	0	0

Figure 20 illustrates the grouping of EYPESA users into cohorts based on the month they started participating in the pilot, taking also into account the subgroup they belonged to. Indicatively, in December 2020 there were no new participants from either of the two subgroups, while in January 2021 there were 6 new participants, all belonging to the basic user group.

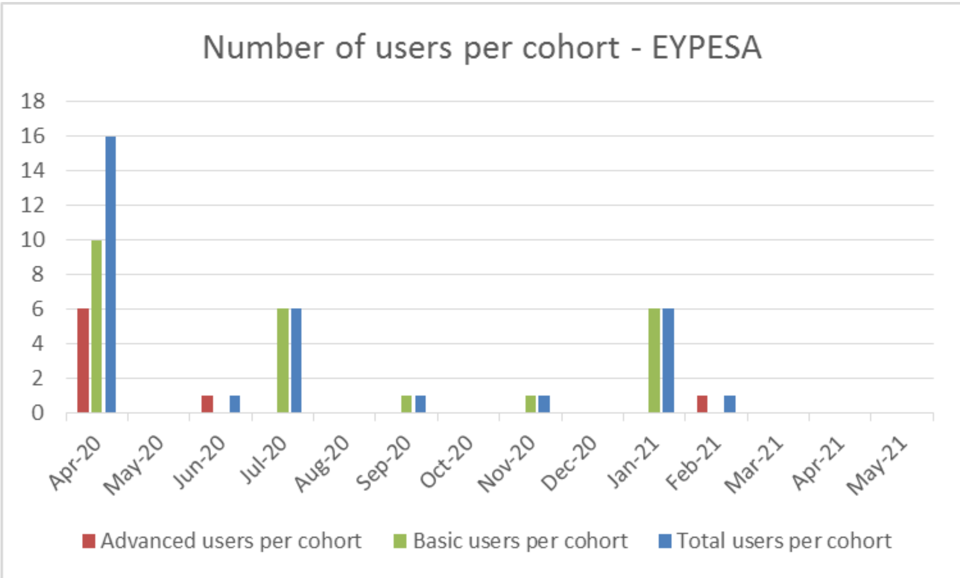

Figure 20: Number of users per cohort – EYPESA

DEXMA (B2B)

Total buildings: 18 (plus 10 bank branches included only in the last month of the pilot)

The 18 buildings that were involved since the beginning of the pilot, were as follows:

- 10 supermarkets
- 7 restaurants (of these 7 restaurants one went out of business in August 2020)
- 1 hotel

Total users (facility managers): 12

It should be noted that 9 out of the 12 total participants were facility managers handling different combinations of the three different types of buildings (supermarkets, restaurants, hotels). In accordance with what was defined in earlier stages of the project, and for the last month of the pilot, 3 additional facility managers with 10 bank branches were also involved.

Target: 7 users (Facility Managers) that will handle a minimum of 20 buildings (hotels, supermarkets and restaurants) – *Partially achieved*

Table 35 shows the total number of recruited participants achieved over time for the DEXMA pilot. For instance, in November 2020 the DEXMA pilot had 9 participants, who were facility managers handling different combinations of 3 types of buildings (supermarkets, restaurants, hotels), while in the last month of the pilot (March 2021), the total number of participants increased to 12, as 3 additional facility managers, handling in total 10 bank branches, joined the pilot.

Table 35: Total recruited participants over time – DEXMA

	April 2020	May 2020	June 2020	July 2020	Aug. 2020	Sept. 2020	Oct. 2020	Nov. 2020	Dec. 2020	Jan. 2021	Feb. 2021	March 2021
Total users	6	6	6	6	6	6	8	9	9	9	9	12

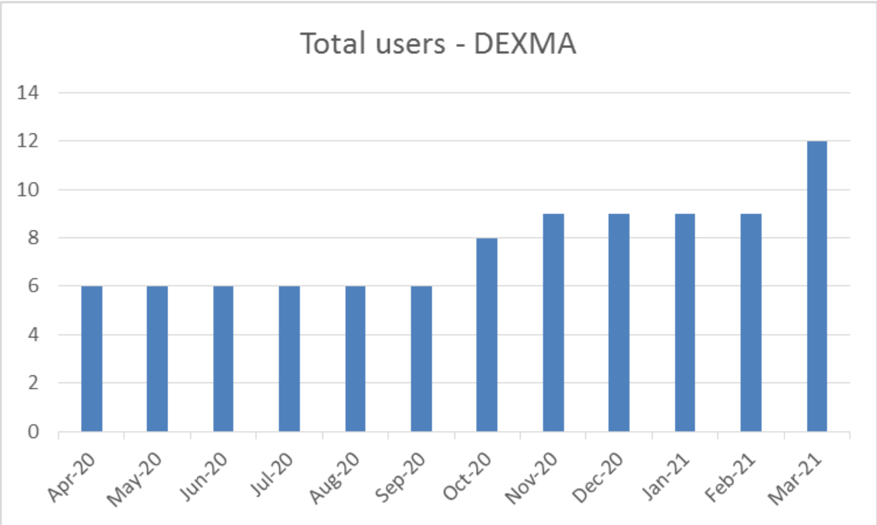

Figure 21 depicts the number of total users (recruited participants) over time for the DEXMA pilot.

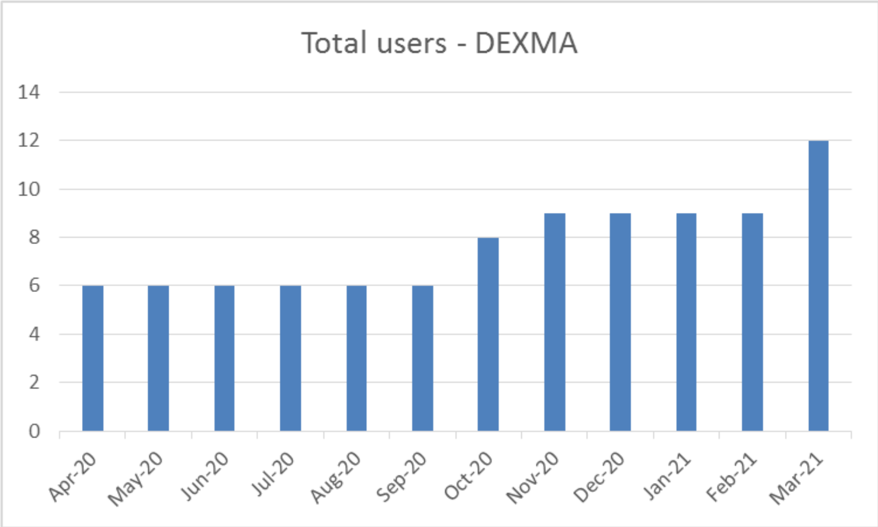

Figure 21: Total Users – DEXMA

Table 36Table 32 shows the number of users per cohort for the DEXMA pilot, i.e. how many users started using Eco-Bot during each specific month. As mentioned above, the facility managers belonging to the 'March 2021' cohort were facility managers of bank branches who joined the pilot for only the last month.

Table 36: Number of users per cohort – DEXMA

	April 2020	May 2020	June 2020	July 2020	Aug. 2020	Sept. 2020	Oct. 2020	Nov. 2020	Dec. 2020	Jan. 2021	Feb. 2021	March 2021
Total users per cohort	6	0	0	0	0	0	2	1	0	0	0	3

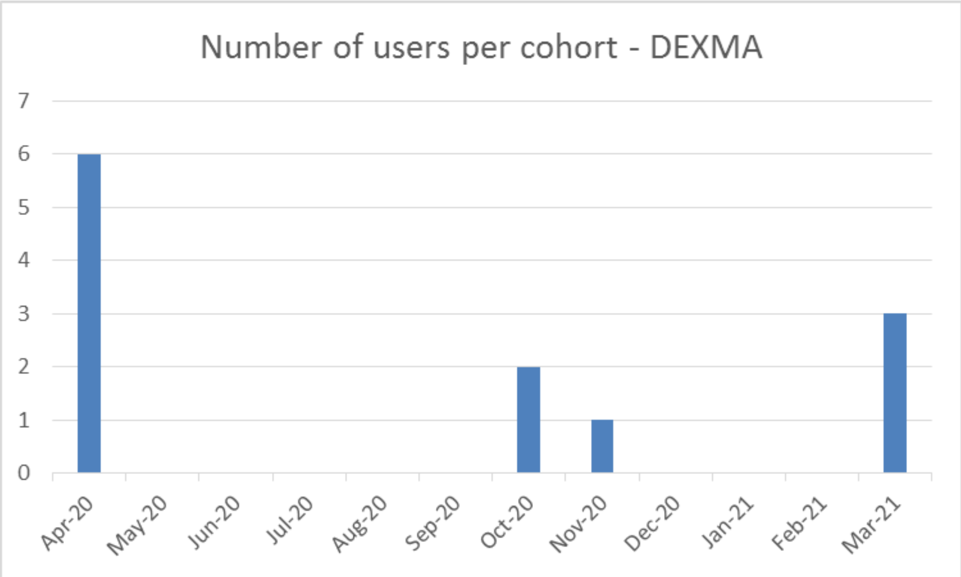

Figure 22 illustrates the grouping of DEXMA users into cohorts based on the month they started participating in the pilot.

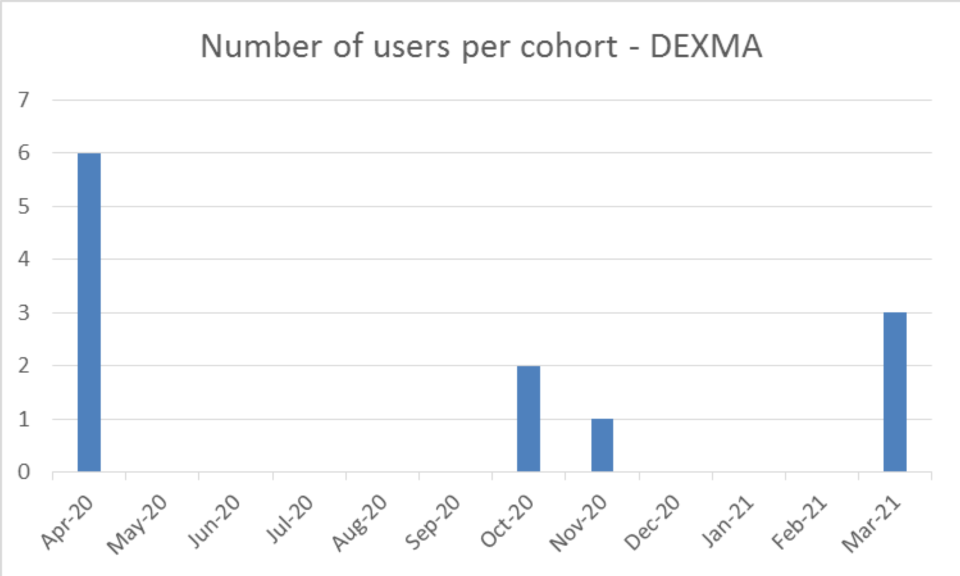

Figure 22: Number of users per cohort – DEXMA

Active users

Active users are defined as the people who log in to Eco-Bot and read a message in a specific month.

Target: Users that logged in Eco-Bot at least once every two months and at least 12 times in total during the 12-month pilot period > 80% of the total users.

It should be noted that the target (for P 11 – active users) was defined in D3.3 considering a 12-month duration for each of the 3 pilots, with all participants involved from the start of the pilots. Unavoidably, given a) the extension of the duration of the 2 out of 3 pilots from 12 to 14 months, and b) the involvement of new participants in much later phases of the pilots and in some cases for only one or two months before the completion of the pilot, the initially defined target should be adapted.

Accordingly, the adapted target is as follows: Users that logged in Eco-Bot at least once every two months and at least for a number of times equal to the total number of participation months > 80% of the total users.

SEnerCon (B2B2C)

Table 37 shows the number of SEnerCon's active users per each month, as well as how many active users per month belonged to each one of the two subgroups, i.e. smart meter users and non-smart meter users. Indicatively, 28 SEnerCon users logged in Eco-Bot in February 2021, 4 of them owning a smart meter and 24 without a smart meter.

Table 37: Active users per month – SEnerCon

	April 2020	May 2020	June 2020	July 2020	Aug. 2020	Sept. 2020	Oct. 2020	Nov. 2020	Dec. 2020	Jan. 2021	Feb. 2021	March 2021	April 2021	May 2021
Active users	6	11	12	2	4	4	16	16	15	19	28	31	19	19
Smart meter users	3	2	2	2	3	2	2	3	6	4	4	3	5	4
Non-smart meter users	3	9	10	0	1	2	14	13	9	15	24	28	14	15

Figure 23 depicts SEnerCon active users per month, as well as SEnerCon active users per subgroup per month.

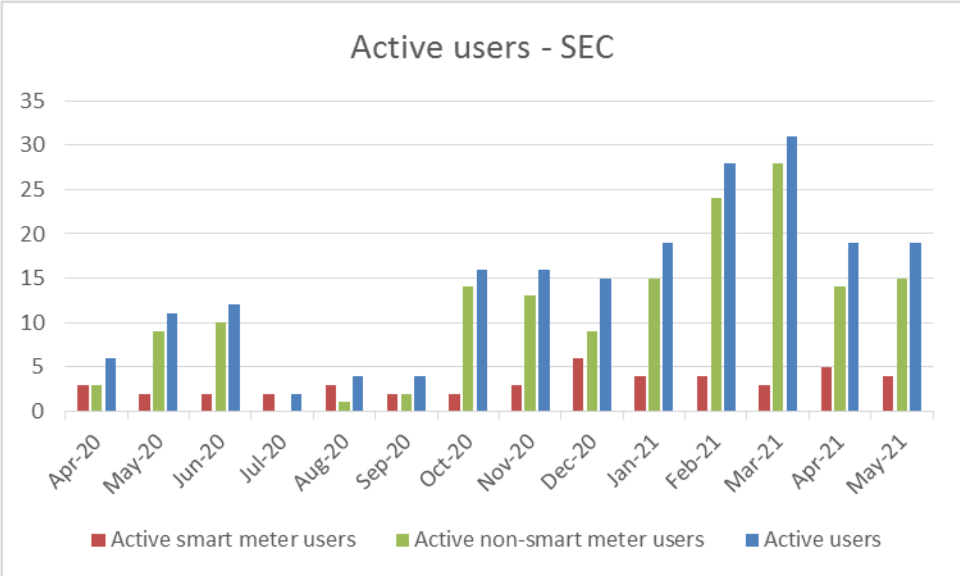

Figure 23: Active users – SEnerCon

Adapted target: SEnerCon users that logged in Eco-Bot at least once every two months and at least for a number of times equal to the total number of participation months > 80% of the total users – *Not achieved*

In order to evaluate the result against the adapted target, we considered the total number of logins of each SEnerCon participant in combination with the total number of his/her participation months (i.e. counting from the month that the specific participant started using Eco-Bot until the completion of the pilot).

Based on the above, the number of SEnerCon participants that met the defined criteria above was 17 out 87 participants in total, i.e. 19.25 percent.

The target was not reached, as there was not always a continuous usage of the bot. Especially during the summer months from June to September 2020 only few users used the bot which is in line with the usage of iESA, which typically decreases during these months as well.

EYPESA (B2C)

Table 38 shows the number of EYPESA active users per each month, as well as how many active users per month belonged to each one of the two subgroups, i.e. advanced users and basic users. Indicatively, in February 2021, 14 EYPESA users logged in Eco-Bot in total, 3 advanced users and 11 basic users.

Table 38: Active users per month - EYPESA

	April 2020	May 2020	June 2020	July 2020	Aug. 2020	Sept. 2020	Oct. 2020	Nov. 2020	Dec. 2020	Jan. 2021	Feb. 2021	March 2021	April 2021	May 2021
Active users	16	4	5	10	0	8	4	4	1	11	14	9	1	7
Advanced users	6	0	3	1	0	3	0	2	0	2	3	1	0	2
Basic users	10	4	2	9	0	5	4	2	1	9	11	8	1	5

Figure 24 depicts EYPESA active users per month, as well as EYPESA active users per subgroup per month.

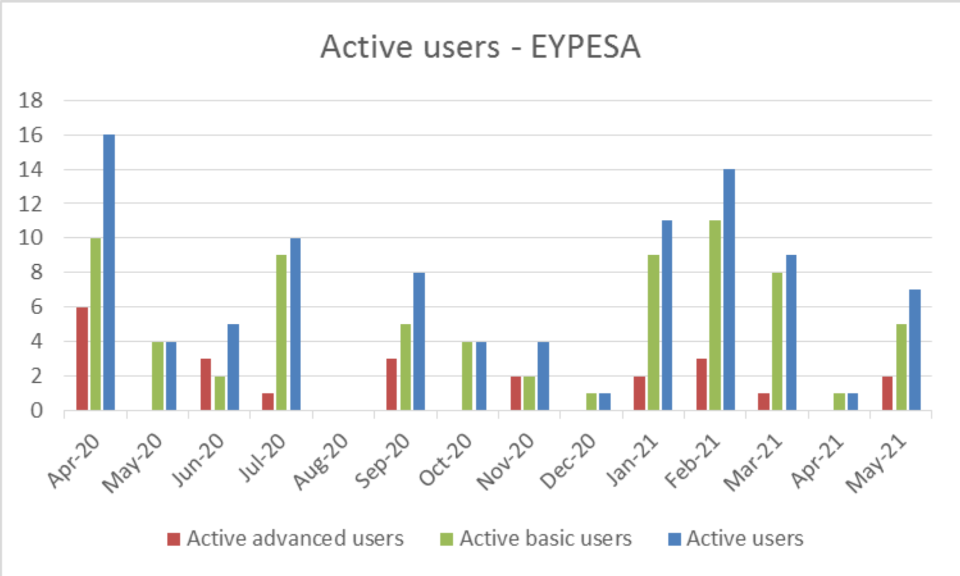

Figure 24: Active users – EYPESA

Adapted target: EYPESA users that logged in Eco-Bot at least once every two months and at least for a number of times equal to the total number of participation months > 80% of the total users – *Not achieved*

The target was not reached. However, it can be seen that users logged in at a different frequency than with the regularity expected at the beginning. For example, for several users there is inactivity during summer months and winter holidays, most likely due to holiday periods. Yet, after these periods, the same users performed more frequent log ins, presenting activity more than once a month for several month in a row. Therefore, even if these users do not fulfil the target set at the beginning, they can still be considered regular – considering the aim of Eco-Bot and the service it wants to give to users.

Therefore, the calculation of active users the target is readapted to "EYPESA users that did not have more than 3 months of inactivity in a row and logged in at least for a number of times equal to the total number of participation months". Consequently, we obtain 47% of active users.

DEXMA (B2B)

The number of DEXMA active users per each month are depicted in Table 39 below. As shown in the table, there was a decrease in the number of facility managers that logged in Eco-Bot in August 2020, more specifically only 1 out of 6 total participants until that time used it during that month, which was expected due to the fact that the specific month was a vacation period for most of the facility managers involved in the pilot.

Table 39: Active users per month – DEXMA

	April 2020	May 2020	June 2020	July 2020	Aug. 2020	Sept. 2020	Oct. 2020	Nov. 2020	Dec. 2020	Jan. 2021	Feb. 2021	March 2021
Active users	6	6	6	6	1	5	8	9	9	9	9	12

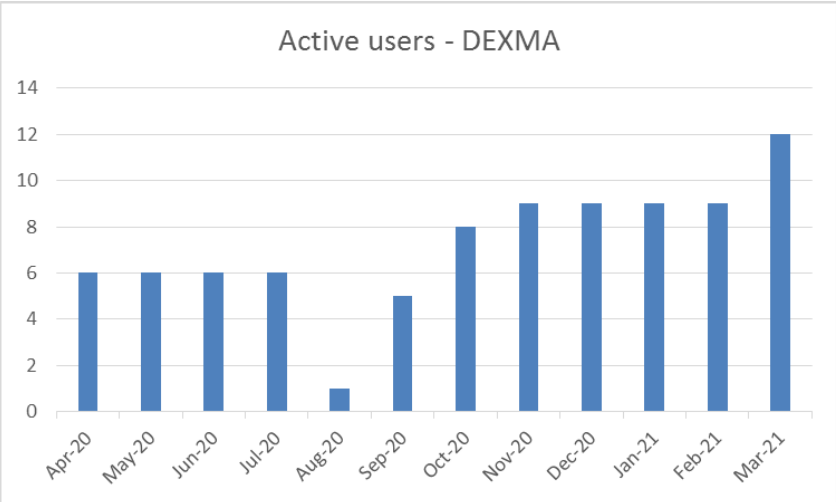

Figure 25 illustrates DEXMA active users per month:

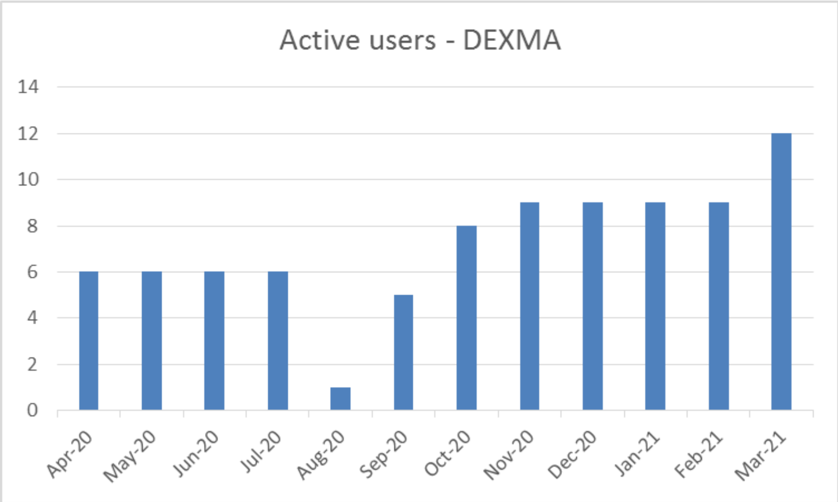

Figure 25: Active users – DEXMA

Adapted target: DEXMA users that logged in Eco-Bot at least once every two months and at least for a number of times equal to the total number of participation months > 80% of the total users – *Achieved*

More specifically, 100% of DEXMA participants met the criteria with regard to the adapted target, as they all used Eco-Bot at least on a monthly basis (with the exception of vacation months, as explained above).

Engaged users

Engaged users are defined as the people who interact with Eco-Bot, i.e. send a message (either making an inquiry or responding to a message from Eco-Bot), in a specific month.

Target: Users that interacted with Eco-Bot at least once every two months and at least 12 times in total during the 12-month pilot period > 80% of the total users.

As in the case of P11 (active users), it should be noted that the above target was defined in D3.3 having in mind that the duration of all 3 pilots would be 12 months and that all participants would be involved from the start of the pilots. For the same reasons discussed in 'P11 – Active users' subsection above, the target regarding P12 should also be adapted so as to consider the different pilot periods as well as the different participation periods of users that were recruited at a later phase.

Accordingly, the adapted target is as follows: Users that interacted with Eco-Bot at least once every two months and at least for a number of times equal to the total number of participation months > 80% of the total users.

SEnerCon (B2B2C)

Table 40 shows the number of SEnerCon engaged users per each month, as well as how many engaged users per month belonged to each one of the two subgroups, i.e. smart meter users and non-smart meter users. Indicatively, 31 SEnerCon users interacted with Eco-Bot in March 2021, 3 of them owning a smart meter and 28 without a smart meter.

It should be noted that the results are the same as in the case of P11 (active users), as all SEnerCon users that were logging in Eco-Bot, were also interacting with it.

Table 40: Engaged users per month – SEnerCon

	April 2020	May 2020	June 2020	July 2020	Aug. 2020	Sept. 2020	Oct. 2020	Nov. 2020	Dec. 2020	Jan. 2021	Feb. 2021	March 2021	April 2021	May 2021
Engaged users	6	11	12	2	4	4	16	16	15	19	28	31	19	19
Smart meter users	3	2	2	2	3	2	2	3	6	4	4	3	5	4
Non-smart meter users	3	9	10	0	1	2	14	13	9	15	24	28	14	15

Figure 26 illustrates SEnerCon engaged users per month, as well as SEnerCon engaged users per subgroup per month.

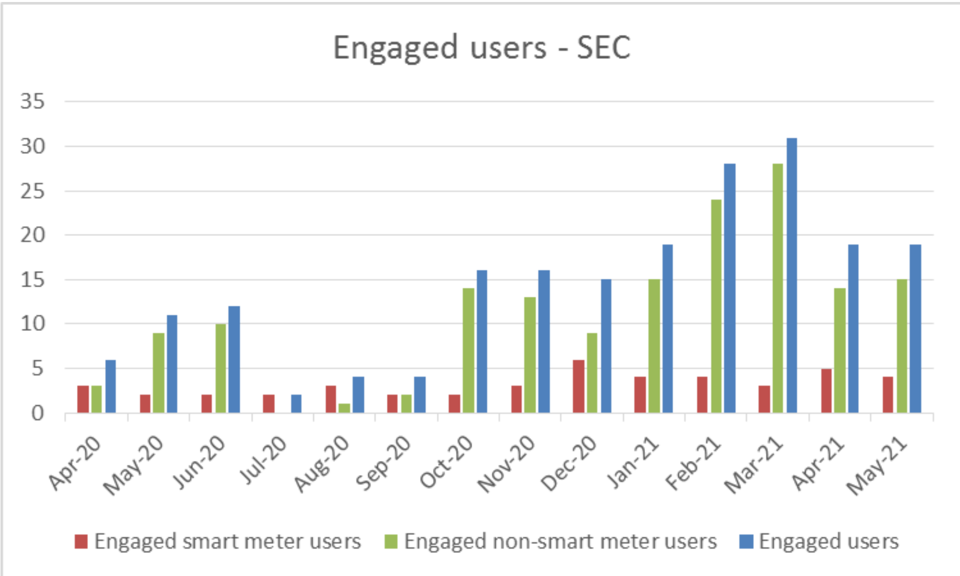

Figure 26: Engaged users – SEnerCon

<u>Adapted target</u>: SEnerCon users that interacted with Eco-Bot at least once every two months and at least for a number of times equal to the total number of participation months > 80% of the total users – *Not achieved*

In order to evaluate the result against the adapted target, we considered the total number of interactions of each SEnerCon participant in combination with the total number of his/her participation months (i.e. counting from the month that the specific participant started using Eco-Bot until the completion of the pilot).

Based on the above, the number of SEnerCon participants that met the defined criteria above was 17 out of 87 participants in total, i.e. 19.25 percent.

EYPESA (B2C)

Table 41 shows the number of EYPESA engaged users per each month, as well as how many engaged users per month belonged to each one of the two subgroups, i.e. advanced users and basic users. Indicatively, in January 2021, 11 EYPESA users interacted with Eco-Bot in total, 2 of them being advanced users and 9 basic users.

It should be noted that the results are the same as in the case of P11 (active users), as all EYPESA users that were logging in Eco-Bot, were also interacting with it.

Table 41: Engaged users per month - EYPESA

	April 2020	May 2020	June 2020	July 2020	Aug. 2020	Sept. 2020	Oct. 2020	Nov. 2020	Dec. 2020	Jan. 2021	Feb. 2021	March 2021	April 2021	May 2021
Engaged users	16	4	5	10	0	8	4	4	1	11	14	9	1	7
Advanced users	6	0	3	1	0	3	0	2	0	2	3	1	0	2
Basic users	10	4	2	9	0	5	4	2	1	9	11	8	1	5

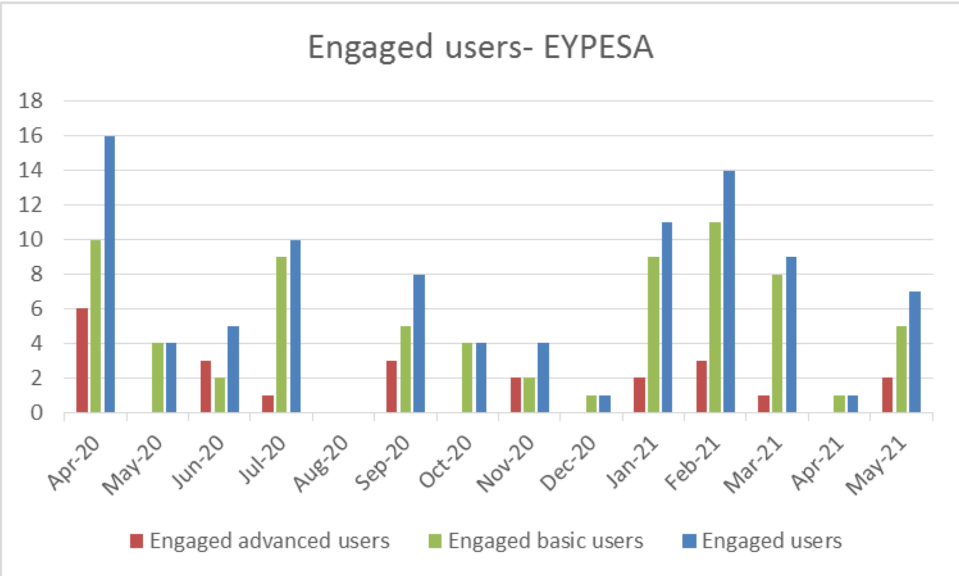

Figure 27 depicts EYPESA engaged users per month, as well as EYPESA engaged users per subgroup per month.

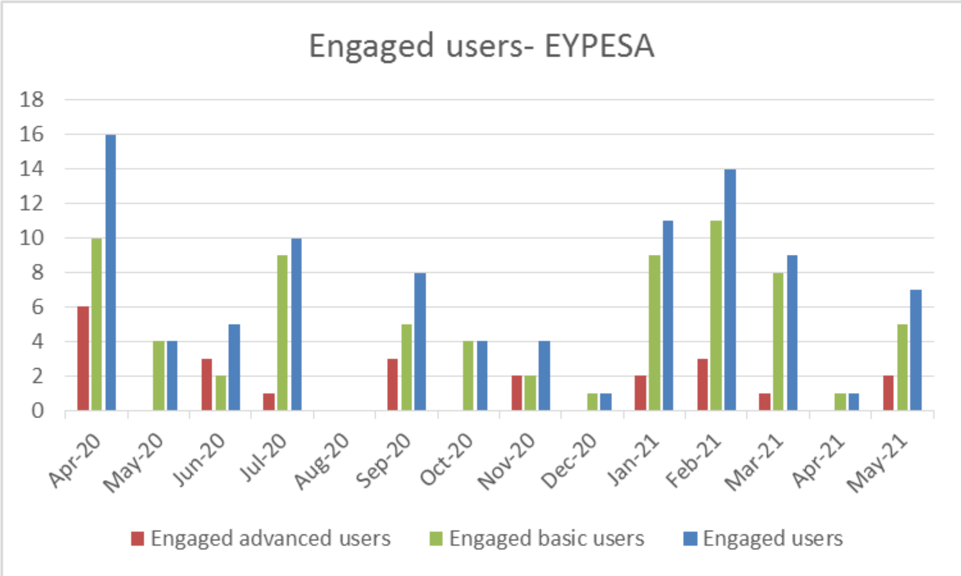

Figure 27: Engaged users – EYPESA

<u>Adapted target</u>: EYPESA users that interacted with Eco-Bot at least once every two months and at least for a number of times equal to the total number of participation months > 80% of the total users – *Not achieved*

For the reasons explained in

Active users, if the engaged users of EYPESA are considered as the *"Users that did not have more than 3 months of inactivity in a row and logged in at least for a number of times equal to the total number of participation months",* 15 users can be considered engaged and correspond to the 47% of the total participants.

DEXMA (B2B)

Table 42 shows the number of DEXMA engaged users per each month. With the exception of a decrease of engaged users in August 2020 and a slight decrease in September 2020, which reflect the vacation periods of the pilot participants during that time, all DEXMA participants interacted with Eco-Bot at least on a monthly basis.

It should be noted that the results are the same as in the case of P11 (active users), as all DEXMA users that were logging in Eco-Bot, were also interacting with it.

Table 42: Engaged users per month – DEXMA

	April 2020	May 2020	June 2020	July 2020	Aug. 2020	Sept. 2020	Oct. 2020	Nov. 2020	Dec. 2020	Jan. 2021	Feb. 2021	March 2021
Engaged users	6	6	6	6	1	5	8	9	9	9	9	12

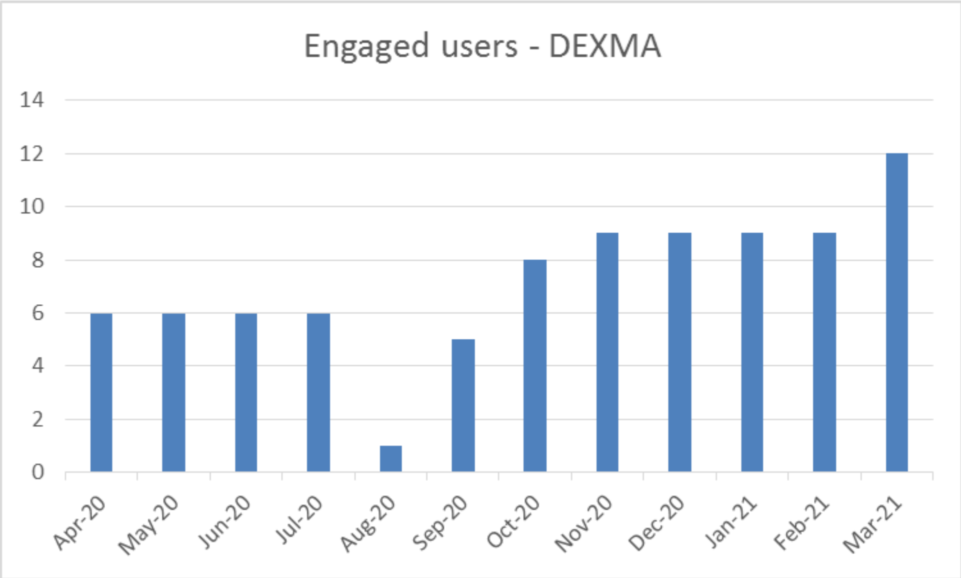

Figure 28 illustrates DEXMA engaged users per month:

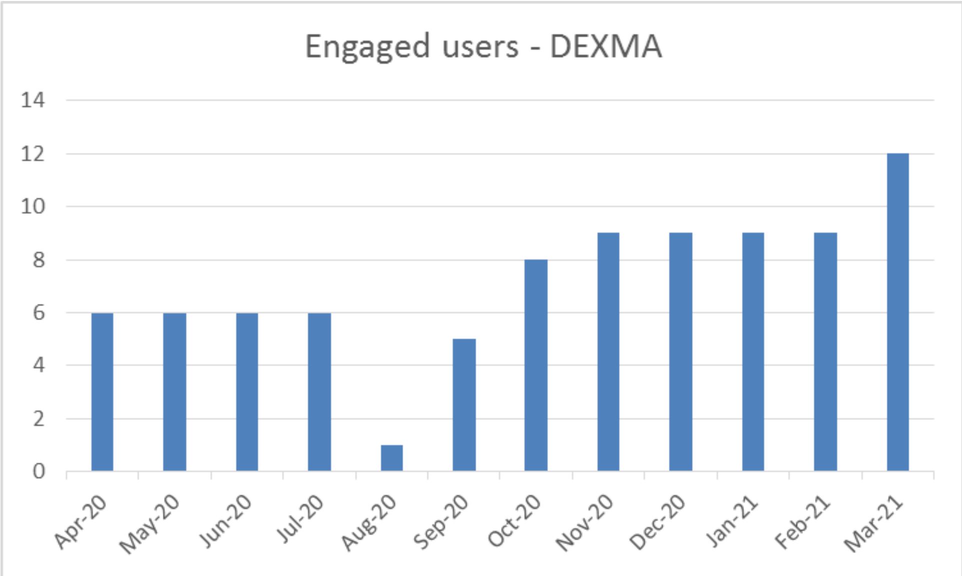

Figure 28: Engaged users – DEXMA

Adapted target: DEXMA users that interacted with Eco-Bot at least once every two months and at least for a number of times equal to the total number of participation months > 80% of the total users – *Achieved*

More specifically, 100% of DEXMA participants met the criteria with regard to the adapted target, as they all interacted with Eco-Bot at least on a monthly basis (with the exception of vacation months, as explained above).

Retention rate

The retention rate shows the percentage of users that returned to Eco-Bot within a certain time period. As explained above (Section 0), the fact that the number of total users (participants) was changing over time due to ongoing recruitment activities during the whole pilot period led us to perform cohort analysis by grouping users into cohorts based on the month they started using Eco-Bot. For instance, the 'June 2020' cohort of a pilot is the group of users of that pilot that used Eco-Bot for the first time in June 2020. The number of users per cohort for all 3 pilots are given in Section 0.

In order to calculate the retention rate per pilot, we considered the user cohorts and examined how many users belonging to the same cohort returned to Eco-Bot 1, 2, 3 etc. months after they started using it. For instance, in the case of SEnerCon's pilot which had a total duration of 14 months, for users belonging to the 'April 2020' cohort, i.e. who started using Eco-Bot in the first month of the pilot, we examined how many of them were returning to Eco-Bot within each of the 13 months that followed after their first month of system use and until the completion of the pilot in May 2021. On the other hand, for users belonging to the 'February 2021' cohort, we examined their retention for the 3 months that were left between their first month of system use and the completion of the pilot. Users that participated only for the last month of the pilots were excluded from the analysis, given that their participation time was limited and did not allow the extraction of any conclusions in terms of retention over time.

SEnerCon (B2B2C)

The retention rate findings for all users of the SEnerCon pilot are given in the following table. The first two columns show the user cohorts and the number of users belonging to each cohort, respectively. 'Month 0' is the month of the first Eco-Bot use per cohort.

Indicatively, as shown in the table, 83% of the users of the 'April 2020' cohort returned to Eco-Bot within the next month and 67% returned also 2 months after their first system use. Moreover, 12 months after the first system use, 50% of the 'April 2020' cohort users continued using Eco-Bot, while 13 months later the corresponding percentage was 33%.

Table 43: Retention rate (percentages) – SEnerCon, All users

Retention (percentages) - SEC - All users															
Cohort	New users	Month 0	Month 1	Month 2	Month 3	Month 4	Month 5	Month 6	Month 7	Month 8	Month 9	Month 10	Month 11	Month 12	Month 13
Apr-20	6	100%	83%	67%	33%	17%	33%	50%	50%	33%	67%	33%	33%	50%	33%
May-20	6	100%	0%	0%	0%	17%	17%	17%	17%	0%	33%	0%	0%	0%	
Jun-20	8	100%	0%	13%	0%	50%	25%	13%	13%	13%	0%	0%	0%		
Aug-20	2	100%	0%	0%	50%	100%	0%	0%	0%	50%	0%				
Sep-20	1	100%	0%	0%	0%	100%	0%	0%	0%	0%					
Oct-20	8	100%	63%	50%	50%	38%	13%	38%	38%						
Nov-20	4	100%	50%	50%	75%	50%	50%	50%							
Dec-20	3	100%	100%	100%	67%	100%	100%								
Jan-21	4	100%	0%	0%	25%	25%									
Feb-21	14	100%	14%	7%	7%										
Mar-21	22	100%	0%	9%											
Apr-21	5	100%	20%												
May-21	4	100%													
		100%	22%	22%	25%	43%	29%	29%	26%	17%	27%	10%	10%	25%	33%

The following two tables present the retention rates for the two SEnerCon subgroups, i.e. the smart meter and the non-smart meter users. Based on the findings, the retention rates of the smart meter users were significantly higher than those of the non-smart meter group. Indicatively, 67% of smart meter users that started using Eco-Bot in April 2020, continued using it 13 months later, i.e. until the end of the pilot.

Table 44: Retention rate (percentages) – SEnerCon, Smart meter users

Retention (percentages) - SEC - Smart meter users															
Cohort	New users	Month 0	Month 1	Month 2	Month 3	Month 4	Month 5	Month 6	Month 7	Month 8	Month 9	Month 10	Month 11	Month 12	Month 13
Apr-20	3	100%	67%	67%	67%	33%	67%	67%	67%	67%	67%	67%	67%	67%	67%
Aug-20	2	100%	0%	0%	50%	100%	0%	0%	0%	50%	0%				
Dec-20	2	100%	100%	100%	50%	100%	100%								
		100%	57%	57%	57%	71%	57%	40%	40%	60%	40%	67%	67%	67%	67%

Table 45: Retention rate (percentages) – SEnerCon, Non-smart meter users

Retention (percentages) - SEC - Non-smart meter users															
Cohort	New users	Month 0	Month 1	Month 2	Month 3	Month 4	Month 5	Month 6	Month 7	Month 8	Month 9	Month 10	Month 11	Month 12	Month 13
Apr-20	3	100%	100%	67%	0%	0%	0%	33%	33%	0%	67%	0%	0%	33%	0%
May-20	6	100%	0%	0%	0%	17%	17%	17%	17%	0%	33%	0%	0%	0%	
Jun-20	8	100%	0%	13%	0%	50%	25%	13%	13%	13%	0%	0%	0%		
Sep-20	1	100%	0%	0%	0%	100%	0%	0%	0%	0%					
Oct-20	8	100%	63%	50%	50%	38%	13%	38%	38%						
Nov-20	4	100%	50%	50%	75%	50%	50%	50%							
Dec-20	1	100%	100%	100%	100%	100%	100%								
Jan-21	4	100%	0%	0%	25%	25%									
Feb-21	14	100%	14%	7%	7%										
Mar-21	22	100%	0%	9%											
Apr-21	5	100%	20%												
May-21	4	100%													
		100%	18%	18%	20%	37%	23%	27%	23%	6%	24%	0%	0%	11%	0%

EYPESA (B2C)

The retention rate findings for all users of the EYPESA pilot are given in the following table. As mentioned above, the first two columns show the user cohorts and the number of users belonging to each cohort, respectively, while 'Month 0' is the month of the first Eco-Bot use per cohort.

Indicatively, as shown in the table, 25% of the users of the 'April 2020' cohort returned to Eco-Bot within the next month and 31% of the users returned also 5 months after their first system use. The retention rate decreased during the months that followed, reaching 19% 13 months after the first system use.

Table 46: Retention rate (percentages) – EYPESA, All users

Retention (percentages) - EYPESA - All users															
Cohort	New users	Month 0	Month 1	Month 2	Month 3	Month 4	Month 5	Month 6	Month 7	Month 8	Month 9	Month 10	Month 11	Month 12	Month 13
Apr-20	16	100%	25%	25%	25%	0%	31%	19%	13%	0%	19%	0%	6%	0%	19%
Jun-20	1	100%	0%	0%	0%	0%	100%	0%	0%	100%	0%	0%	0%		
Jul-20	6	100%	0%	33%	17%	0%	0%	17%	50%	50%	17%	0%			
Sep-20	1	100%	0%	0%	0%	100%	100%	100%	0%	0%					
Nov-20	1	100%	100%	0%	0%	0%	0%	0%							
Jan-21	6	100%	83%	67%	0%	17%									
Feb-21	1	100%	0%	0%	100%										
		100%	31%	31%	19%	6%	28%	20%	21%	17%	17%	22%	6%	0%	19%

The following two tables present the retention rates for the two EYPESA subgroups: the basic smart meter and the advanced smart meter users. Based on the findings, there is no significant difference on the retention rate comparing the two subgroups.

Table 47: Retention rate (percentages) – EYPESA, Advanced users

Retention (percentages) - EYPESA - Advanced users															
Cohort	New users	Month 0	Month 1	Month 2	Month 3	Month 4	Month 5	Month 6	Month 7	Month 8	Month 9	Month 10	Month 11	Month 12	Month 13
Apr-20	6	100%	0%	33%	17%	0%	50%	0%	17%	0%	33%	17%	17%	0%	17%
Jun-20	1	100%	0%	0%	0%	0%	100%	0%	0%	100%	0%	0%	0%		
Feb-21	1	100%	0%	0%	100%										
		100%	0%	25%	25%	0%	57%	0%	14%	14%	29%	14%	14%	0%	17%

Table 48: Retention rate (percentages) – EYPESA, Basic users

Retention (percentages) - EYPESA - Basic users															
Cohort	New users	Month 0	Month 1	Month 2	Month 3	Month 4	Month 5	Month 6	Month 7	Month 8	Month 9	Month 10	Month 11	Month 12	Month 13
Apr-20	10	100%	40%	20%	30%	0%	20%	30%	10%	0%	10%	20%	0%	0%	20%
Jul-20	6	100%	0%	33%	17%	0%	0%	17%	50%	50%	17%	33%			
Sep-20	1	100%	0%	0%	0%	100%	100%	100%	0%	0%					
Nov-20	1	100%	100%	0%	0%	0%	0%	0%							
Jan-21	6	100%	83%	67%	0%	17%									
		100%	42%	33%	17%	8%	17%	28%	24%	18%	13%	25%	0%	0%	20%

DEXMA (B2B)

As mentioned earlier, DEXMA's pilot duration was 12 months, from April 2020 to March 2021. The retention rate findings for all users of the DEXMA pilot are given in the following table, taking also into account the different user cohorts.

As shown in the table, the retention rate of the facility managers was extremely high, reaching 100% in most of the cases, which means that all facility managers that used Eco-Bot continued using it at least once every month. The only exceptions took place in Months 4 and 5 for the 'April 2020' cohort, i.e. August and September 2020, when the retention rate fell to 17% and 83%, respectively, which was expected anyway given that this period coincided with the facility managers' vacation time.

Table 49: Retention rate (percentages) – DEXMA

Retention (percentages) - DEXMA													
Cohort	New users	Month 0	Month 1	Month 2	Month 3	Month 4	Month 5	Month 6	Month 7	Month 8	Month 9	Month 10	Month 11
Apr-20	6	100%	100%	100%	100%	17%	83%	100%	100%	100%	100%	100%	100%
Oct-20	2	100%	100%	100%	100%	100%	100%						
Nov-20	1	100%	100%	100%	100%	100%							
Mar-21	3	100%											
		100%	100%	100%	100%	44%	88%	100%	100%	100%	100%	100%	100%

Sessions per day

This indicator shows the average number of total sessions per day. The results for each of the three pilots are given below.

SEnerCon **(B2B2C)**

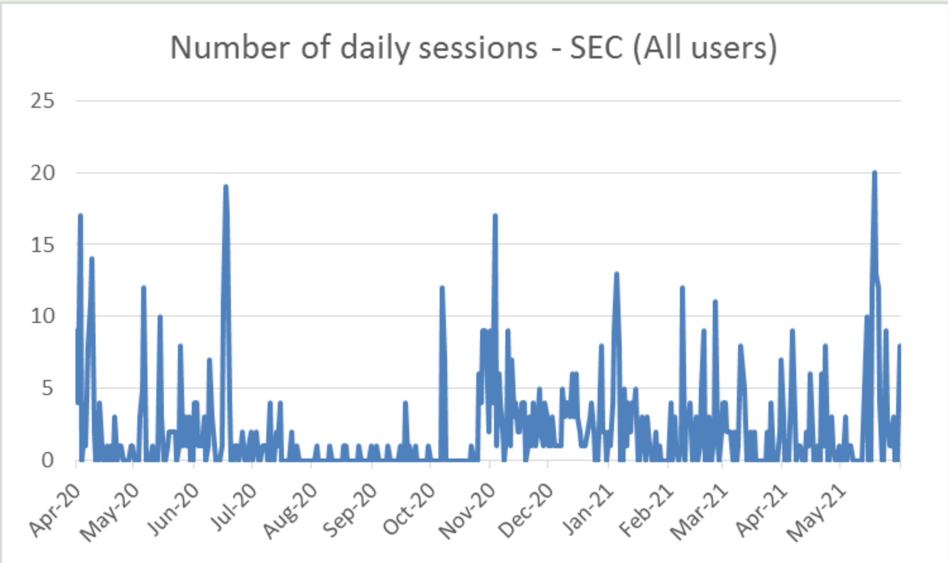

Figure 29 shows the number of total sessions per day for the whole pilot period (April 2020 – May 2021).

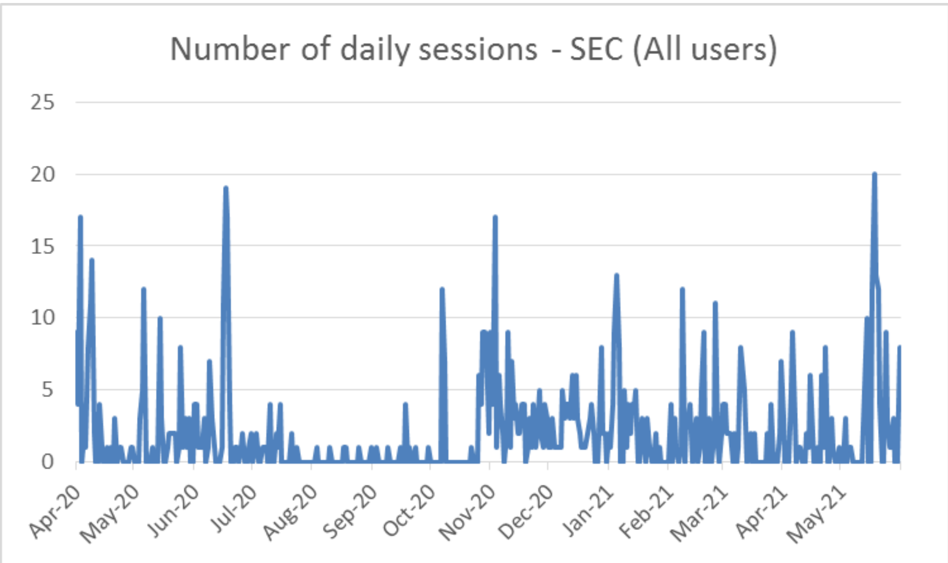

Figure 29: Number of daily sessions – SEnerCon

Figure 30 and Figure 31 below illustrate the number of daily sessions of the smart meter users and the non-smart meter users, respectively. During the summer months July to September 2020 the user activity was rather low, which is in line with the general usage of the iESA in this period. The peak in October 2020 can be related to the start of the heating period in Germany when iESA users usually increase their activity in using the iESA.

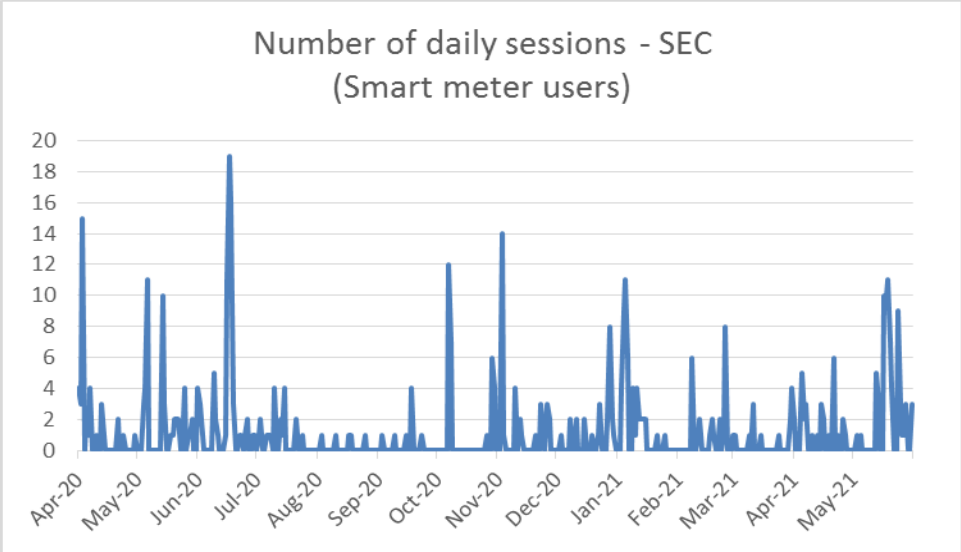

Figure 30: Number of daily sessions – SEnerCon (smart meter users)

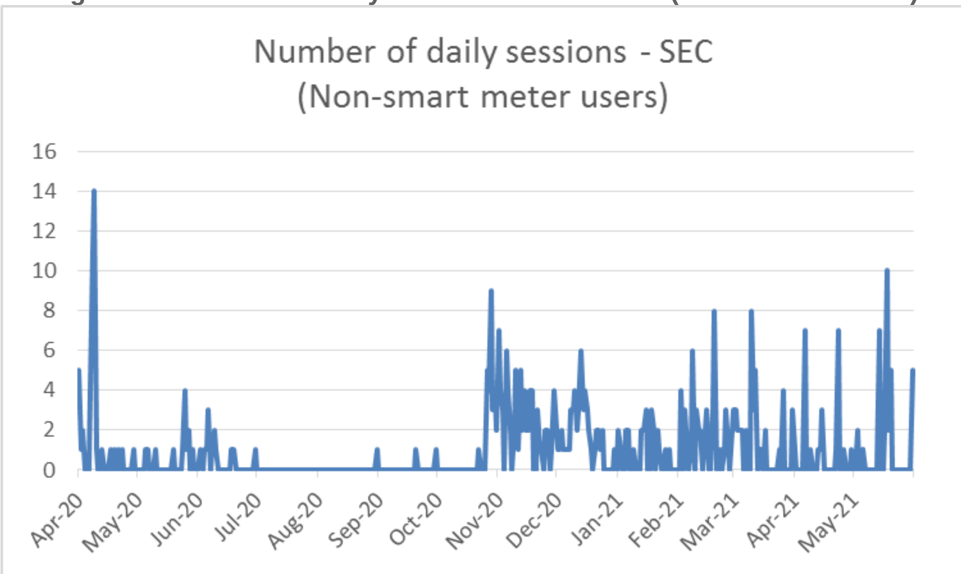

Figure 31: Number of daily sessions – SEnerCon (non-smart meter users)

It can be seen that during the summer months July to September 2020 the user activity was rather low, which is in line with the general usage of the iESA in this period. The peak in October 2020 can be related to the start of the heating period in Germany when iESA users usually increase their activity in using the iESA. If the activity of smart meter users and non-smart meter users is regarded separately, it is interesting that smart meter users have a higher activity in using the bot (although the group of smart meter users is much smaller).

The total number of sessions for the whole pilot period (14 months) for **all SEnerCon users** was **865**.

The total number of days for the whole pilot period of SEnerCon (April 2020 – May 2021) was **426** days.

Based on the above, the average number of total sessions per day for the whole SEnerCon group was: **2.03**.

The total number of sessions for the whole pilot period for **smart meter users** was **469**, and the average number of total sessions per day for this group was **1.10**.

The total number of sessions for the whole pilot period for **non-smart meter users** was **396**, and the average number of total sessions per day for this group was **0.93**.

It should be noted that although the smart meter users were notably smaller in numbers than the non-smart meter ones, they had more sessions in total than the latter group. A reason for this could be that more Eco-Bot features are provided for this group making Eco-Bot more attractive for these users. Also, for the non-smart meter users it is a pre-condition for reliable Eco-Bot answers to enter a certain amount of energy data into the iESA, which is not necessary for the smart meter users. Given the fact that the number of total users (participants) was changing over time, the average number of total sessions per day was calculated on a monthly basis as well, so as to be able to monitor this parameter taking also into account the total number of users participating per month.

Table 50 shows the total number of sessions per month and the corresponding average number of sessions per day for the specific month.

Table 50: Sessions per day – SEnerCon, all users

	April 2020	May 2020	June 2020	July 2020	Aug. 2020	Sept. 2020	Oct. 2020	Nov. 2020	Dec. 2020	Jan. 2021	Feb. 2021	March 2021	April 2021	May 2021
Number of sessions	82	60	86	20	6	10	58	108	76	74	69	56	54	106
Average number of sessions per day	2.73	1.94	2.87	0.65	0.19	0.33	1.87	3.60	2.45	2.39	2.46	1.81	1.80	3.42

Table 51 and Table 52 show the total number of sessions per month and the corresponding average number of sessions per day for the specific month for each one of the two subgroups.

Table 51: Sessions per day – SEnerCon, smart meter users

	April 2020	May 2020	June 2020	July 2020	Aug. 2020	Sept. 2020	Oct. 2020	Nov. 2020	Dec. 2020	Jan. 2021	Feb. 2021	March 2021	April 2021	May 2021
Number of sessions	38	47	73	20	5	8	30	34	25	49	24	14	30	72
Average number of sessions per day	1.27	1.52	2.43	0.65	0.16	0.27	0.97	1.13	0.81	1.58	0.86	0.45	1.00	2.32

Table 52: Sessions per day – SEnerCon, non-smart meter users

	April 2020	May 2020	June 2020	July 2020	Aug. 2020	Sept. 2020	Oct. 2020	Nov. 2020	Dec. 2020	Jan. 2021	Feb. 2021	March 2021	April 2021	May 2021
Number of sessions	44	13	13	0	1	2	28	74	51	25	45	42	24	34
Average number of sessions	1.47	0.42	0.43	0.00	0.03	0.07	0.90	2.47	1.65	0.81	1.61	1.35	0.80	1.10

Figure 32 depicts the total number of sessions per month for all SEnerCon users, as well as for each one of the two subgroups. As was also indicated above, the figure shows that smart meter users had more sessions than the non-smart meter users.

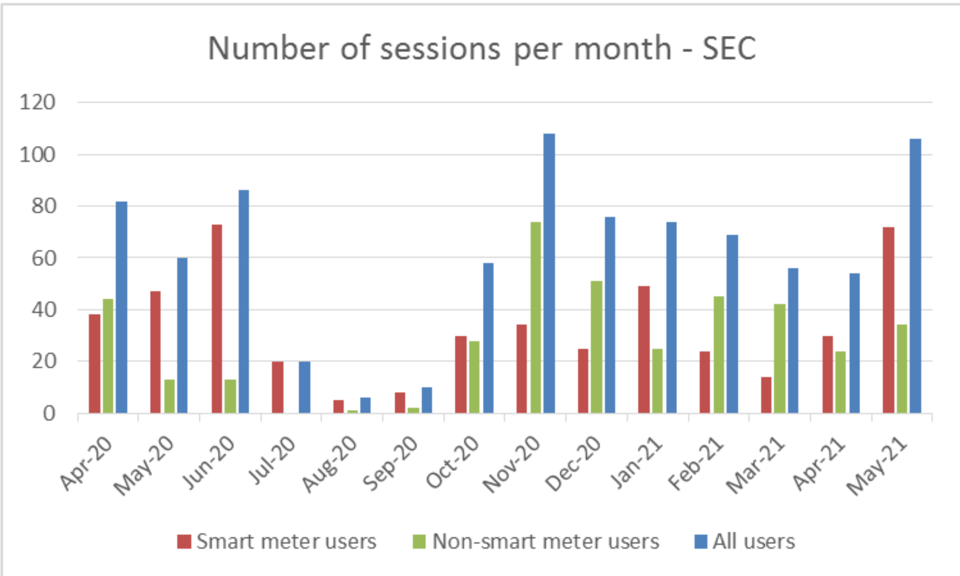

Figure 32: Number of sessions per month – SEnerCon

As a consequence of the low number of logins during the summer months July to September 2020, the number of sessions is similarly low and starts increasing in October 2020. The peaks in November 2020 and in May 2021 may be related to communication activities during these months with challenges including prize lotteries for the usage of the bot and for filling in several user surveys.

EYPESA (B2C)

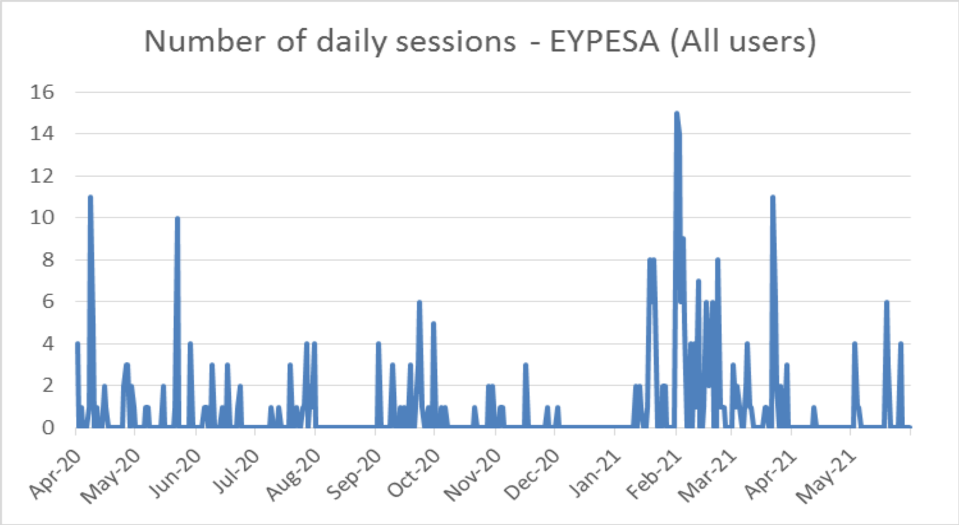

Figure 33 shows the number of total sessions per day for the whole pilot period (April 2020 – May 2021).

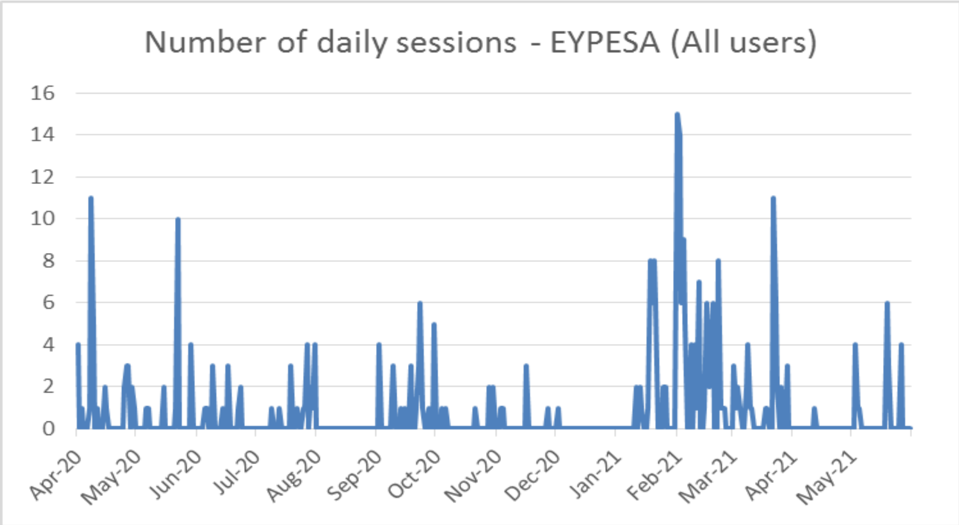

Figure 33: Number of daily sessions – EYPESA

Figure 34 and Figure 35 below illustrate the number of daily sessions of the advanced and the basic users, respectively.

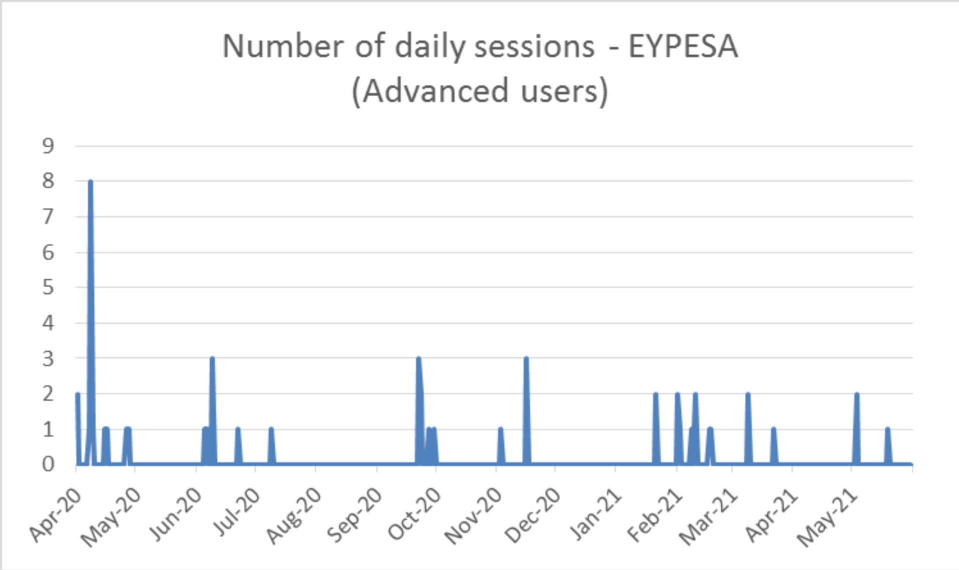

Figure 34: Number of daily sessions – EYPESA, advanced users

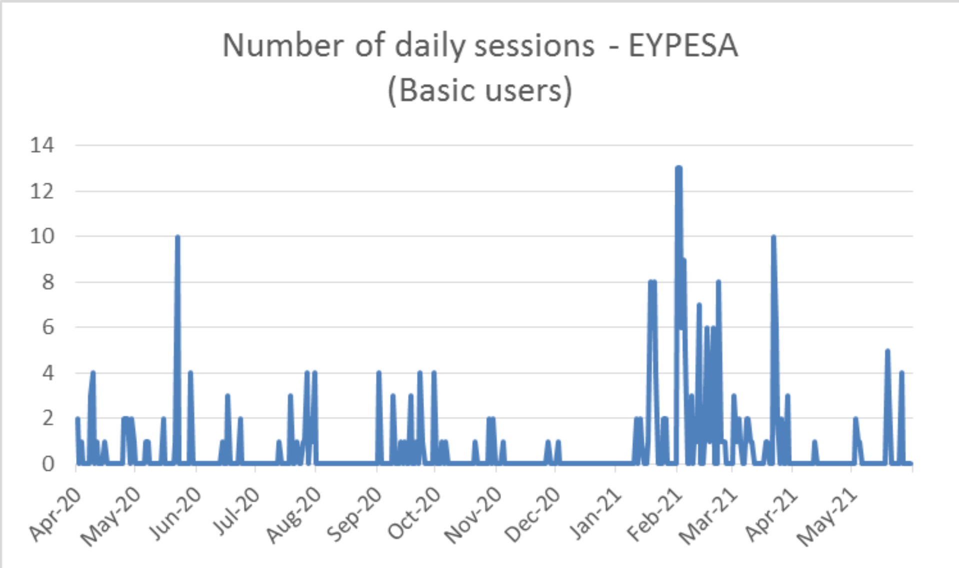

Figure 35: Number of daily sessions – EYPESA, basic users

The total number of sessions for the whole pilot period (14 months) for **all EYPESA users** was **323**.

The total number of days for the whole pilot period of EYPESA (April 2020 – May 2021) was **426** days.

Based on the above, the average number of total sessions per day for the whole EYPESA group was: **0.76**.

The total number of sessions for the whole pilot period for **advanced users** was **50**, and the average number of total sessions per day for this group was **0.12**.

The total number of sessions for the whole pilot period for **basic users** was **273**, and the average number of total sessions per day for this group was **0.64**.

It should be noted that throughout the whole pilot, the number of basic users was approximately three times larger than the number of advanced users, which could partially explain the much larger number of sessions of the basic users' group.

Given the fact that the number of total users (participants) was also changing over time for the EYPESA pilot, the average number of total sessions per day was calculated on a monthly basis as well. This allowed for the monitoring of this parameter, while also considering the total number of users participating per month.

Table 53 shows the total number of sessions per month and the corresponding average number of sessions per day for the specific month.

Table 53: Sessions per day – EYPESA, all users

	April 2020	May 2020	June 2020	July 2020	Aug. 2020	Sept. 2020	Oct. 2020	Nov. 2020	Dec. 2020	Jan. 2021	Feb. 2021	March 2021	April 2021	May 2021
Number of sessions	37	19	12	19	0	29	7	6	1	38	94	41	1	19
Average number of sessions per day	1.23	0.61	0.40	0.61	0.00	0.97	0.23	0.20	0.03	1.23	3.36	1.32	0.03	0.61

Table 54 and Table 55 show the total number of sessions per month and the corresponding average number of sessions per day for the specific month for each one of the two subgroups.

Table 54: Sessions per day – EYPESA, advanced users

	April 2020	May 2020	June 2020	July 2020	Aug. 2020	Sept. 2020	Oct. 2020	Nov. 2020	Dec. 2020	Jan. 2021	Feb. 2021	March 2021	April 2021	May 2021
Number of sessions	16	0	6	1	0	7	0	4	0	2	8	3	0	3
Average number of sessions per day	0.53	0.00	0.20	0.03	0.00	0.23	0.00	0.13	0.00	0.06	0.29	0.10	0.00	0.10

Table 55: Sessions per day – EYPESA, basic users

	April 2020	May 2020	June 2020	July 2020	Aug. 2020	Sept. 2020	Oct. 2020	Nov. 2020	Dec. 2020	Jan. 2021	Feb. 2021	March 2021	April 2021	May 2021
Number of sessions	21	19	6	18	0	22	7	2	1	36	86	38	1	16
Average number of sessions per day	0.70	0.61	0.20	0.58	0.00	0.73	0.23	0.07	0.03	1.16	3.07	1.23	0.03	0.52

Figure 36 depicts the total number of sessions per month for all EYPESA users, as well as for each one of the two subgroups.

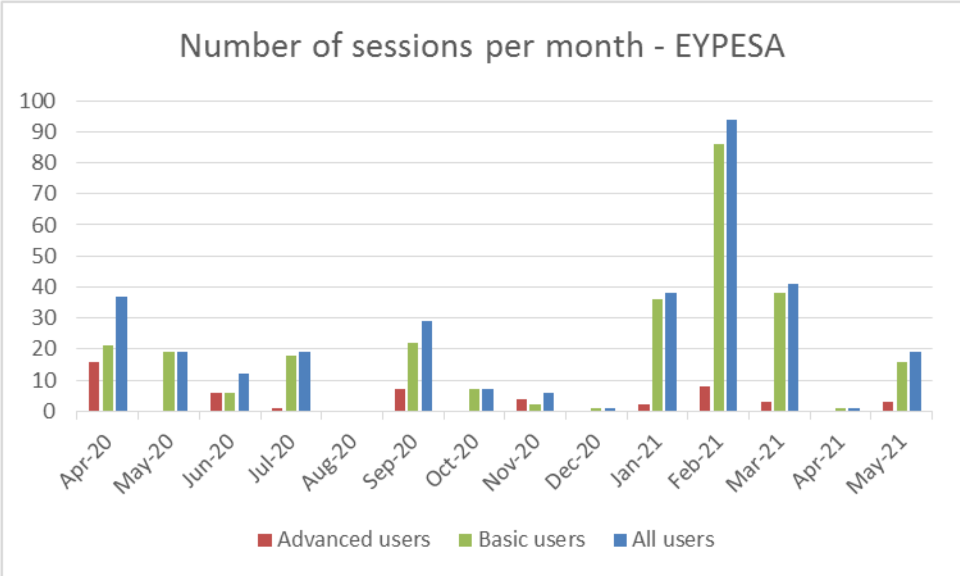

Figure 36: Number of sessions per month – EYPESA

As it can be seen that the months of January, February and March register a peak in the number of sessions. This is most likely caused by the incentive-laden challenge that was launched at the end of January and lasted till the end of March.

DEXMA (B2B)

Figure 37 shows the number of total sessions per day for the whole pilot period (April 2020 – March 2021).

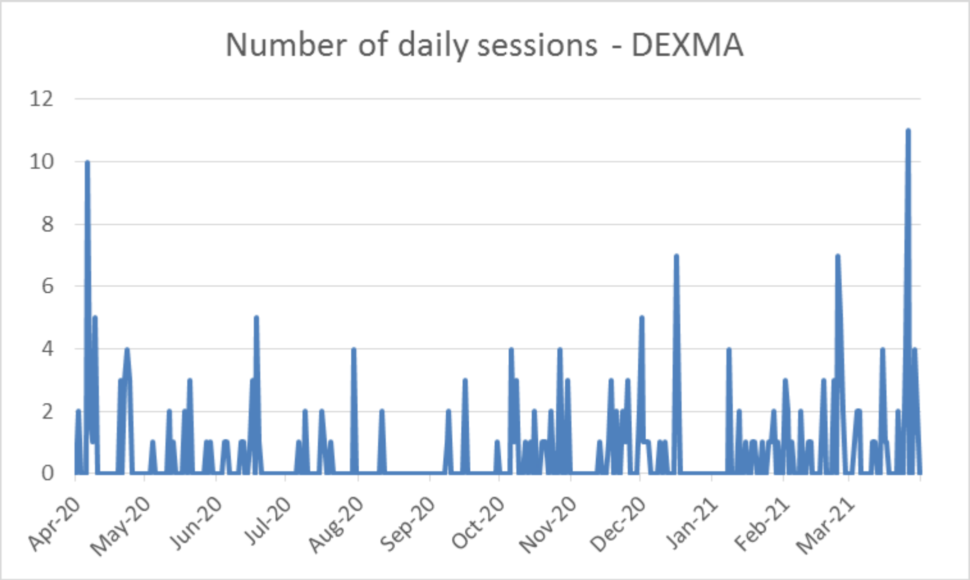

Figure 37: Number of daily sessions – DEXMA

The total number of sessions for the 12-month pilot period was **219**.

Especially in the case of DEXMA users, who are facility managers and used Eco-Bot exclusively on working days, in order to evaluate the average number of sessions per day, we considered an approximate value of **252** working days per year. Based on the above, the average number of total sessions per day was: **0.87**.

However, given the fact that the number of total users (participants) was changing over time also for the DEXMA pilot, the average number of total sessions per day was calculated on a monthly basis as well, so as to be able to monitor this parameter taking also into account the total number of users participating per month. In order to evaluate the average values on a monthly basis, we considered an approximate value of **21** working days per month.

Table 56 shows the total number of sessions per month and the corresponding average number of sessions per day for the specific month.

Table 56: Sessions per day – DEXMA

	April 2020	May 2020	June 2020	July 2020	Aug. 2020	Sept. 2020	Oct. 2020	Nov. 2020	Dec. 2020	Jan. 2021	Feb. 2021	March 2021
Number of sessions	33	11	14	11	2	7	24	15	19	15	32	36

	April 2020	May 2020	June 2020	July 2020	Aug. 2020	Sept. 2020	Oct. 2020	Nov. 2020	Dec. 2020	Jan. 2021	Feb. 2021	March 2021
Average number of sessions	1.57	0.52	0.67	0.52	0.10	0.33	1.14	0.71	0.90	0.71	1.52	1.71

Figure 38 depicts the total number of sessions per month. As can be seen, there was a decrease in the number of sessions that took place in August 2020, while the total number of sessions in March 2021 was the highest one, as during that month the number of participants was also increased from 9 to 12.

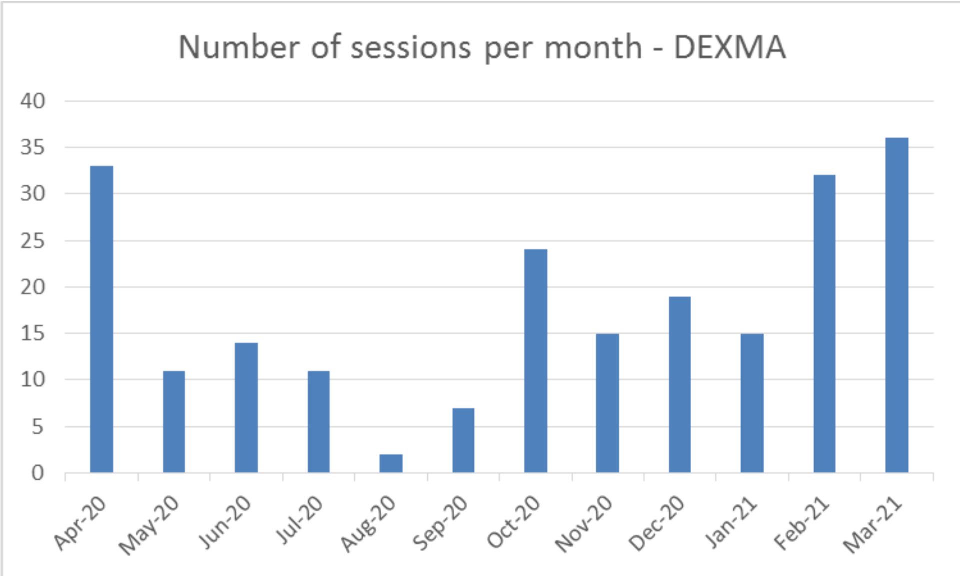

Figure 38: Number of sessions per month – DEXMA

Sessions per user

This indicator shows the average number of total sessions per user. The results for each of the three pilots are given below.

SEnerCon (B2B2C)

The total number of sessions for the whole pilot period for **all SEnerCon users** was **865**.

The total number of users for the whole period was **87**.

Based on the above, the average number of total sessions per user for the whole SEnerCon group was: **9.94**.

The total number of sessions for the whole pilot period for **smart meter users** was **469**, while the total number of smart meter users for the whole period was **7**, thus the average number of total sessions per smart meter user was: **67.00**.

The total number of sessions for the whole pilot period for **non-smart meter users** was **396**, while the total number of non-smart meter users for the whole period was **80**, thus the average number of total sessions per non-smart meter user was: **4.95**.

As evident, smart meter users made on average a significantly higher number of sessions than the non-smart meter users. A reason for this could be that Eco-Bot provided more features for smart meter users than for users with conventional meters.

Given the fact that the number of total users (participants) was changing over time, the average number of total sessions per user was calculated on a monthly basis as well, so as to be able to evaluate this parameter taking also into account the number of users per month.

Table 57 shows the total number of sessions per month, the corresponding number of users and the resulting average number of sessions per user per month.

Table 57: Sessions per user per month – SEnerCon, all users

	April 2020	May 2020	June 2020	July 2020	Aug. 2020	Sept. 2020	Oct. 2020	Nov. 2020	Dec. 2020	Jan. 2021	Feb. 2021	March 2021	April 2021	May 2021
Number of sessions	82	60	86	20	6	10	58	108	76	74	69	56	54	106
Number of users	6	11	12	2	4	4	16	16	15	19	28	31	19	19
Sessions per user	13.67	5.45	7.17	10.00	1.50	2.50	3.63	6.75	5.07	3.89	2.46	1.81	2.84	5.58

Table 58 shows the total number of sessions made by smart meter users per month, the corresponding number of smart meter users and the resulting average number of sessions per smart meter user per month.

Table 58: Sessions per user per month – SEnerCon, smart meter users

	April 2020	May 2020	June 2020	July 2020	Aug. 2020	Sept. 2020	Oct. 2020	Nov. 2020	Dec. 2020	Jan. 2021	Feb. 2021	March 2021	April 2021	May 2021
Number of sessions	38	47	73	20	5	8	30	34	25	49	24	14	30	72
Number of users	3	2	2	2	3	2	2	3	6	4	4	3	5	4
Sessions per user	12.67	23.50	36.50	10.00	1.67	4.00	15.00	11.33	4.17	12.25	6.00	4.67	6.00	18.00

Table 59 shows the total number of sessions made by non-smart meter users per month, the corresponding number of non-smart meter users and the resulting average number of sessions per non-smart meter user per month.

Table 59: Sessions per user per month – SEnerCon, non-smart meter users

	April 2020	May 2020	June 2020	July 2020	Aug. 2020	Sept. 2020	Oct. 2020	Nov. 2020	Dec. 2020	Jan. 2021	Feb. 2021	March 2021	April 2021	May 2021
Number of sessions	44	13	13	0	1	2	28	74	51	25	45	42	24	34
Number of users	3	9	10	0	1	2	14	13	9	15	24	28	14	15
Sessions per user	14.67	1.44	1.30	-	1.00	1.00	2.00	5.69	5.67	1.67	1.88	1.50	1.71	2.27

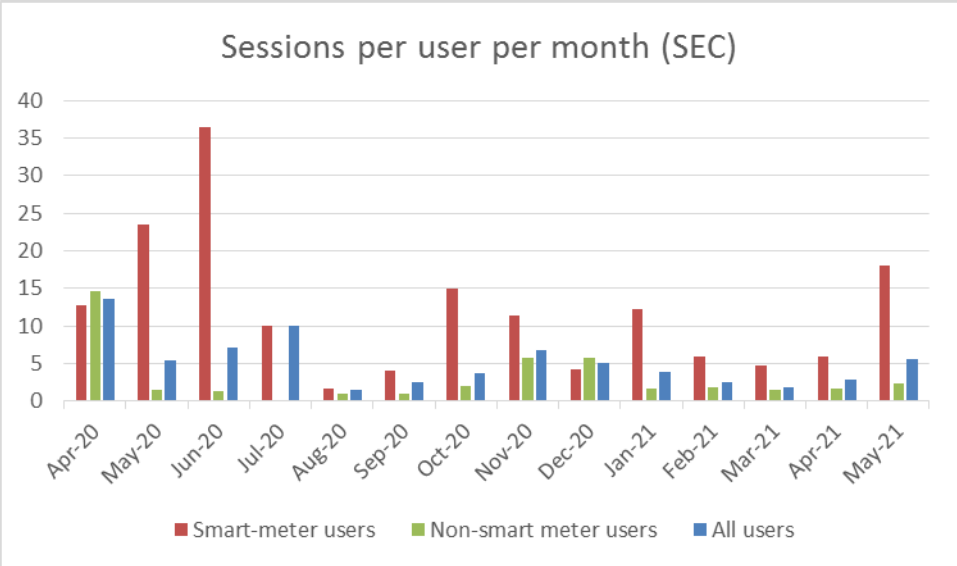

Figure 39 depicts the total number of sessions per user per month for all SEnerCon users, as well as for each one of the two subgroups.

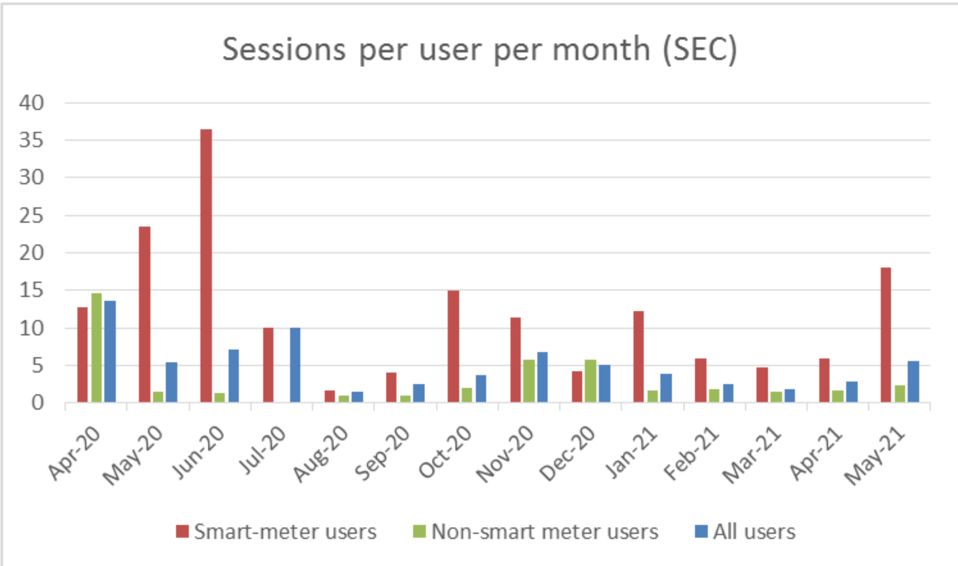

Figure 39: Number of sessions per user per month – SEnerCon

As illustrated in the figure above, in June 2020 there was a high peak in the number of sessions per smart meter user. According to Table 58 above, 2 smart meter users made 73 sessions in total within that month, thus resulting in an average value of 36.50 sessions per user. The peak in June 2020 could be related to new users that were recruited by a collaboration mailing with Discovergy end of May.

EYPESA (B2C)

The total number of sessions for the whole pilot period for **all EYPESA users** was **323**.

The total number of users for the whole period was **32**.

Based on the above, the average number of total sessions per user for the whole EYPESA group was: **10.09**.

The total number of sessions for the whole pilot period for **advanced users** was **50**, while the total number of advanced users for the whole period was **8**, thus the average number of total sessions per advanced user was: **6.25**.

The total number of sessions for the whole pilot period for **basic users** was **273**, while the total number of basic users for the whole period was **24**, thus the average number of total sessions per basic user was: **11.38**.

Given the fact that the number of total users (participants) was changing over time, the average number of total sessions per user was calculated on a monthly basis as well, so as to be able to evaluate this parameter taking also into account the number of users per month.

Table 60 shows the total number of sessions per month, the corresponding number of users and the resulting average number of sessions per user per month.

Table 60: Sessions per user per month – EYPESA, all users

	April 2020	May 2020	June 2020	July 2020	Aug. 2020	Sept. 2020	Oct. 2020	Nov. 2020	Dec. 2020	Jan. 2021	Feb. 2021	March 2021	April 2021	May 2021
Number of sessions	37	19	12	19	0	29	7	6	1	38	94	41	1	19
Number of users	16	4	5	10	0	8	4	4	1	11	14	9	1	7
Sessions per user	2.31	4.75	2.40	1.90	-	3.63	1.75	1.50	1.00	3.45	6.71	4.56	1.00	2.71

Table 61 shows the total number of sessions made by advanced users per month, the corresponding number of advanced users and the resulting average number of sessions per advanced user per month.

Table 61: Sessions per user per month – EYPESA, advanced users

	April 2020	May 2020	June 2020	July 2020	Aug. 2020	Sept. 2020	Oct. 2020	Nov. 2020	Dec. 2020	Jan. 2021	Feb. 2021	March 2021	April 2021	May 2021
Number of sessions	16	0	6	1	0	7	0	4	0	2	8	3	0	3

	April 2020	May 2020	June 2020	July 2020	Aug. 2020	Sept. 2020	Oct. 2020	Nov. 2020	Dec. 2020	Jan. 2021	Feb. 2021	March 2021	April 2021	May 2021
Number of users	6	0	3	1	0	3	0	2	0	2	3	1	0	2
Sessions per user	2.67	-	2.00	1.00	-	2.33	-	2.00	-	1.00	2.67	3.00	-	1.50

Table 62 shows the total number of sessions made by basic users per month, the corresponding number of basic users and the resulting average number of sessions per basic user per month.

Table 62: Sessions per user per month – EYPESA, basic users

	April 2020	May 2020	June 2020	July 2020	Aug. 2020	Sept. 2020	Oct. 2020	Nov. 2020	Dec. 2020	Jan. 2021	Feb. 2021	March 2021	April 2021	May 2021
Number of sessions	21	19	6	18	0	22	7	2	1	36	86	38	1	16
Number of users	10	4	2	9	0	5	4	2	1	9	11	8	1	5
Sessions per user	4.75	3.00	2.00	-	4.40	1.75	1.00	1.00	4.00	7.82	4.75	1.00	3.20	4.40

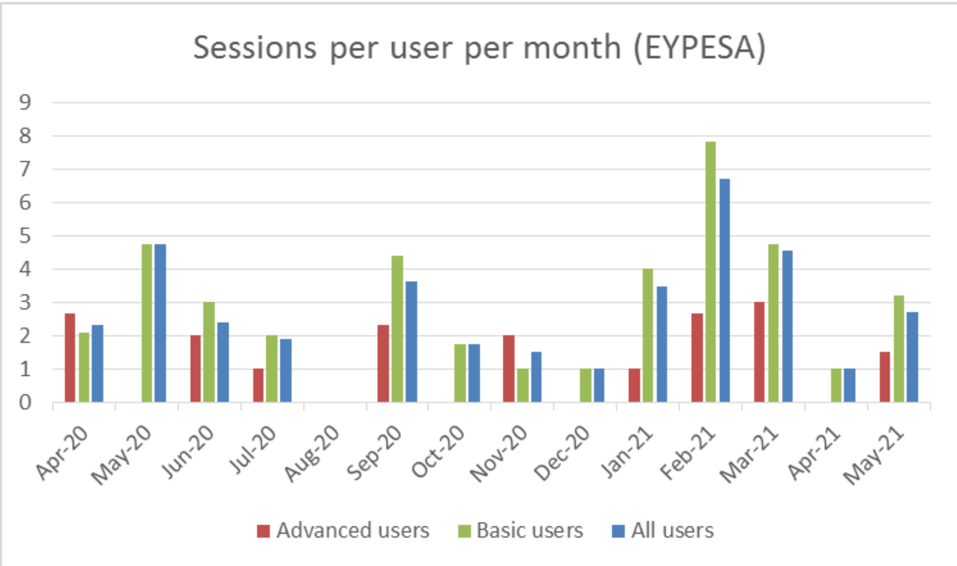

Figure 40 depicts the total number of sessions per user per month for all EYPESA users, as well as for each one of the two subgroups. Again, this is probably due to the challenge with incentives launched at the end of January.

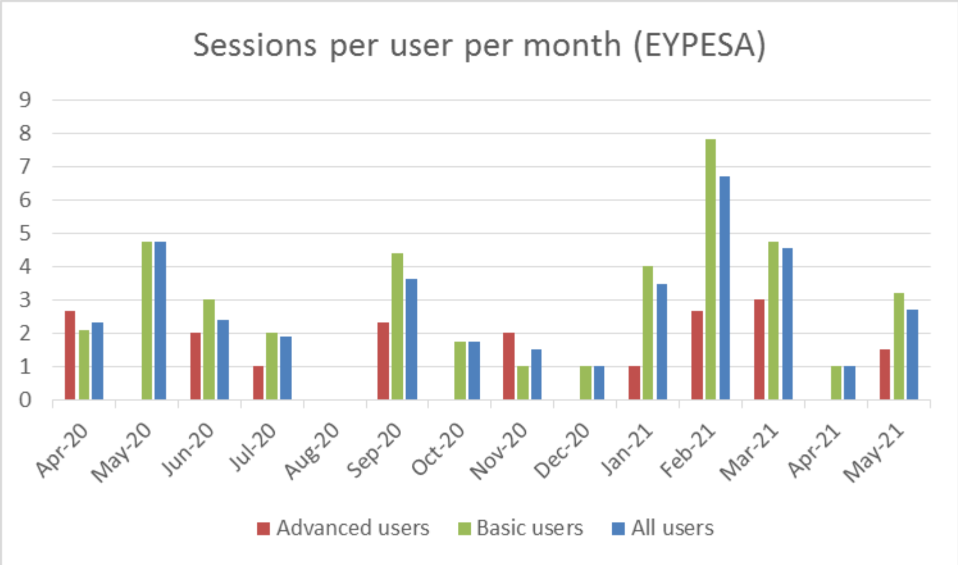

Figure 40: Number of sessions per user per month – EYPESA

DEXMA (B2B)

The total number of sessions for the whole pilot period was **219**.

The total number of users for the whole period was **12**.

Based on the above, the average number of total sessions per user for the DEXMA users was: **18.25**.

Given the fact that the number of total users (participants) was changing over time, the average number of total sessions per user was calculated on a monthly basis as well, so as to be able to evaluate this parameter taking also into account the number of users per month.

Table 63 shows the total number of sessions per month, the corresponding number of users and the resulting average number of sessions per user per month.

Table 63: Sessions per user per month - DEXMA

	April 2020	May 2020	June 2020	July 2020	Aug. 2020	Sept. 2020	Oct. 2020	Nov. 2020	Dec. 2020	Jan. 2021	Feb. 2021	March 2021
Number of sessions	33	11	14	11	2	7	24	15	19	15	32	36
Number of users	6	6	6	6	1	5	8	9	9	9	9	12
Sessions per user	5.50	1.83	2.33	1.83	2.00	1.40	3.00	1.67	2.11	1.67	3.56	3.00

Figure 41 depicts the total number of sessions per user per month for DEXMA users.

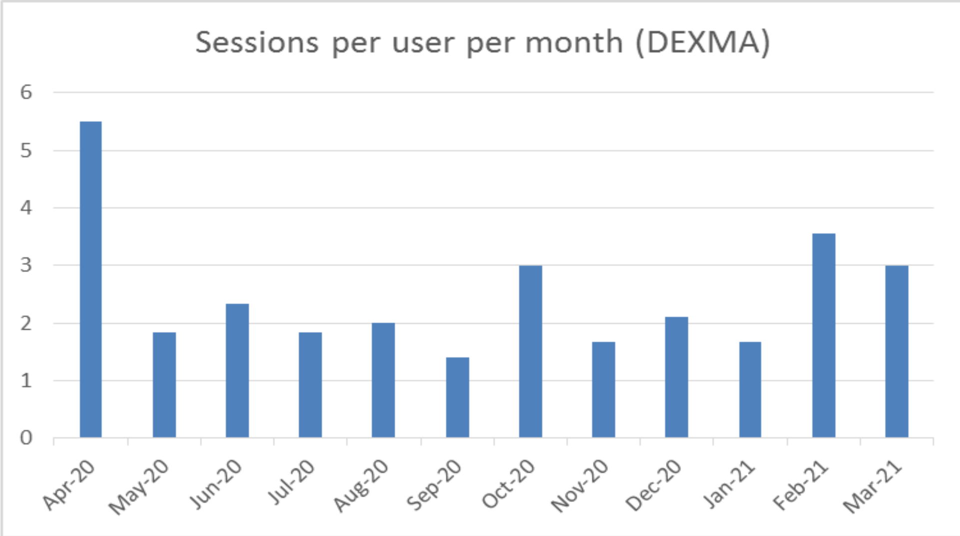

Figure 41: Number of sessions per user per month – DEXMA

Time per session

This indicator shows the average duration of the sessions. The results for each of the three pilots are given below.

SEnerCon (B2B2C)

Table 64 shows the average duration of the sessions for the SEnerCon pilot, as well as the average durations of the sessions made by each of the two subgroups.

Table 64: Time per session - SEnerCon

SEnerCon	Time per Session (in seconds)
All users	150
Smart meter users	155
Non-smart meter users	144

As shown in the table, smart meter users spent more time in their session (on average) than users with conventional meters. A reason for this could be the more complex features that Eco-Bot provides for this group, which take a longer dialogue time for inquiring and providing data.

EYPESA (B2C)

Table 65 shows the average duration of the sessions for the EYPESA pilot, as well as the average durations of the sessions made by each of the two subgroups.

Table 65: Time per session - EYPESA

EYPESA	Time per Session (in seconds)
All users	162
Advanced users	288
Basic users	140

As shown in the table, the advanced users spent on average in their sessions twice the time that the basic users did.

DEXMA (B2B)

The average duration of the sessions of the DEXMA pilot was: 219 seconds.

Messages per sessions

This indicator shows the average number of messages exchanged between the user and the chatbot per session. The results for each of the three pilots are given below, followed by a summary of these results.

SEnerCon (B2B2C)

Table 66 shows the total number of messages exchanged between all SEnerCon users and the chatbot, the total number of sessions, and the resulting average number of messages per session, along with the corresponding information for each of the two subgroups.

Table 66: Messages per session per group - SEnerCon

SEnerCon	Total number of messages	Total number of sessions	Messages per session
All users	11724	865	13.55
Smart meter users	6229	469	13.28
Non-smart meter users	5495	396	13.88

Although the smart meter user group is much smaller compared to the group of users with conventional meters (7 compared to 80 by the end of the pilot), smart meter users only had a slightly lower number of sessions. A reason for this could be that Eco-Bot offers more features for smart meter users than for users with conventional meters.

EYPESA (B2C)

Table 67 shows the total number of messages exchanged between all EYPESA users and the chatbot, the total number of sessions, and the resulting average number of messages per session, along with the corresponding information for each of the two subgroups.

Table 67: Messages per session per group - EYPESA

EYPESA	Total number of messages	Total number of sessions	Messages per session
All users	4831	323	14.96
Advanced users	1122	50	22.44
Basic users	3709	273	13.59

As shown in the table, advanced users exchanged with the chatbot more messages on average than the basic users.

DEXMA (B2B)

Table 68 shows the total number of messages exchanged between DEXMA users and the chatbot, the total number of sessions, and the resulting average number of messages per session.

Table 68: Messages per session per group - DEXMA

DEXMA	Total number of messages	Total number of sessions	Messages per session
All users	6438	219	29.40

TOTAL RESULTS (ALL PILOTS)

Table 69 summarises the results of all three pilots and presents the total results.

Table 69: Messages per session - all pilots

	Total number of messages	Total number of sessions	Messages per session
SEnerCon	11724	865	13.55
EYPESA	4831	323	14.96
DEXMA	6438	219	29.40
TOTAL	22993	1407	16.34

Bot messages per session

This indicator shows the average number of messages sent by the chatbot in one session. The results for each of the three pilots are given below, followed by a summary of these results.

SEnerCon (B2B2C)

Table 70 shows the total number of messages sent by the bot to all SEnerCon users, the total number of SEnerCon users' sessions, and the resulting average number of bot messages per session, along with the corresponding information for each of the two subgroups.

Table 70: Bot Messages per session per group - SEnerCon

SEnerCon	Total number of Bot Messages	Total number of sessions	Bot Messages per session
All users	5611	865	6.49
Smart meter users	2971	469	6.33
Non-smart meter users	2640	396	6.67

EYPESA (B2C)

Table 71 shows the total number of messages sent by the bot to all EYPESA users, the total number of EYPESA users' sessions, and the resulting average number of bot messages per session, along with the corresponding information for each of the two subgroups.

Table 71: Bot Messages per session per group - EYPESA

EYPESA	Total number of Bot Messages	Total number of sessions	Bot Messages per session
All users	2475	323	7.66
Advanced users	562	50	11.24
Basic users	1913	273	7.01

DEXMA (B2B)

Table 72 shows the total number of messages sent by the bot to DEXMA users, the total number of DEXMA users' sessions, and the resulting average number of bot messages per session.

Table 72: Bot Messages per session per group - DEXMA

DEXMA	Total number of Bot Messages	Total number of sessions	Bot Messages per session
All users	3381	219	15.44

TOTAL RESULTS (ALL PILOTS)

Table 73 summarises the results of all three pilots and presents the total results.

Table 73: Bot Messages per session - all pilots

	Total number of Bot Messages	Total number of sessions	Bot Messages per session
SEnerCon	5611	865	6.49
EYPESA	2475	323	7.66
DEXMA	3381	219	15.44
TOTAL	11467	1407	8.15

In messages per session

This indicator shows the average number of messages sent by the user in one session. The results for each of the three pilots are given below, followed by a summary of these results.

SEnerCon (B2B2C)

Table 74 shows the total number of messages sent by all SEnerCon users, the total number of SEnerCon users' sessions, and the resulting average number of In messages per session, along with the corresponding information for each of the two subgroups.

Table 74: In Messages per session per group - SEnerCon

SEnerCon	Total number of In Messages	Total number of sessions	In Messages per session
All users	6113	865	7.07
Smart meter users	3258	469	6.95
Non-smart meter users	2855	396	7.21

Table 75: Bot and In Messages per session per group - SEnerCon

SEnerCon	In Messages per session	Bot Messages per session	Messages per session
All users	7.07	6.49	13.55
Smart meter users	6.95	6.33	13.28
Non-smart meter users	7.21	6.67	13.88

EYPESA (B2C)

Table 76 shows the total number of messages sent by all EYPESA users, the total number of EYPESA users' sessions, and the resulting average number of In messages per session, along with the corresponding information for each of the two subgroups.

Table 76: In Messages per session per group - EYPESA

EYPESA	Total number of In Messages	Total number of sessions	In Messages per session
All users	2356	323	7.29
Advanced users	560	50	11.20
Basic users	1796	273	6.58

DEXMA (B2B)

Table 77 shows the total number of messages sent by DEXMA users, the total number of DEXMA users' sessions, and the resulting average number of In messages per session.

Table 77: In Messages per session per group - DEXMA

DEXMA	Total number of Bot Messages	Total number of sessions	Bot Messages per session
All users	3057	219	13.96

TOTAL RESULTS (ALL PILOTS)

Table 78 summarises the results of all three pilots and presents the total results.

Table 78: In Messages per session - all pilots

	Total number of In Messages	Total number of sessions	In Messages per session
SEnerCon	6113	865	7.07
EYPESA	2356	323	7.29
DEXMA	3057	219	13.96
TOTAL	**11526**	**1407**	**8.19**

Results of Chatbot usability related parameters

Parameters 20 to 23 are monitored to assess Eco-Bot's usage, its efficiency and effectiveness. They gain insights into the usability of the chatbot, assessing its efficiency and effectiveness performance as well as its usage, thus enabling system refinement and optimisation. By analysing the topics popularity index, the most popular questions can be identified, as well as possible new use cases. The Fall-Back Rate gives insight into the amount of times the chatbot fails to complete a task, similar to the Confusion Triggers parameters that measure the times Eco-Bot fails to correctly understand a user message. The Task Completion Time indicates the time needed for Eco-Bot to successfully complete a user request.

These metrics were derived by examining parameters suggested in papers and articles in the field (Neff 2019; AIMultiple 2019; Radziwill & Benton 2017; Newlands 2017; Discover.Bot 2019; ChatbotPack 2019; Lee 2018).

For the Eco-Bot needs, two chatbot agents have been created; one for the pilots of SEnerCon and EYPESA that covers the use cases foreseen for the consumers and one for the pilot of DEXMA that covers the use cases foreseen for the facility managers. The following metrics on the usability are given per chatbot agent.

Topics Popularity Index

The Topics Popularity Index parameter portrays the most popular questions asked by users. The analysis of this metric allows for an assessment of the overall trend of inquiries, making it possible for Eco-Bot to refine the user experience. Furthermore, this parameter helps identifying new use cases and features to support.

The following Table 79 and Table 80 illustrate the Popularity Index for both consumer application of SEnerCon and EYPESA and facility managers in DEXMA's case, respectively.

SEnerCon (B2B2C) and EYPESA (B2C)

Table 79 presents the popularity index of the detected intents and of the use cases for the consumers' pilot. Intents are grouped under the relevant use case and the popularity index of each UC is the sum of the related intents. Most popular use cases were UC-CO-PL-09 ("Check consumption and cost"), UC-CO-PL-10 ("Check consumption of appliances") and UC-CO-PL-18 ("Check consumption in environmental units"). The intents that are grouped under the categories "menu" and "general" refer to general queries as part of the interaction and navigation.

It is valuable to highlight that the table provides insight to the details of each use case through the popularity of each intent. For instance, UC-CO-PL-07 that allows users to enter new energy saving events has a popularity index of 3.82 but the details reveal that only 0.29 reached the confirmation of the event.

Table 79: Popularity Index for consumers - SEnerCon/EYPESA

USE CASE	Detected intent	PI per intent	Total PI per use case
UC-CO-PL-01	contracted.power	0.52	0.52
UC-CO-PL-02	offers.general	0.17	0.17
UC-CO-PL-03	change.contract	0.08	0.08
UC-CO-PL-04	change.title.holder	0.10	0.10
UC-CO-PL-05	change.tariff	0.14	0.14
UC-CO-PL-06	contact.helpdesk	0.02	0.02
UC-CO-PL-07	indicate.event.timestamp	0.05	3.82
	indicate.event.timestamp.confirm	0.00	
	indicate.event.timestamp.wrong.date	0.00	
	save.event.other	1.11	
	save.event.other.category	0.78	
	save.event.other.get	0.87	
	save.event.other.get.cancel	0.04	
	save.event.other.get.confirmation	0.29	
	save.event.other.get.date	0.40	
	save.event.other.other	0.15	
UC-CO-PL-08	save.event.recommendation	1.49	4.99
	save.event.recommendation.get.cancel	0.00	
	save.event.recommendation.get.confirmation	0.08	

USE CASE	Detected intent	PI per intent	Total PI per use case
	save.event.recommendation.get.recommendation	0.85	
	save.event.recommendation.indirect	0.00	
	save.event.recommendation.more	0.11	
	save.event.recommendation.push	2.46	
	save.event.recommendation.save.event.wrong.date	0.00	
	save.event.recommendation.saving.events	0.00	
	save.event.recommendation.wrong.date	0.00	
	save.event.wrong.date	0.00	
UC-CO-PL-09	consumption.total	8.86	14.53
	consumption.total.another.period	4.67	
	consumption.total.wrong.date	1.00	
UC-CO-PL-10	consumption.energy.appliance.specific	2.63	9.48
	consumption.energy.appliance.specific.another.appliance	0.12	
	consumption.energy.appliance.specific.another.period	0.15	
	consumption.energy.appliance.specific.wrong.date	0.66	
	consumption.energy.appliances.general	3.06	
	consumption.energy.appliances.general.another.period	1.21	
	consumption.energy.appliances.general.wrong.date	1.44	
	consumption.total.analysis	0.21	

USE CASE	Detected intent	PI per intent	Total PI per use case
UC-CO-PL-11	compare.appliances.all.first.period	0.98	3.13
	compare.appliances.all.first.period.wrong.date	0.79	
	compare.appliances.all.second.period	0.64	
	compare.appliances.all.second.period.wrong.date	0.03	
	compare.appliances.specific.first.period	0.42	
	compare.appliances.specific.second.period	0.09	
	compare.appliances.specific.wrong.date	0.00	
	consumption.energy.appliance.specific.comparison	0.00	
	consumption.energy.appliances.general.comparison	0.09	
	compare.appliances.specific.first.period.wrong.date	0.10	
UC-CO-PL-12	compare.general	1.13	4.09
	compare.high_low_consumption	2.90	
	compare.high_low_consumption.analysis	0.00	
	compare.high_low_consumption.more	0.00	
	compare.high_low_consumption.other.number	0.01	
	compare.high_low_consumption.other.period	0.05	
	compare.high_low_consumption.wrong.date	0.00	
UC-CO-PL-13	consumption.standby	0.32	0.34
	consumption.standby.period	0.02	

USE CASE	Detected intent	PI per intent	Total PI per use case
UC-CO-PL-14	information.schedule.domestic.activities	0.26	0.26
UC-CO-PL-15	advice.menu	1.39	3.97
	advice.menu.aa	0.04	
	advice.menu.dishwasher	0.07	
	advice.menu.dryer	0.03	
	advice.menu.fridge	0.14	
	advice.menu.heating	0.15	
	advice.menu.oven	0.40	
	advice.menu.washing.machine	0.10	
	advice.more	0.00	
	advice.specific.aa	0.07	
	advice.specific.dishwasher	0.26	
	advice.specific.dryer	0.16	
	advice.specific.fridge	0.28	
	advice.specific.heating	0.28	
	advice.specific.lighting	0.05	
	advice.specific.oven	0.21	
	advice.specific.washing.machine	0.36	
	advice.user.knows	0.00	
UC-CO-PL-16	compare.total.consumption.event	0.75	1.06
	compare.total.consumption.event.periods	0.31	
UC-CO-PL-17	explanation.bill	0.19	0.19
UC-CO-PL-18	consumption.environment.general	2.63	9.09

USE CASE	Detected intent	PI per intent	Total PI per use case
	consumption.environment.general.another.period	1.75	
	consumption.environment.general.another.unit	0.97	
	consumption.environment.general.direct	0.46	
	consumption.environment.general.unit	2.40	
	consumption.environment.general.wrong.date	0.87	
UC-CO-PL-19	compare.total.consumption.event.more	0.02	2.94
	depreciation.compare.new	0.52	
	depreciation.general	1.09	
	depreciation.general.devicesNew	0.76	
	depreciation.general.dont.know	0.00	
	depreciation.general.new	0.29	
	depreciation.oldComparison.new	0.26	
UC-CO-PL-20	information.pv.installation	0.11	0.11
UC-CO-PL-21	compare.total.consumption.event.periods.get	0.44	1.94
	depreciation.general.oldComparison	0.47	
	depreciation.general.oldComparison.old	0.35	
	depreciation.general.oldComparison.old.new	0.67	
UC-CO-PL-22	set.goal	0.71	1.80
	set.goal.direct	0.41	
	set.goal.value	0.19	

USE CASE	Detected intent	PI per intent	Total PI per use case
	set.goal.value.get	0.25	
	set.goal.value.get.cancel	0.02	
	set.goal.value.get.confirmation	0.23	
UC-CO-PL-23	settings.change	0.67	0.67
UC-CO-PL-24	appliance.survey	0.61	0.61
UC-CO-PL-25	time.diary	0.43	0.43
feedback	advice.useful	1.66	3.33
	advice.useless	0.89	
	advice.useless.reason	0.71	
	advice.useless.reason.when	0.07	
general	cancel.general	0.78	11.60
	chat.dissatisfied.customer	0.05	
	chat.mood	0.52	
	explanation.cups	0.00	
	explanation.distributor.contact	0.01	
	explanation.distributor.role	0.01	
	explanation.ecobot	0.06	
	explanation.fallback.kindof	0.02	
	explanation.meter.wrong	0.00	
	Goodbye	0.06	
	Greetings	1.12	
	greetings.bad	0.00	
	greetings.fine	0.01	
	leave.company	0.07	

USE CASE	Detected intent	PI per intent	Total PI per use case
	present.bot	6.45	
	smalltalk.appraisal.good	0.09	
	speak.languages	0.01	
	thank.bot	0.65	
	Unsupported	0.04	
	want.to.save	0.33	
	welcome.message	1.33	
menu	appliances.menu	0.56	14.8
	change.contract.menu	0.00	
	change.menu	0.20	
	compare.appliances	1.53	
	consumption.general	5.29	
	contract.menu	0.19	
	explanation.high.bill	0.06	
	monitoring.menu	1.61	
	recommendation.menu	2.91	
	save.event	1.94	

DEXMA (B2C)

For the facility managers, Table 80 presents the popularity index of the detected intents and of the use cases. The format of the table is the same as above. Most popular use cases were UC-FM-PL-05 (“Check consumption of devices”), UC-FM-PL-04 (“Check total consumption of buildings”) and UC-FM-PL-12 (“Set goals”). It appears that the final intents usually have similar popularity with others in each use case. The difference with the consumers’ cases above can be explained by the constant guidance of the discussions with the facility managers through menus.

Table 80: Popularity Index for facility managers - DEXMA

USE CASE	Detected intent	PI per intent	Total PI per use case
UC-FM-PL-01	save.mvp	0.28	1.04
	save.mvp.building	0.19	
	save.mvp.confirm	0.09	
	save.mvp.date	0.09	
	save.mvp.event	0.09	
	save.mvp.event.category	0.14	
	save.mvp.event.other	0.14	
UC-FM-PL-02	save.recommendation	0.90	2.33
	save.recommendation.event	0.62	
	save.recommendation.more	0.81	
UC-FM-PL-03	check.gateway	0.52	1.28
	check.gateway.location	0.38	
	check.gateway.sensors.final	0.14	
	check.gateway.sensors.question	0.24	
UC-FM-PL-04	consumption.total	2.99	8.02
	consumption.total.add.building	0.24	
	consumption.total.add.building.question	1.14	
	consumption.total.final	1.90	
	consumption.total.period	1.76	
UC-FM-PL-05	consumption.appliances.period	4.56	18.47
	consumption.appliances.sensors.add	0.09	
	consumption.appliances.sensors.question	0.14	
	consumption.appliances.buildings	4.56	

USE CASE	Detected intent	PI per intent	Total PI per use case
	consumption.appliances.final	4.18	
	consumption.appliances	4.94	
UC-FM-PL-06	compare.consumption.portofolio	1.71	4.56
	compare.consumption.portofolio.period	1.52	
	compare.consumption.portofolio.ranking	1.33	
UC-FM-PL-07	compare.consumption.event	0.47	0.66
	compare.consumption.event.period	0.09	
	compare.consumption.event.period.get	0.09	
UC-FM-PL-08	consumption.cost	1.28	4.56
	consumption.cost.final	1.14	
	consumption.cost.period	1.14	
	consumption.cost.add.building	0.24	
	consumption.cost.buildings.question	0.76	
UC-FM-PL-09	consumption.environment	1.00	3.09
	consumption.environment.add.building	0.05	
	consumption.environment.buildings.question	0.28	
	consumption.environment.final	0.90	
	consumption.environment.period	0.85	
UC-FM-PL-10	check.sensors	0.52	1.61
	check.sensors.buildings	0.43	
	check.sensors.period	0.47	
	check.sensors.sensors.question	0.19	

USE CASE	Detected intent	PI per intent	Total PI per use case
UC-FM-PL-11	check.contracts	0.47	0.47
UC-FM-PL-12	set.goal	1.99	7.55
	set.goal.building	1.61	
	set.goal.value	1.52	
	set.goal.value.get.cancel	0.09	
	set.goal.value.get	1.23	
	set.goal.value.get.confirmation	1.09	
UC-FM-PL-13	settings.change	0.19	0.19
check.menu	check.menu	1.52	1.52
Feedback	advice.useful	3.75	15.43
	advice.useless	5.18	
	advice.useless.reason	3.89	
	save.event.recommendation.push	2.61	
General	Cancel	0.05	21.42
	compare.menu	2.75	
	consumption.menu	6.84	
	Greetings	0.33	
	present.bot.menu	9.54	
	save.event.menu	1.19	
	Thanks	0.14	
	welcome.message	0.57	

Fall Back Rate

The Fall Back Rate presents the percentage of times that the chatbot failed to complete a task for all pilots. It refers to cases that either the user made a query that the chatbot was designed to support (designed used cases) but failed to do so or to user queries that the chatbot was not designed to support.

Table 81 presents the values for this metric for all three pilots. For SEnerCon and EYPESA the values match, as the fall-back rate is calculated for the chatbot agent created for the consumers' case. During the pilot, new intents were added in an effort to reduce this percentage and refine responses to users, in case requests were beyond the scope of the project, e.g. referring to wind energy.

Table 81: Fall Back Rate – all pilots

	SEnerCon	EYPESA	DEXMA
Fall Back Rate (percentage)	6.46	6.46	7.79

Confusion Triggers

Analysis of conversations in the respective languages revealed the percentage that the chatbot wrongly matched a user query to a specific intent, including the fall-back intents.

Table 82: Confusion triggers – all pilots

	SEnerCon	EYPESA	DEXMA
Confusion triggers (percentage)	2.42	2.8	0.56

Task completion time

The next two tables present an indicative response time per intent for each of the agents. The first column refers to the relevant use case. Response time depends on the detection of the intent and on the creation the proper response; responses can be simple predefined text, formatted text (e.g. menus, html, pictures) and possibly also the result of communication with the backend. For UC-CO-PL-04 ("Get guidance on changing the meter title holder"), the agent responds in 7ms as it replies with a predefined response. In case of UC-CO-PL-09 ("Check consumption and cost"), when the agent detects the intent for total consumption, it replies within 4ms asking for the period that the user is interested in. Once the period is given, either as a second step or directly in the initial query, the agent indicatively needs 1327ms to reply after communicating with the backend.

The response time for each user may also depend on their details, for instance the portfolio of each facility manager or the available appliances of each consumer but the variations are minor.

Table 83: Task completion time for consumers

USE CASE	Intent	Response time (ms)
UC-CO-PL-01	contracted.power	14
UC-CO-PL-02	offers.general	3174
UC-CO-PL-03	change.contract	11
UC-CO-PL-04	change.title.holder	7
UC-CO-PL-05	change.tariff	9
UC-CO-PL-06	contact.helpdesk	4
UC-CO-PL-07	indicate.event.timestamp	5
	indicate.event.timestamp.confirm	11
	indicate.event.timestamp.wrong.date	6
UC-CO-PL-08	save.event.other	7
	save.event.other.category	5
	save.event.other.get	4
	save.event.other.get.cancel	4
	save.event.other.get.confirmation	5
	save.event.other.get.date	1351
	save.event.other.other	4
	save.event.recommendation	913
	save.event.recommendation.get.cancel	3
	save.event.recommendation.get.confirmation	5
	save.event.recommendation.get.recommendation	1117
	save.event.recommendation.indirect	7
	save.event.recommendation.more	931
	save.event.recommendation.push	8
	save.event.recommendation.save.event.wrong.date	5

USE CASE	Intent	Response time (ms)
	save.event.recommendation.saving.events	7
	save.event.recommendation.wrong.date	5
	save.event.wrong.date	5
UC-CO-PL-09	consumption.total	4
	consumption.total.another.period	1327
	consumption.total.wrong.date	6
UC-CO-PL-10	consumption.energy.appliance.specific	17
	consumption.energy.appliance.specific.another.appliance	1324
	consumption.energy.appliance.specific.another.period	3804
	consumption.energy.appliance.specific.wrong.date	1324
	consumption.energy.appliances.general	6
	consumption.energy.appliances.general.another.period	1487
	consumption.energy.appliances.general.wrong.date	10
	consumption.total.analysis	1463
UC-CO-PL-11	compare.appliances.all.first.period	3
	compare.appliances.all.first.period.wrong.date	13
	compare.appliances.all.second.period	3450
	compare.appliances.all.second.period.wrong.date	13
	compare.appliances.specific.first.period	3
	compare.appliances.specific.second.period	3
	compare.appliances.specific.wrong.date	13
	consumption.energy.appliance.specific.comparison	1361
	consumption.energy.appliances.general.comparison	1533

USE CASE	Intent	Response time (ms)
UC-CO-PL-12	compare.appliances.specific.first.period.wrong.date	13
	compare.general	6
	compare.high_low_consumption	5
	compare.high_low_consumption.analysis	9
	compare.high_low_consumption.more	1652
	compare.high_low_consumption.other.number	8
	compare.high_low_consumption.other.period	3
	compare.high_low_consumption.wrong.date	4
UC-CO-PL-13	consumption.standby	3571
	consumption.standby.period	1406
UC-CO-PL-14	information.schedule.domestic.activities	6
UC-CO-PL-15	advice.menu	8
	advice.menu.aa	3
	advice.menu.dishwasher	3
	advice.menu.dryer	3
	advice.menu.fridge	3
	advice.menu.heating	3
	advice.menu.oven	3
	advice.menu.washing.machine	3
	advice.more	4
	advice.specific.aa	3
	advice.specific.dishwasher	3
	advice.specific.dryer	3
	advice.specific.fridge	3

USE CASE	Intent	Response time (ms)
	advice.specific.heating	3
	advice.specific.lighting	3
	advice.specific.oven	3
	advice.specific.washing.machine	3
	advice.user.knows	10
UC-CO-PL-16	compare.total.consumption.event	650
UC-CO-PL-17	compare.total.consumption.event.periods	5
	explanation.bill	3
UC-CO-PL-18	consumption.environment.general	7
	consumption.environment.general.another.period	1487
	consumption.environment.general.another.unit	8
	consumption.environment.general.direct	7
	consumption.environment.general.unit	11
	consumption.environment.general.wrong.date	7
UC-CO-PL-19	compare.total.consumption.event.more	718
	depreciation.compare.new	6
	depreciation.general	5
	depreciation.general.devicesNew	4
	depreciation.general.dont.know	3
	depreciation.general.new	4
	depreciation.oldComparison.new	9
UC-CO-PL-20	information.pv.installation	3
UC-CO-PL-21	compare.total.consumption.event.periods.get	1326
	depreciation.general.oldComparison	4
	depreciation.general.oldComparison.old	4

USE CASE	Intent	Response time (ms)
	depreciation.general.oldComparison.old.new	4
UC-CO-PL-22	set.goal	78
	set.goal.direct	4
	set.goal.value	4
	set.goal.value.get	7
	set.goal.value.get.cancel	4
	set.goal.value.get.confirmation	883
UC-CO-PL-23	settings.change	6
UC-CO-PL-24	appliance.survey	662
UC-CO-PL-25	time.diary	2
Feedback	advice.useful	1484
	advice.useless	2180
	advice.useless.reason	850
	advice.useless.reason.when	833
General	cancel.general	6
	chat.dissatisfied.customer	3
	chat.mood	4
	explanation.cups	3
	explanation.distributor.contact	3
	explanation.distributor.role	3
	explanation.ecobot	3
	explanation.fallback.kindof	3
	explanation.meter.wrong	3
	Goodbye	
	Greetings	2

USE CASE	Intent	Response time (ms)
	greetings.bad	2
	greetings.fine	2
	leave.company	6
	present.bot	5
	smalltalk.appraisal.good	3
	speak.languages	3
	thank.bot	2
	Unsupported	3
	want.to.save	3
	welcome.message	2
Menu	appliances.menu	5
	change.contract.menu	8
	change.menu	7
	compare.appliances	5
	consumption.general	6
	contract.menu	3
	explanation.high.bill	6
	monitoring.menu	5
	recommendation.menu	7
	save.event	5

Table 84: Task completion time for facility managers

USE CASE	Intent	Response time (ms)
UC-FM-PL-01	save.mvp	682
	save.mvp.building	2461

USE CASE	Intent	Response time (ms)
	save.mvp.confirm	1630
	save.mvp.date	99
	save.mvp.event	1158
	save.mvp.event.category	61
	save.mvp.event.other	10
UC-FM-PL-02	save.recommendation	816
	save.recommendation.event	2068
	save.recommendation.more	904
UC-FM-PL-03	check.gateway	2299
	check.gateway.location	824
	check.gateway.sensors.final	905
	check.gateway.sensors.question	735
UC-FM-PL-04	consumption.total	2450
	consumption.total.add.building	1003
	consumption.total.add.building.question	101
	consumption.total.final	1199
	consumption.total.period	104
UC-FM-PL-05	consumption.appliances.period	976
	consumption.appliances.sensors.add	740
	consumption.appliances.sensors.question	1005
	consumption.appliances.buildings	1262
	consumption.appliances.final	2760
	consumption.appliances	5

USE CASE	Intent	Response time (ms)
UC-FM-PL-06	compare.consumption.portofolio	748
	compare.consumption.portofolio.period	5
	compare.consumption.portofolio.ranking	1283
UC-FM-PL-07	compare.consumption.event	671
	compare.consumption.event.period	5
	compare.consumption.event.period.get	682
UC-FM-PL-08	consumption.cost	884
	consumption.cost.final	1262
	consumption.cost.period	5
	consumption.cost.add.building	1003
	consumption.cost.buildings.question	98
UC-FM-PL-09	consumption.environment	5
	consumption.environment.add.building	1101
	consumption.environment.buildings.question	99
	consumption.environment.final	2780
	consumption.environment.period	5
UC-FM-PL-10	check.sensors	905
	check.sensors.buildings	1005
	check.sensors.period	886
	check.sensors.sensors.question	801
UC-FM-PL-11	check.contracts	1105
UC-FM-PL-12	set.goal	963

USE CASE	Intent	Response time (ms)
	set.goal.building	7
	set.goal.value	5
	set.goal.value.get.cancel	5
	set.goal.value.get	6
	set.goal.value.get.confirmation	745
UC-FM-PL-13	settings.change	5
check.menu	check.menu	4
Feedback	advice.useful	1484
	advice.useless	2180
	advice.useless.reason	850
	save.event.recommendation.push	8
General	Cancel	1095
	compare.menu	5
	consumption.menu	5
	Greetings	2
	present.bot.menu	5
	save.event.menu	5
	Thanks	3
	welcome.message	5

ANNEX C: Detailed analysis of Metric: Quality of new energy feedback and itemized billing practices

P7[18] assesses the suitability of the time diary feedback we receive from the pilots. This was reported in D4.5 in relation to P6, and is further discussed here to explain metric P6.

The figures below, namely Figure 42, Figure 43, Figure 44, Figure 45, Figure 46, and Figure 47 show the six instances where time diaries might have been recorded inaccurately with human error when we look at the smart meter data readings for the day the diaries were reported. The highlighted bars in the plots indicate the hourly slots reported by the time diaries overlaid on the smart meter readings.

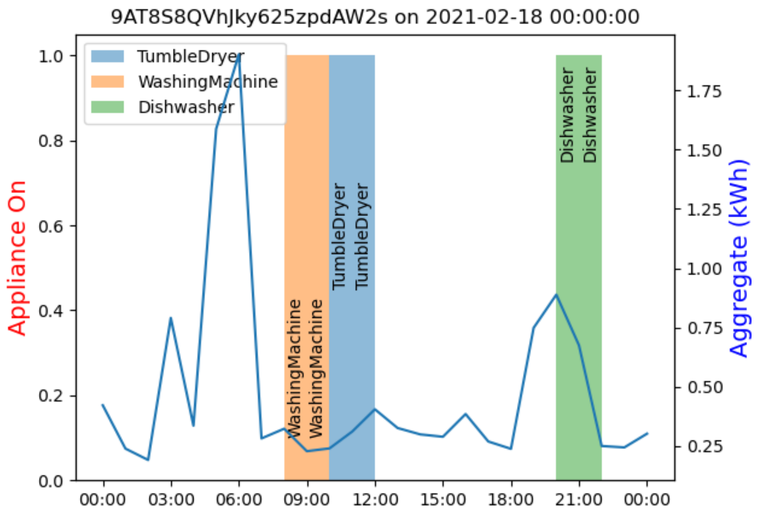

Figure 42: Correct detection of tumble dryer and dishwasher for hourly EYPESA user (missing washing machine due to peak)

[18] Quality of new energy feedback and itemized billing practices based on energy disaggregation and integrating user feedback into NILM algorithm design

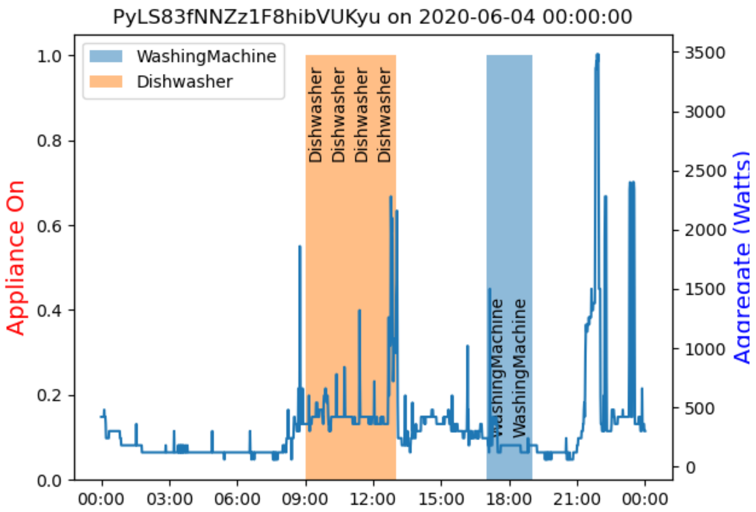

Figure 43: Example of incorrect time diary entry for washing machine

In this case, the dishwasher was missed due to lower heating wattage than encountered in the same EYPESA high frequency user.

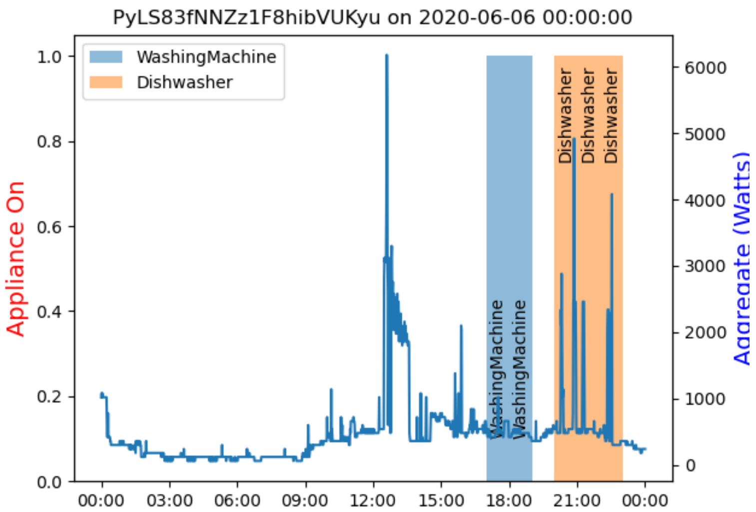

Figure 44: Time diary indicates unusual signature

In this case, the signature is not corresponding to a typical washing machine signature for a high frequency EYPESA user.

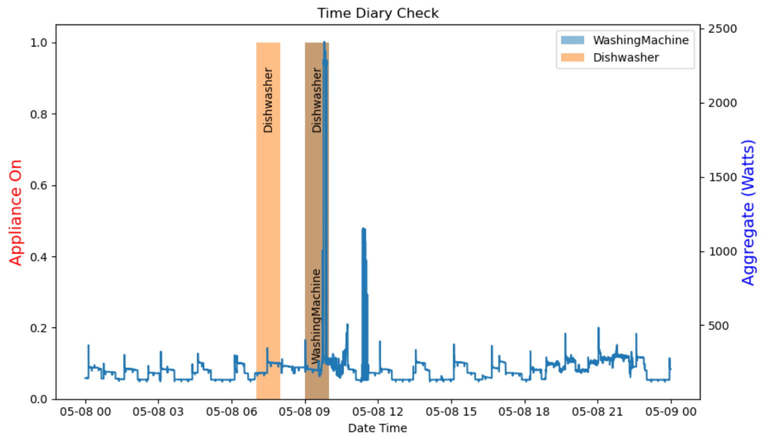

Figure 45: Incorrect time diary entry around 7am and 9am for dishwasher.

However, in this case the washing machine was detected correctly for this SEnerCon user.

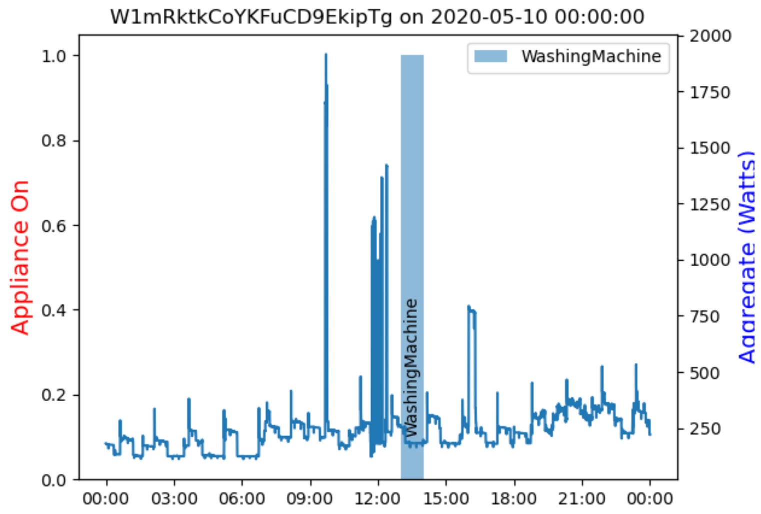

Figure 46: Example of possible incorrect time diary entry for washing machine for SEnerCon user

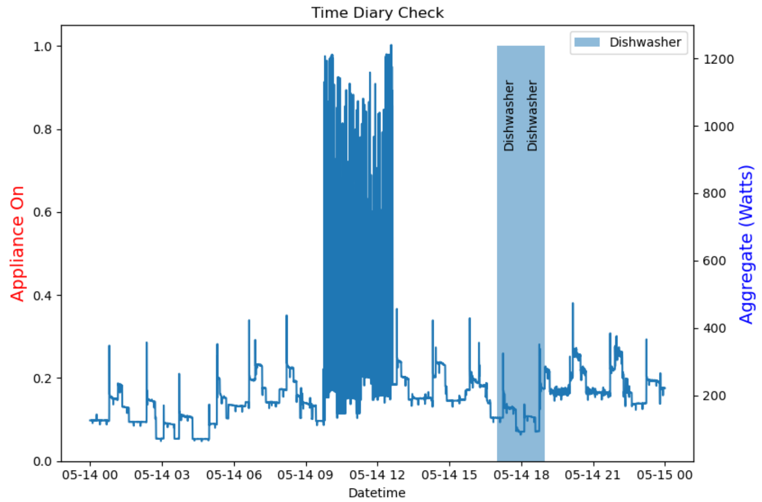

Figure 47: Incorrect time diary entry for dishwasher for SEnerCon user

ANNEX D: Detailed analysis of Metric: Learning performance of NILM algorithms

P6[19] (related to KPI_4.4) was evaluated during Task 4.2 and reported in D4.2 on public datasets, as well as during Task 4.5 (reported in D4.5 for the small-scale validation). The final P6 evaluation was carried out as part of the final demonstration phase in WP5. This metric, namely classification accuracy (the accuracy in appliance detection), is available for the residential pilot participants who fill in time diaries. This measure must be considered together with metric P7. P6 is analysed here for the first evaluation phase 01 April 2020 to 31 January 2021, as well as the second evaluation phase 01 February 2021 to 31 May 2021.

Referring to Table 4, UID K2SjajF93vtBCT7kcQSesD resulted in 4 FPs, likely caused by prolonged periods of consumption. This is well above what we had expected to see in a residential building (Figure 52 and Figure 53). FPs present in E-LF houses are more likely due to the nature of 1 hour data granularity, especially where consumption was not fluctuating much during that day and hence it was impossible for the NILM model to identify unique activations of an appliance and tended to overestimate consumption (Figure 54, Figure 55, Figure 56, Figure 57, Figure 58, Figure 59, Figure 60). Further false positives for UIDs NWT (Figure 61), PYL (Figure 62), and W1 (Figure 63) are due to model errors where similar signatures have caused false positives, which are an expected issue with NILM.

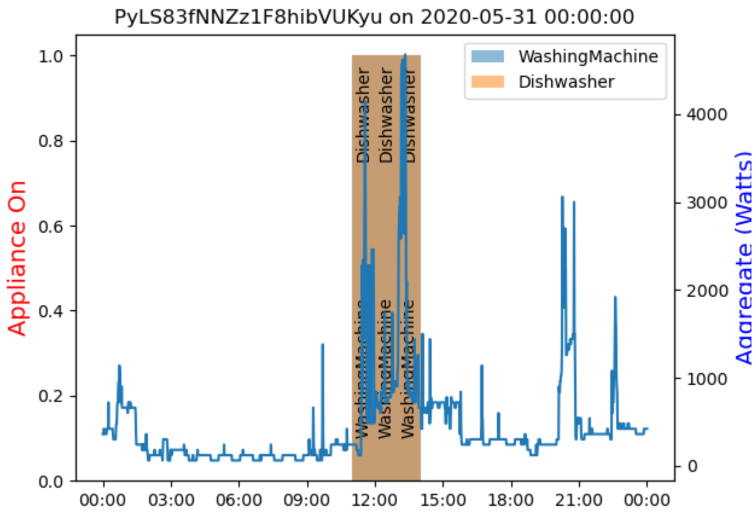

Figure 48: Washing machine is missed by NILM for EYPESA high frequency user [1]

[19] Learning performance of NILM algorithms with regard to appliance signatures from user feedback a.k.a classification accuracy

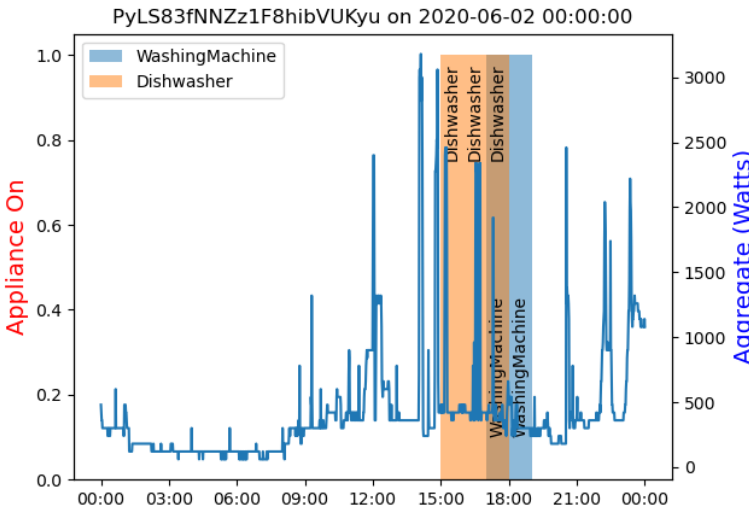

Figure 49: Washing machine is missed by NILM for EYPESA high frequency user [2]

As in figure 15, the washing machine being missed was possibly caused by the dishwasher being run at the same time.

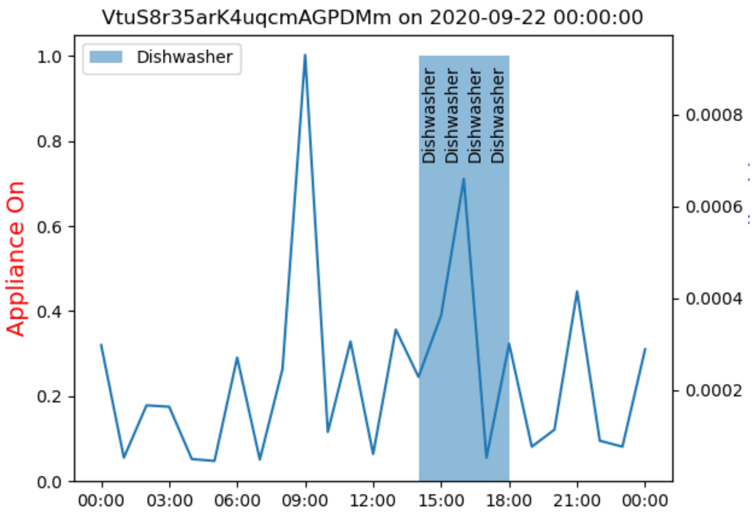

Figure 50: Dishwasher is missed for EYPESA low frequency user

In this case, the missed dishwasher was likely missed due to another large load operating at the same time.

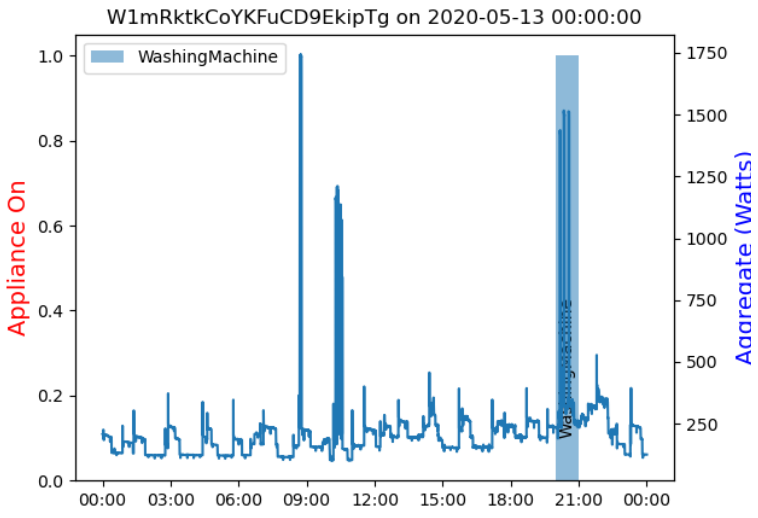

Figure 51: Washing machine is missed for SEnerCon user

Here, the very short duration signature of a seldom used washing machine cycle caused the washing machine being missed for this SEnerCon user.

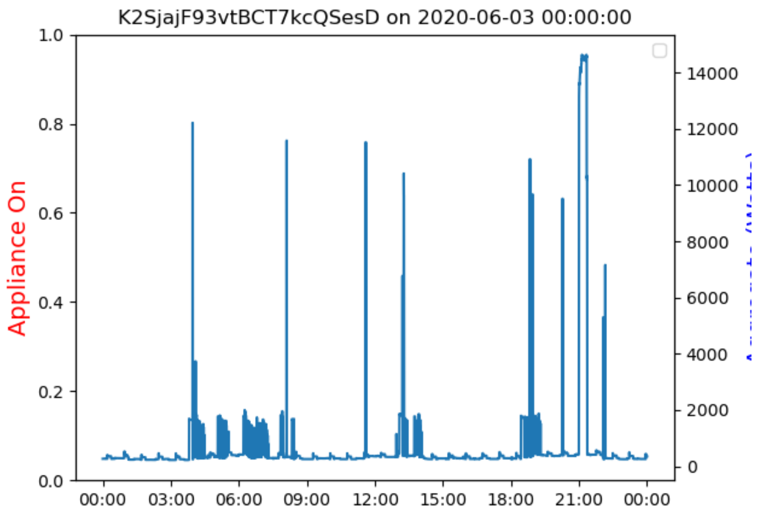

Figure 52: Washing machine and dishwasher detected but not recorded in time diary [1]

Especially around 9pm both appliances were detected but were not recorded in the time diary for this SEnerCon user.

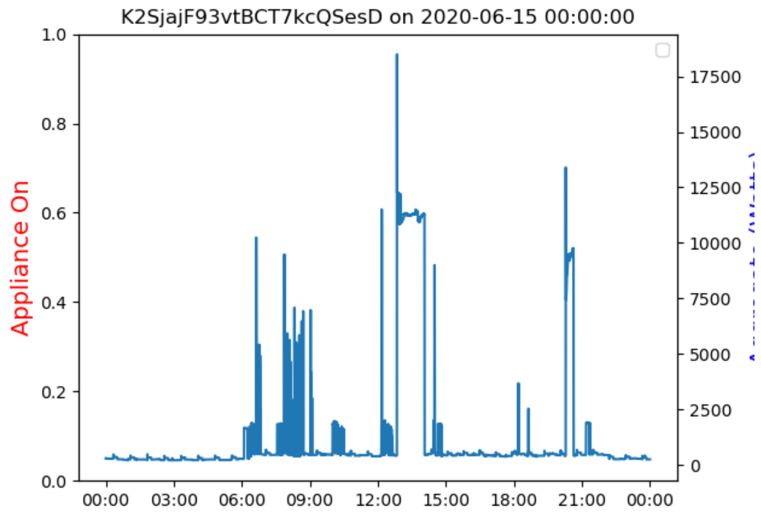

Figure 53: Washing machine and dishwasher detected but not recorded in time diary [2].

In contrast to Figure 19, here the detection peaks around midday, but is similarly not recorded.

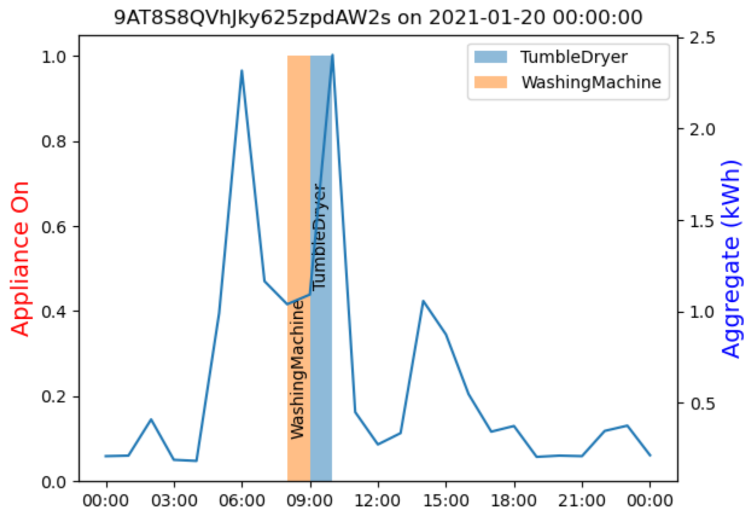

Figure 54: NILM detected combination of washing machine and tumble dryer usage as dishwasher for low frequency EYPESA user

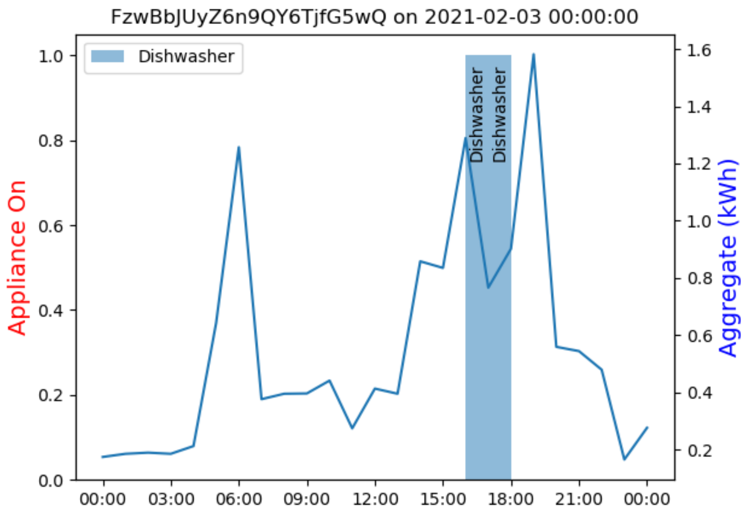

Figure 55: NILM detected tumble dryer and washing machine instead of dishwasher in the evening for low frequency EYPESA user

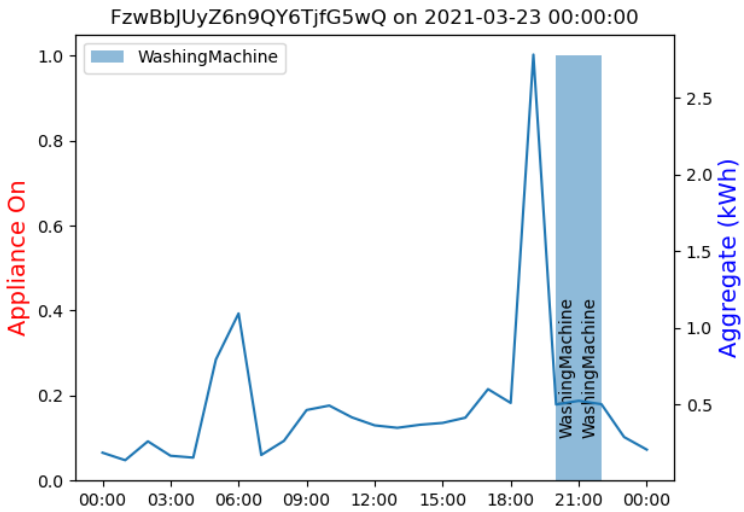

Figure 56: NILM detected tumble dryer and dishwasher as well as washing machine for low frequency SEnerCon user

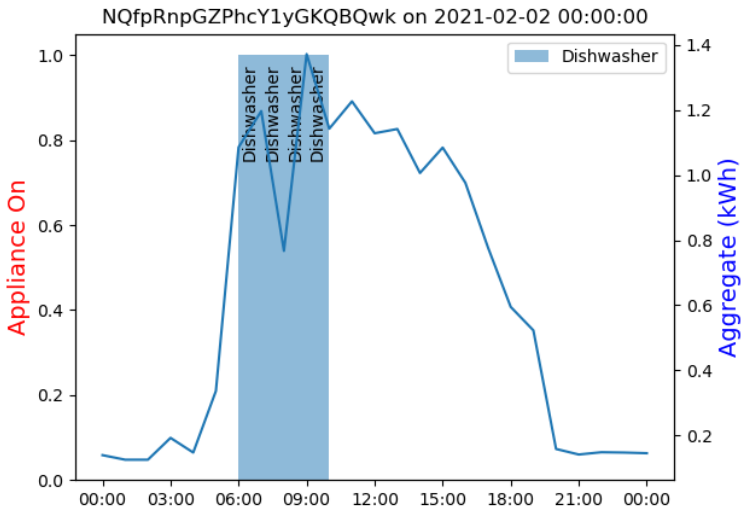

Figure 57: NILM detected washing machine instead of dishwasher for low frequency SEnerCon user

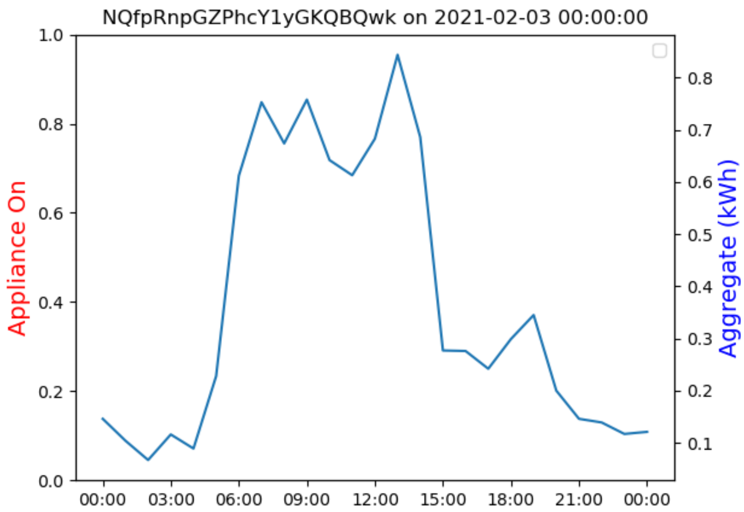

Figure 58: NILM detected dishwasher and washing machine, although low frequency SEnerCon user had no entries

Similar to Figure 24, Figure 25 shows that the NILM detected both dishwasher and washing machine, although this SEnerCon user had no entries.

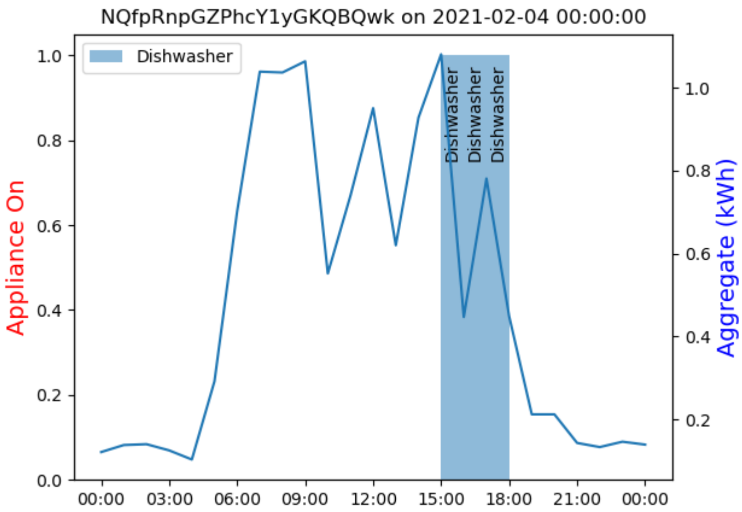

Figure 59: Washing machine detected instead of dishwasher for low frequency EYPESA user

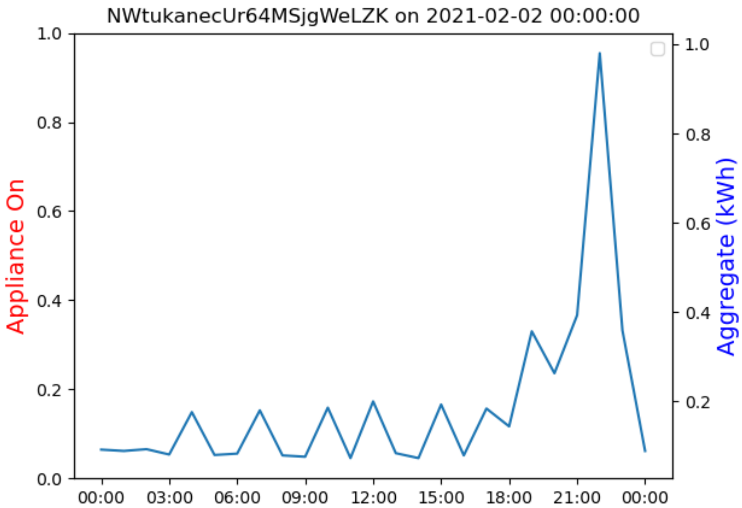

Figure 60: Washing machine and dishwasher detected at 7pm and 10pm for low frequency EYPESA user without corresponding time diary entry.

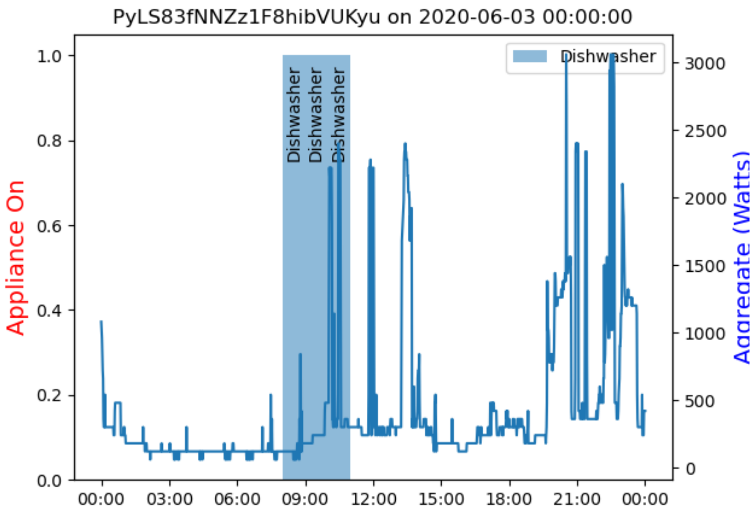

Figure 61: Washing machine detected at midday instead of dishwasher for high frequency EYPESA user

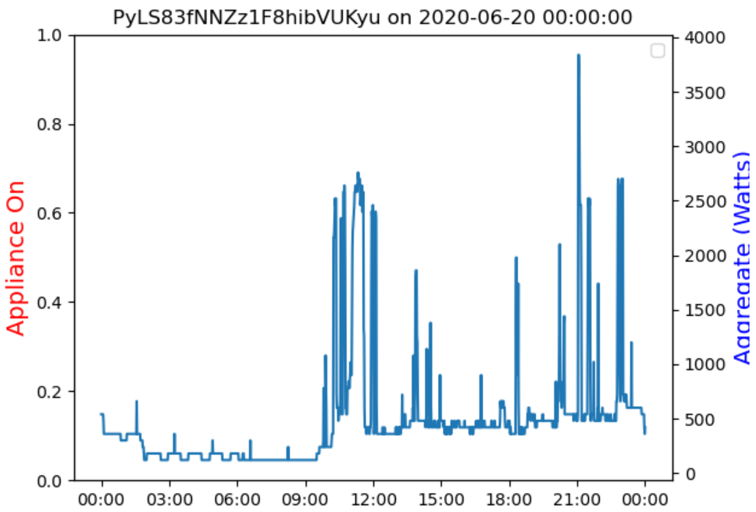

Figure 62: NILM detected both dishwasher and washing machine for EYPESA high frequency user

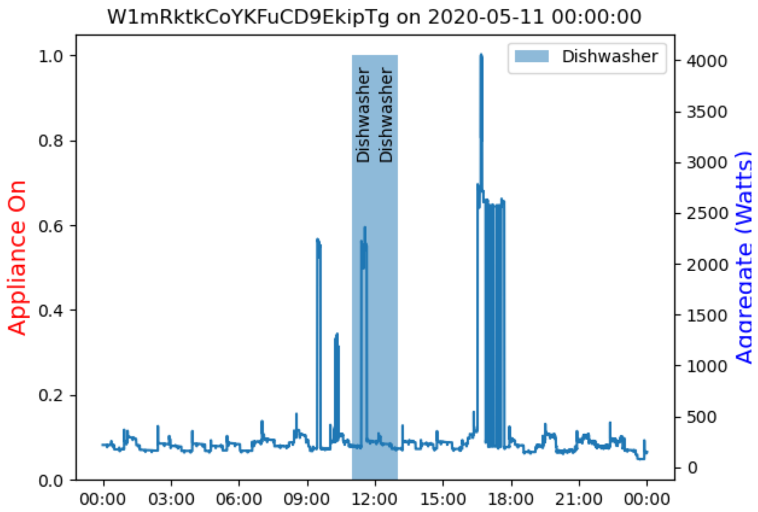

Figure 63: Washing machine detected between 9AM spike and 12PM spike, where dishwasher is recorded by time diary for SEnerCon user

W1MRKTKCOYKFUCD9EKIPTG had a low accuracy rate. This is because the washing machine in this house had a different power draw to the average washing machine load profile that the NILM model was trained on. The washing machine load consumption was considerably lower and appeared to miss the distinctive "heating" phase at the beginning of the duty cycle which can be seen in Figure 61 It also had a much longer profile with a steadier low power consumption.

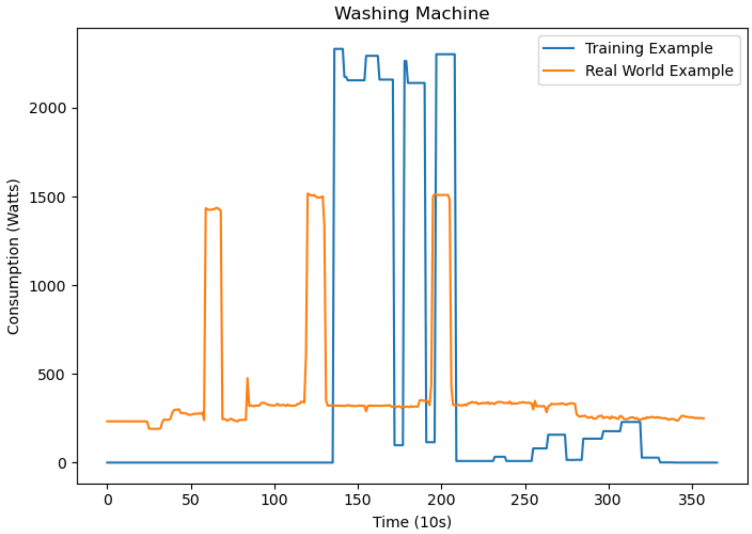

Figure 64: Comparison of training data for washing machine and equivalent SEnerCon data on 13/05/2020.

This visualization explains why this appliance was not detected.

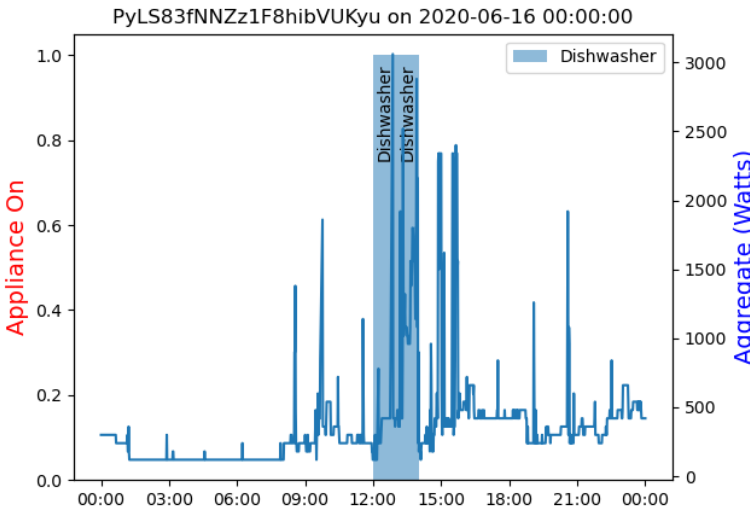

Figure 65: Missed dishwasher for high frequency EYPESA user

The previous visualizations show that the F1-score standard classification accuracy or P6 metric – as calculated by TP / (TP+ 0.5FP +0.5FN) – is 68% for the evaluation phase of April 2020 until May 2021. This is equivalent to 70% accuracy for LF households, and 65% for HF households. However, due to the reliability of time diaries, as discussed in section 0, this metric value cannot be taken as true accuracy of the NILM. Accounting for these FPs, a more realistic accuracy metric value of 72% is obtained over 42 time-diaries. In comparison, for the small-scale evaluation as reported in D4.5, we reported 25 TPs, 4 FNs and 3 FPs, resulting in an F1-score classification accuracy of 88% over 16 time-diary entries.

List of References

AIMultiple (2019, January 1): 15 Key Metrics for Chatbot Conversational Analytics in 2019. Retrieved on July 22, 2019 from https://blog.aimultiple.com/chatbot-analytics/.

ChatbotPack (2019). Why Chatbots Need Quality Standards? Retrieved on July 4, 2019 from https://www.chatbotpack.com/chatbot-quality.

Discover.Bot (2019, January 1): Improving Your Chatbot Performance Through A/B Testing. Retrieved on July 2, 2019 from https://discover.bot/bot-talk/improving-chatbot-performance-through-ab-testing/.

Lee, J. (2018, June 28): The practical guide to chatbot metrics and analytics. Growthbot. Retrieved on July 4, 2019 from https://blog.growthbot.org/the-practical-guide-to-chatbot-metrics-and-analytics.

Neff, A. (2019, February 12): Metrics uncover and track the value of chatbots within customer service. ICMI. Retrieved on July 18, 2019 from https://www.icmi.com/Resources/2019/metrics-uncover-the-value-of-chatbots-within-customer-service.

Newlands, M. (2017, June 13): 10 Metrics Your Chatbot Should Track to Optimize User Experience. Entrepreneur Media, Inc. Retrieved on July 2, 2019 from https://www.entrepreneur.com/article/295447.

Radziwill, N. M., & Benton, M. C. (2017). Evaluating quality of chatbots and intelligent conversational agents. Software Quality Professional, 19(3), 25-36.