35th International Electric Vehicle Symposium and Exhibition (EVS35)
Oslo, Norway, June 11-15, 2022

"Charging process chain": Holistic approach for identifying weaknesses in the electric vehicle charging ecosystem

Felix Roeckle[1] (corresponding author), Marco Amorim[1],
Lukas Keicher[2], Mira Kern[2], Jens Rupprecht[2]

[1]*Fraunhofer Institute for Industrial Engineering IAO, Nobelstr. 12, 70569 Stuttgart, Germany,* *felix.roeckle@iao.fraunhofer.de*

[2] *University of Stuttgart, Institute of Human Factors and Technology Management, Stuttgart, Germany*

Summary

As part of the "Wirkkette Laden" funding project, a holistic approach for analysing weak points in the charging ecosystem was developed and applied in collaboration between research and industry. The approach is based on four building blocks: Statistical data analysis, a diary study, a charging experiment, and an AI-based analysis of online user comments. The results show that there are challenges both in terms of the individual components of the charging ecosystem, such as charging stations, and in terms of the interfaces between the components within the system.

Keywords: charging; consortium; market development; mass market; user behaviour

1 Motivation & prior research

Charging faults and their causes can only be solved quickly if all market participants (charge point operators, e-mobility service providers, vehicle manufacturers, charging station manufacturers) work together in tracing down weak points and links in the ecosystem. To address this challenge, the German Federal Ministry for Digital and Transport (BMDV) has funded the research project "Wirkkette Laden", the overall objective of which is to look at the complete charging chain, considering all above-mentioned market participants and focusing on user-friendly, successful charging of the vehicle.

Numerous previous research papers have addressed electric vehicle charging and the challenges associated with it. For instance, in a previous EVS paper, Röckle and Litauer identified that failed authorization is a major problem for EV drivers [1]. Since then, EV adoption has increased significantly and with it the charging problems, as current studies show [2,3,4]. However, the authors are not aware of any research that specifically analyses the root causes of erratic charging sessions. Travel diaries have for sure been used in e-mobility research, however not focusing on charging errors [5,6,7]. Natural language processing analysis of user comments has been used to identify pain points from a customer perspective, e.g. in the airline industry [8], but there is no published work in the domain of EV charging. Finally, traditional statistical methods as well as machine learning algorithms are known from the quality management literature [9], but to the best of the authors' knowledge, they have not been applied to the charging process of electric vehicles. Accordingly, a combination of all or parts of the methods mentioned is not known either.

2 Objectives

The overall goal of the project "Wirkkette Laden" is to pinpoint issues that need to be addressed to improve the quality of service for charging electric vehicles. For this purpose, firstly, a "big picture" has been developed to provide a common understanding and base for discussion. This big picture includes components (physical, IT back-end, IT front-end) as well as the interfaces (energy flow, data flow) between them. The second step in the project is the holistic analysis of charging errors and problems to identify weaknesses in the ecosystem, e.g. frequent problems with remote authentication. These supporting research and analysis activities are the core of this paper and are described in detail in the following section. In the final step of the project, the results of this applied research should then be mapped to the overall picture to show which issues need to be addressed first to improve the quality of charging processes.

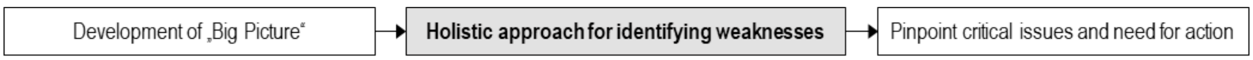

Figure 1: Placement of the present research in the overall project context

3 Methodology

3.1 Overview

To achieve a holistic view on the charging process and charging errors, the research methodology relies on both qualitative and quantitative methods that are applied to data from different sources, namely back-end data exports on the one hand, and user observations and perceptions at the front-end on the other hand (cf. Figure 2).

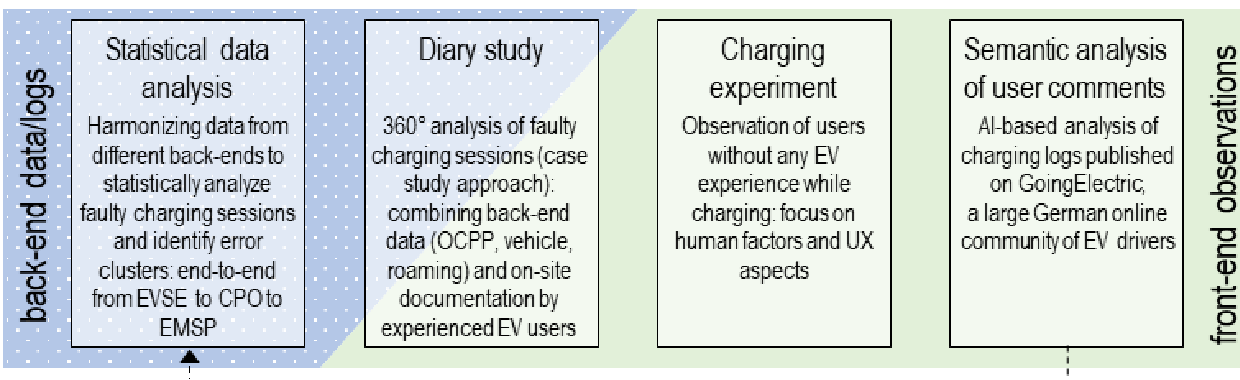

Figure 2: Research methodologies applied in the project

3.2 Statistical data analysis

To identify error patterns, the harmonization of data records derived from different sources (cf. Table 1) is a first crucial step. The way this harmonization is done depends on how the link between charge point operator (CPO) and e-mobility service provider (EMSP) is implemented. The following description relates to a connection based on the Hubject roaming platform and the associated OICP roaming protocol.

Table 1: Data records used for analysis

	provided by CPO	provided by EMSP
Authentication messages		x
RFID requests		x
Remote start messages		x
Remote stop messages		x
Charge detail records (CDR)	x	x
Errors as logged via OCPP	x	

In the first step, remote stop messages are linked to the corresponding remote start message based on the Hubject session-ID. Then, RFID requests and remote start/stop messages are concatenated before both are linked to the CDRs provided by the EMSP (hereafter referred to as EMSP CDR), again based on the session-ID. Then, the entire EMSP record is linked to the CPO records based on an exact match between the charging start timestamp and the EVSE-ID contained in the CPO CDR and EMSP CDR, respectively. The OCPP error data is already associated with the CPO CDR in the input record.

In a second step, linked records ("observations") are aggregated to "charging events" (cf. Figure 3). Thereby each charging event is understood as a user's attempt to charge their vehicle on one occasion. As challenges may occur during authentication and/or during charging, several authentication messages and CDRs (and corresponding errors) may be associated with one and the same charging event. Based on a preceding frequency analysis of the time between requests per user, 40 minutes was selected as a threshold for grouping observations. Thus, if records share the same user-ID or contract-ID and the timestamp of the request lies within 40 minutes after the timestamp of the previous request, they are combined into the same charging event.

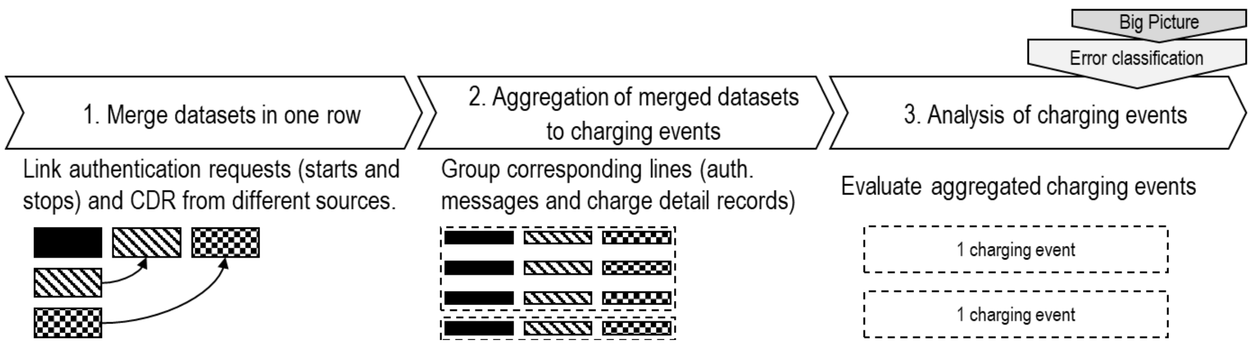

Figure 3: Procedure for data preparation and evaluation

In the third and final step, the charging sessions are analysed by first dividing them into "flawless" and "potentially faulty" charging events before adding additional viewpoints, such as, the number of stations involved in an event, the number of authentication messages, or the presence of CDRs with null kWh of charged energy. In addition, the error messages or error codes associated with each event are considered. For further analysis, an error classification is used (cf. Figure 3) that assigns fault codes to specific elements of the big picture developed in the project, e.g., the charging station (EVSE) or the interface between EVSE and vehicle.

"Flawless" charging events were defined as those that consisted of exactly one authentication message (one RFID or one remote start message) and exactly one CDR with no associated OCPP error messages. The rest of the charging events were categorized as "potentially faulty" and further classified as follows:

- one to one relation of authentication to CDR, but at least one error message: like flawless charging events, with the only difference that here OCPP error messages were present
- multiple authentication messages and/or CDR and no error message: several authentication messages and/or CDRs can be associated with the event, but no OCPP error message has been assigned
- multiple authentication messages and/or CDRs; at least one error message: several authentication messages and/or CDRs can be associated with the event, and at least one OCPP error message has been assigned

3.3 Diary study

While the statistical data analysis described above leaves the actual events in front of charging stations unobserved, the diary study aims to add this perspective by integrating electric vehicle users into the data collection process. To this end, experienced electric car drivers with their own electric vehicles were recruited to document the public charging sessions they perform during their daily driving. Each participant received an individual link to an online survey that was used to document the charging processes - by individualizing the link, it was possible to store the participant's master data (in particular on the vehicle used) in the questionnaire already.

However, it was also possible to specify an alternative vehicle for each diary entry. Participants were asked to document both successful and faulty charging sessions, even though more questions were posed in the case of unsuccessful authentication and/or charging events (cf. Table 2). An additional data collection channel with a simplified questionnaire was implemented via GoingElectric, a large online community and charging station directory for e-drivers in Germany. A link to the survey was included for each directory entry related to the project partners' charging infrastructure.

Table 2: Main questions to assess the success of documented charging events (question numbers for subsequent reference)

Q1 On the first attempt, which of the following authentication methods did you use for the charging process? [*request to specify the provider used in a free text box]	• RFID card / dongle* • EMSP app / e-mobility app* • SMS • Web app of charging station operator (usually QR code on charging station) / Interchange Direct • Payment terminal with card slot and PIN entry (EC/credit card) • Contactless payment terminal w/o PIN entry (EC/credit card, Apple/Google Pay) • Plug & Charge* • Other*	
Q2 Was the authentication method successful?	Yes, the operation was successful.	No, the following problem has occurred: [free text box]
Q3 What authentication method did you use after the first attempt failed?	*See above.*	
Q5 Was the second authentication method successful?	Yes, the operation was successful.	No, the following problem has occurred: [free text box]
Q6 How do you evaluate the success of the charging session performed?	Charging has worked without any flaws.	Charging process only worked imperfectly.
Q7 Which of the deficiencies listed below applies to your situation? Please explain the occurrences within the comment field as precisely as possible.	• The charging session has been interrupted/terminated. • There were challenges/problems starting the charging process. • There were challenges/problems when terminating the charging process. • Maximum charging power lower than expected. • Charging power dropped more than expected during the charging process. • Other shortcomings	

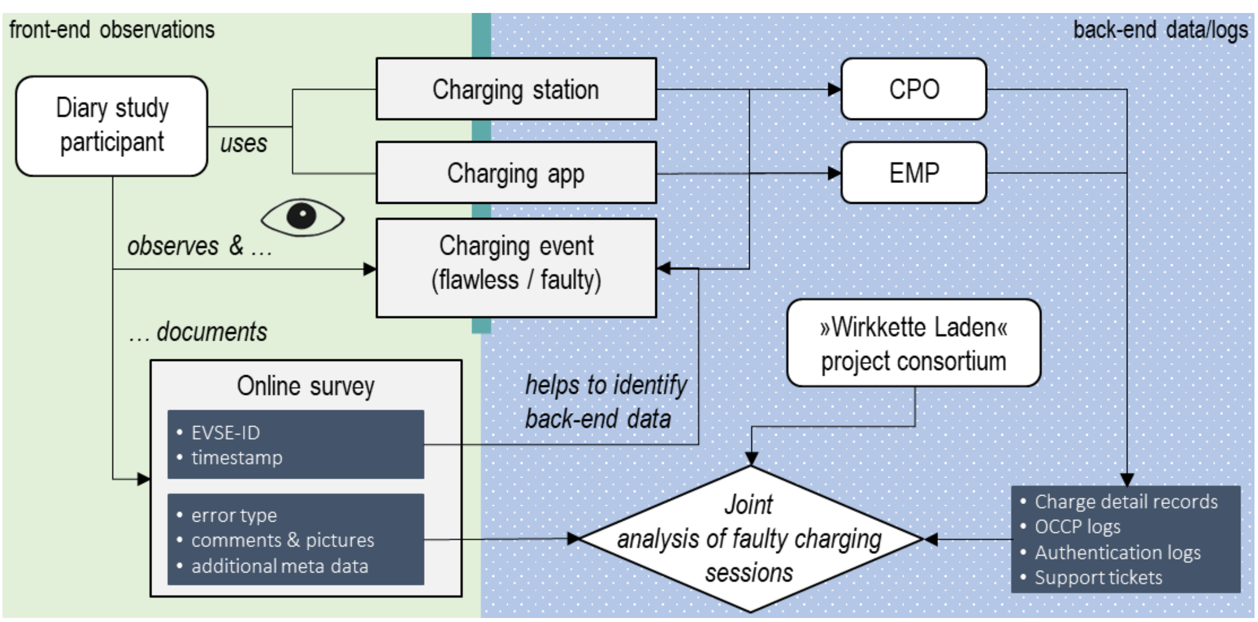

Figure 4: Methodological approach in the diary study

Furthermore, participants were asked to provide the time and date of their charging process as well as additional information to help identify the process later. For this purpose, they had the option of uploading images of the EVSE-ID, screenshots from their smartphone showing a charging app, and the nameplate of the respective charging station. Alternatively, they could also enter the EVSE-ID manually. Other questions were related to contextual information about the charging process, such as weather conditions or whether other vehicles were charging at the same time. Finally, some additional questions were asked about the general satisfaction with the charging process and the condition of the charging station.

Periodically, the recorded diary entries have been exported from the LimeSurvey system that was used to host the questionnaire. After the data had been checked and cleaned by Fraunhofer IAO, it was made available to the EVSE manufacturers, CPO and EMSP involved in the project so that they could localize the relevant processes in their systems and thus perform a joint error analysis within the project consortium. In addition to charging station logs (OCPP logs), authentication message logs and charge detail records, support tickets were also considered (cf. Figure 4). As an incentive for the targeted use of the project partners' products, i.e. a specific charging app and specific charging infrastructure, participants were offered free charging for a limited period of time if these products were used.

3.4 Charging experiment

In addition to the actual technical problems in the process chain of the charging session that were addressed in the diary study, the project also investigated possible challenges in the human-machine interaction that can lead to unsuccessful charging operations. Therefore, people who had no experience with EV charging were recruited for a charging experiment, i.e. they were asked to perform several charging processes at different charging stations using different means of authentication. The test persons were observed during the charging process and their feedback and reaction were constantly reflected and collected via a personal guideline-based interview. For some of the participants, the experiment was conducted with the aid of eye-tracking glasses that closely track the subjects' gaze to detect, for example, deficiencies in the information displayed on the station screens. Fixed cameras were also used to record facial expressions and gestures as well as the general behaviour of the subjects. Relevant observation points were defined in advance (e.g., entering the parking lot, plug attachment) to facilitate the documentation of such key factors. The accompanying survey was based on guiding questions, focusing on identifying missing or misleading information, general challenges in charging station operation, and discussion of (missing) requirements from the non-user perspective. The analysis of the data followed the facilitated process of a qualitative content analysis. First, the information generated from the survey and observation was bundled and reflected. In a second step followed the generalisation in universal challenges as well as individual mentions. Particularly striking scenes or challenges were then validated using eye-tracking recordings and film footage.

3.5 Semantic analysis of user comments

The main goal behind the approach of analyzing user comments for possible error causes lies in the identification of error causes that are not apparent from the information in the operators' back-end. Therefore, the comments of EV drivers, published on GoingElectric (www.goingelectric.de) were analyzed semantically by an Artificial Intelligence. The software used can understand the content meaning of texts and can evaluate large amounts of data in a short time. For this semantic analysis of user comments, the commercial Natural Language Processing (NLP) tool Sherpa was used [10]. This tool allows easy training and application of both established and the latest language models in NLP, e.g., transformer models such as BERT [11], to a custom document corpus.

To identify the error causes, the following approach, based on the information processing process by Greitemann [12], was developed: Firstly, the information needs were defined by the project consortium, in this case the causes of errors from the users' point of view. Secondly, the information was collected, i.e., an export of approx. 13,000 charge logs documents by users of GoingElectric between January 1st 2019 and April 30th 2022 was retrieved. This so-called document corpus includes user comments about both successful and unsuccessful charging processes. The language of the comments is German. Thirdly, the document corpus was analyzed. To make sense of the information, the project consortium collected and structured known error causes as categories. In a first step, 49 categories were defined. However, after first experience made with the algorithm, it turned out that at least 30 but preferably more than 50 example assignments per category are required for a reliable result (meaning a total of 2,450 manually assigned examples would be required in a best-case scenario). As there were not enough examples for all categories in the document corpus, the number of categories was reduced to 21 during two revision loops. These remaining categories include the process dimension (e.g., “start only after several attempts,”), components of the charging system (e.g., “display is defective, but charging works”) but also error causes that can be named directly by the user (e.g., “at 63% charge termination and totally heated plug”). These defined categories were assigned to a couple of hundred comments manually by the project team to train the algorithm. In the fourth step, the algorithm uses an active learning approach and makes suggestions for new assignments based on the learnings. These suggestions require feedback from the user of the software, which makes the algorithm learn even more efficient than assigning categories manually.

To reduce the number of unassigned documents, an additional classifier separating comments about positive experiences from negative comments was trained and added to the first steps of the process. Nevertheless, 100% correct classification of all data is not possible in reality, neither by a human nor by an algorithm [13], as e.g., Kokkinos [14] has shown for the boundary detection case. Thus, there remains a residue of unassigned user comments that either cannot be processed by the algorithm correctly or do not match any of the predefined error causes. To reduce the number of wrong autonomous assignments, the F-Score was introduced as quality gate for the accuracy of the classification made by the system. The F-Score summarizes the precision (= confidence, meaning the proportion of predicted positive cases that are correctly real positives) and recall (= sensitivity, meaning proportion of all real positive cases that are correctly predicted as positive case) values of the algorithm [15] and is defined by their harmonic mean [16]. In this case the boundary was set to a F-Score of 80 %, which is a conventionally used value [17].

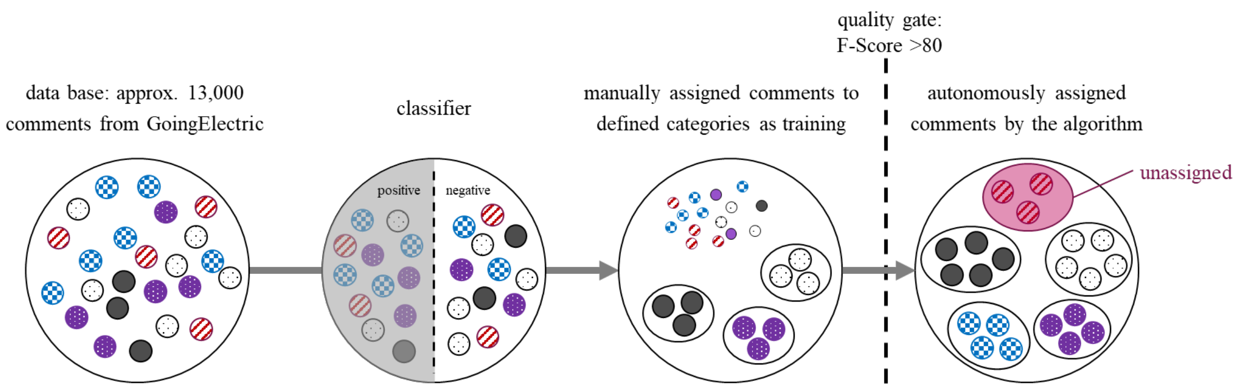

Figure 5: Representation of the process for the semantic analysis of user comments

4 Results

4.1 Statistical data analysis

The basis for the following analysis was a data export for the period 01/01-01/03/2022, which included 46,900 RFID authentication messages, 5,480 remote start messages, and 1,509 remote stop messages relating to one of the CPOs in the project, provided by one of the EMSP companies involved in the project, as well as 50,417 CDRs that were provided by both parties involved (EMSP and CPO). Based on the methodology described above, these input data were aggregated to 42,573 individual charging events, which were then further analysed. As Table 3 shows, charging events with RFID authentication appear to be more inconspicuous, with 87 % of operations rated as "flawless," compared to only 79 % for remote authentication. However, a closer look at the total of 14% "potentially faulty" events reveals that 74 % of the remote events have no error message assigned (cf. Table 4) – thus, the bulk of events is not categorized as flawless due to the existence of multiple authentication messages and/or CDRs.

Table 3: Flawless vs. potentially faulty charging sessions according to statistical data analysis (n=42,573)

	flawless charging events	potentially faulty charging events	total
RFID authentication	87 %	13 %	38,733
remote authentication	79 %	21 %	3,840
total	**86 %**	**14 %**	**42,573**
	36,543	**6,030**	

Table 4: More detailed analysis of potentially faulty charging sessions (n=6,030)

	potentially faulty charging sessions (n=6,030)		
	one:one authentication:CDR but at least 1 error message	multiple auth messages and/or CDRs; no error message	multiple auth messages and/or CDRs; at least 1error message
RFID	7%	52%	41%
remote	8%	74%	18%
total	**7%**	**55%**	**38%**

The analysis of events that do have at least one assigned error message provides further insight regarding the potential causes of errors and the differences between RFID and remote authentication. Table 5 shows how many events with errors involved the different elements of the big picture (based on the error messages and the classification of these), where multiple elements can be assigned to one event. Some error messages could not be classified, so the row sum is not 100 %. Nevertheless, a few trends can be identified: In remote events, there are more problems between the charging station and the CPO back-end, which is quite plausible. For RFID events, on the other hand, more errors relate directly to the charging station itself (57 %). To better understand what is behind these results, a more in-depth analysis is needed and will be conducted later in the project.

Table 5: Evaluation of all events that have error messages and assignment to elements of the big picture (n=2,726)

	driver	vehicle	EVSE - vehicle	EVSE	EVSE - CPO back-end
RFID authentication (n=2,518)	0,2 %	6 %	34 %	57 %	1 %
remote authentication (n=208)	0,5 %	9 %	38 %	46 %	5 %
total (n=2,726)	**0,2 %**	**7 %**	**35 %**	**56 %**	**2 %**

4.2 Diary study

Eventually, 57 persons have been invited to participate in the study, covering 18 different makes of electric vehicles. As of April 19, 31 registered participants have documented at least one charging session. A total of 250 charging sessions (42 via GoingElectric and 208 from registered participants) were documented between January 1 and April 18, 2022, of which 76 % were classified as successful (Q6), although 5 % had authentication issues (Q5), cf. Figure 6. Accordingly, 60 charging sessions (24 %) were classified as faulty.

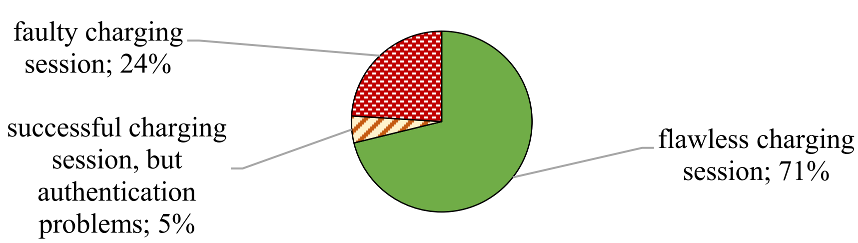

Figure 6: Evaluation of documented charging sessions (Q5 x Q6 | n=250)

Of the 72 charging sessions in which a problem occurred, the project partners were involved in 29 (cf. underlined values in Table 6) - these are the charging sessions that were subjected to a joint error analysis in the consortium. In about 10 of these sessions, a charging station fault was identified as the cause of the error, and in a few others, the problem could not be traced at all. Select case studies that shed light on some weaknesses in the ecosystem are discussed below.

Table 6: Success of documented charging sessions and involvement of project partners

	project partner involved	no project partner involved	**total**
flawless charging session	127	51	**178**
successful charging session, but authentication problems	<u>8</u>	4	**12**
faulty charging session	<u>21</u>	39	**60**
total	**156**	**94**	**250**

4.2.1 QR codes are not standardized

In several cases, EV drivers were not able to scan the QR code affixed to the charging station. The reason was found in the discussion within the consortium. The e-mobility service provider searches the URL for the string “evseid=” followed by the EVSE-ID of the respective charging station in the format specified by BDEW [18]:

<EVSEID> = <Country Code> <S> <EVSE Operator ID> <S> <ID Type> <Power Outlet ID>

However, one of the charge point operators in the project does use a different format for its charging stations. The EMSP app was thus not able to read the QR code. If the URL stored in the QR code always contained a standardized character string, e.g. because this is mandatory by national or international charging infrastructure regulations, this problem could be improved.

4.2.2 POI data is inaccurate

Certainly a classic in EV charging research and likely a regular part of the EV driving experience, inaccurate POI data, i.e. charging station meta data, has also been identified as a problem in at least one charging session documented in the diary study. The used EMSP app displayed a particular charging station as open 24/7. In fact, however, the parking lot where the charging station was located was closed at night, so the driver of the electric car who documented the charging process parked their car outside the parking lot and blocked the sidewalk next to it. The analysis of this problem in the consortium revealed that the (restricted) opening hours were indeed properly maintained in the back-end, but the 24/7 attribute was not set correctly.

4.2.3 Swapped cables go unnoticed by the charging station

In one case, it was indeed the human-machine-interface that caused some trouble for the EV user. The charging station used had two CCS plugs located directly next to each other. When the diary study participant wanted to use the station, they selected one of the two available plugs and started the session with the charging app, which all went fine. The socket of the plug that was selected lit up to indicate that this cable was to be used. However, charging did not start. The reason for this was that the two plugs were interchanged, i.e. plug 1 was plugged into socket 2 and plug 2 into socket 1 - thus, the coloured highlighting did not help, as only the socket was illuminated and not the plug itself. The EV driver finally noticed this themselves, but the charging pause had been extended from 30 to 40 minutes as a result.

4.3 Charging experiment

Before the deadline for submission of this paper, one of the two planned charging experiments could be carried out. The focus of the first experiment, conducted on April 7-8, 2022, was on fast-charging infrastructure; the demographic composition of the sample is shown in Figure 7. As the first performance of the experiment showed, there are only a few hurdles that cause a charging session to be unsuccessful, but this does not mean that there is nothing to improve about the charging experience from the subjects' point of view. In general, it can be stated that initially most test persons compare charging to fuelling an ICEV and use this process as a reference. However, a learning effect sets in quite quickly, so that a certain routine is established by the third or fourth charging process at the latest. Nevertheless, the experiment revealed some shortcomings that can affect the user experience to a greater or lesser extent, or even cause a failure in charging an EV. Initial findings are briefly described below.

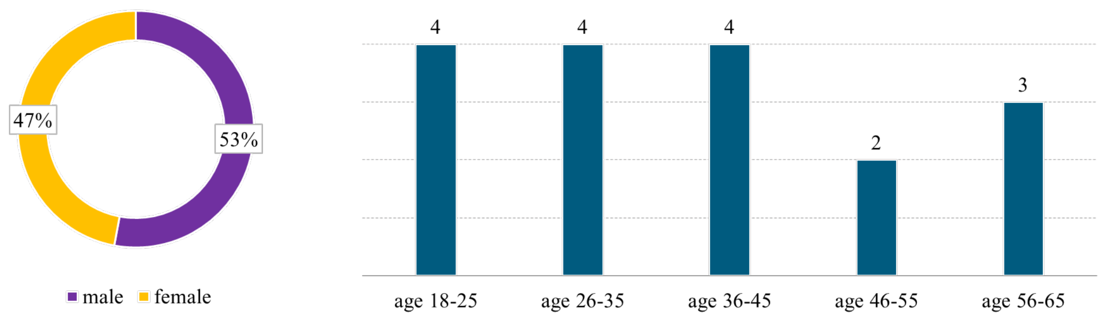

Figure 7: Demographics of the sample of the first iteration of the charging experiment (n=17)

Actual reasons for non-successful charging

- **Remote authentication (native smartphone app or web-app):** In the experiment, it happened that a timeout occurred from the vehicle side because the plug was first inserted into the vehicle and authentication (e.g., entering credit card data) took longer than anticipated.
- **Insufficient feedback from the charging station:** In a few cases, especially when paying directly via a web app, the test subjects were lost because neither the charging station nor the (web) app told them what was happening and what the next step was.

User experience issues

- **Clear instructions:** Subjects found it very helpful when clear step-by-step instructions or clear instructions on the following steps were given, e.g. step 1-2-3 explanations at the charging station.
- **Feedback from the charging station:** Even if this does not lead to an unsuccessful charging process, it has been shown that unclear feedback leads to dissatisfaction among the test subjects. Each successful step, e.g. plugging in the vehicle, authentication, starting the charging process, ending the charging process or returning the cable to the charging station, should be confirmed by the charging station, ideally by a combination of acoustic and visual feedback. In addition, the status and progress of the charging process should be visible at all times (see the following bullets).
- **Price information:** Accustomed to direct feedback on the current price when refuelling a conventional vehicle, the test subjects also expected immediate price information at the charging station. One person even raised the question of whether it would be possible to set a final price (e.g., €20) for charging - which may sound surprising to experienced EV drivers, but is an interesting perspective, nonetheless.
- **Remaining time to the desired SoC (state of charge):** Almost all the subjects wanted to know how long it would take for their vehicle to fully charge, or how long it would take for them to charge enough to complete their planned trip. The authors are aware that some charging stations show the remaining charging time to 80 % SoC, but the stations used in the experiment only displayed the current SoC in % and the elapsed charging time.
- **Technical terms and abbreviations:** Additional technical information shown on the charging station display, such as current and voltage, does not provide any significant added value for most test subjects. In addition, technical abbreviations, such as "A" for ampere or "V" for volt, cause confusion.
- **Consistency of displayed information:** Some test persons noticed the discrepancy between the SoC shown in the charging app and the SoC shown on the charging station display and were confused by this.

- **NFC vs. QR codes:** Finding the QR code scanner in the smartphone, finding the QR code at the charging station, and scanning the QR code with the smartphone generally seemed to be significantly more difficult than using an RFID card for authentication. Moreover, it even seemed more natural to simply swipe the smartphone over the card reader, as is familiar from NFC-enabled credit cards or mobile payment systems (e.g. Apple Pay).

4.4 Semantic analysis of user comments

Applying a manually trained classifier (based on Support Vector Machine, F-Score = 81.3 %) to the database of approximately 13,000 user comments resulted in 5,852 comments that were classified as "negative". This means that 46 % of all user comments on the charging processes were not satisfactory or even reported failed charging processes in the period studied. In a second step, the manually trained language model (based on a conditional random field algorithm and limited to a focus on error categories that were trained with a minimum number of 30 examples each – a total of 784) was used to categorize these negative user comments into defined error categories. By applying the language model - trained on both negative and positive comments - solely on negative comments, the F-score could be increased to a value of 90.8 %. With this F-score, the language model autonomously assigned the negative subset of user comments to the defined error categories. The application of the language model resulted in 2,036 (of 5,852) autonomously assigned user comments by the algorithm.

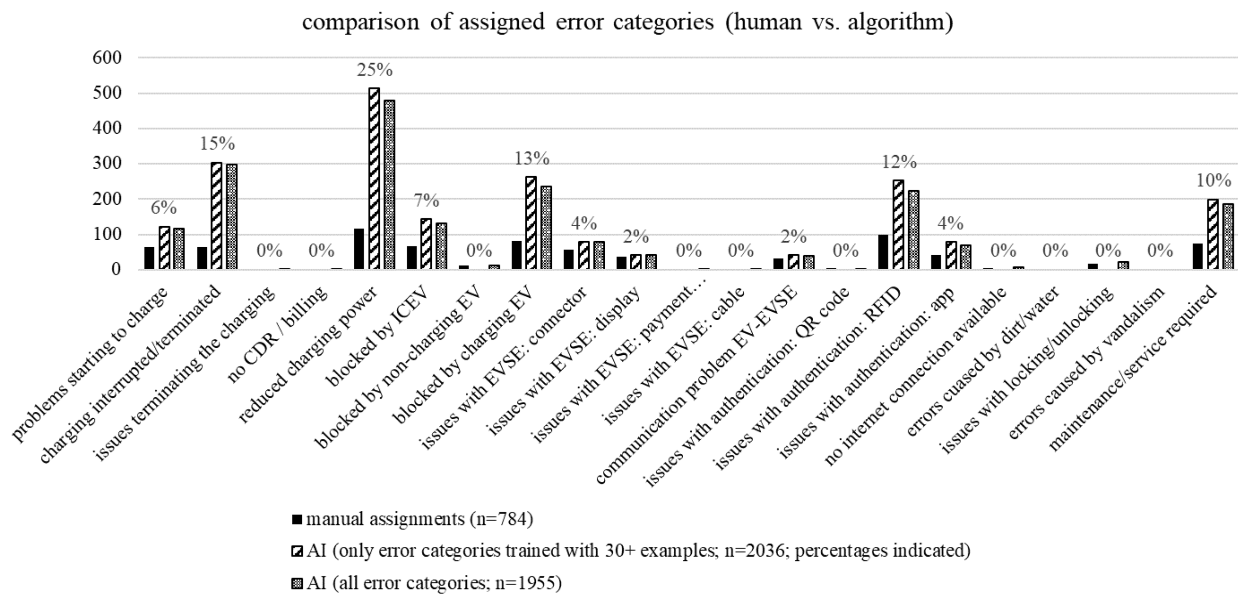

Figure 8: Comparison of assigned error categories (manual/human vs. algorithm/AI)

Looking at the distribution of the manually assigned classification and the assignments by the AI, a similar picture emerges (cf. Figure 8). If the 784 manually assessed comments are reasonably representative, it can be deduced that the language model can reliably categorize faults. To validate the quality of the language model a second language model containing all error categories (not only the one that were at least trained by 30 manual assignments) was applied to the data set. As expected, this second model was able to classify only 1,955 user comments while also having a slightly lower F-score (90.1 %).

The results of applying the optimized language model to user comments from EV drivers for semantic analysis of frequency distribution of perceived failure causes from the users' perspective are shown in Figure 8 (striped bars). The top three ("reduced charging power", "charging interrupted/terminated", "blocked by charging EV") do not come as a surprise. It is noticeable, however, that authentication with RFID cards is mentioned comparatively frequently as the cause of errors. There may be two reasons for this: Firstly, the number of charging sessions in which an RFID card was used is significantly higher (this is suggested by the statistical data analysis, cf. section 4.1), which is not accounted for in this analysis of user comments. Therefore, even though the absolute number of user comments with RFID authentication problems may be higher, there may be fewer problems relative to the total number of charging transactions than with other authentication methods. Secondly, the F-Score given above represents an average across all 21 error categories - but the F-Score for the error item "issues with authentication: app" is only 65.8 % (lowest value of all error causes). Therefore, a significant number of comments related to app authentication issues may not be assigned by the language model.

5 Discussion

5.1 Conclusion

In the "Wirkkette Laden" research project, a holistic approach to identifying weak points in the electric vehicle charging ecosystem is being tested. The initial results of the different research approaches applied in the project and presented in this paper show that EV charging is still far from providing a perfect service. Depending on the source and method, the percentage of error-free loadings ranges from 54% in the semantic analysis of user comments to 71% in the diary study to 86% in the statistical analysis of back-end data. In the case of user comments, however, it can be assumed that there is a bias due to self-selection, i.e., negative experiences tend to be recorded in charge logs that were the basis for the analysis, while positive experiences are not documented.

Moreover, this research demonstrates some insights into the types of problems when charging and causes of errors. From the user perspective, problems starting the charging session and lower than expected charging power can be identified as main problems. This is evident in the results of the semantic analysis of user comments but can also be derived from a more detailed analysis of the sessions documented in the diary study, which is not presented here. Problems starting the charging process are mostly due to authentication issues. In this regard, both the statistical data analysis and the charging experiment show that RFID authentication seems to work more robustly than remote authentication via native smartphone apps or web apps. Moreover, as initial results from the diary study show, incorrect POI data and the inconsistent use of QR codes are further causes of starting problems.

5.2 Limitations & further research

The methods and results of this study have several limitations. Firstly, the statistical data analysis to date provides only preliminary evidence. Relevant factors for evaluating charging processes, such as the amount of energy charged or the average charging power, have not yet been considered. Also, faulty authentication processes without an assigned CDR can only be evaluated inadequately so far, since the error messages necessary for a better analysis have so far been provided in connection with CDRs only. These steps as well as further analyses, e.g., on correlations, are to be carried out in the further course of the project. Secondly, not all conducted studies can be considered representative. Except for the semantic analysis of user comments, all methods were limited to a reduced number of EMSPs and CPOs or charging stations. However, especially the diary study proved to be a valuable tool for identifying weaknesses in the ecosystem. In this respect, it is advisable for the industry to jointly pursue this cross-stakeholder approach further.

Acknowledgments

The authors would like to extend their gratitude to all members of the “Wirkkette Laden” consortium (ABB, BMW, CharIN, Digital Charging Solutions, EcoG, EWE Go, Ionity, P3, Porsche Engineering, Stromnetz Hamburg) for the fruitful collaboration, and the German Federal Ministry for Digital and Transport (BMDV) for funding this project.

References

[1] F. Röckle, R. Litauer, *Current and potential future EV driver charging needs*, 2018

[2] F. Röckle, *Begleitforschung USP-BW*, not published, 2021

[3] *Focus*, https://www.focus.de/auto/elektroauto/tesla-mit-hoechster-zufriedenheit-nur-jeder-zweite-elektroauto-fahrer-wuerde-seinen-wagen-weiterempfehlen_id_13399926.html, accessed on 2021-12-20

[4] S. Eisenkrämer, *Feldversuch: Hohe Ladeleistung viele Fehler*, ATZelektronik Ausgabe 11/2021, Wiesbaden, Springer Fachmedien Wiesbaden GmbH, 2021

[5] T. Franke, J.F. Krems, *Understanding charging behaviour of electric vehicle users*, Transportation Research Part F: Traffic Psychology and Behaviour, 21 (2013), 75-89

[6] N. Daina et al., *Electric vehicle charging choices: Modelling and implications for smart charging services.* Transportation Research Part C: Emerging Technologies, 81 (2017), 36-56

[7] J. Dixon, K. Bell, *Electric vehicles: battery capacity, charger power, access to charging and the impacts on distribution networks*. ETransportation, 4 (2020), 100059

[8] A. Alibasic, T. Popovic, *Applying natural language processing to analyze customer satisfaction*, 25th International Conference on Information Technology (IT), 2021, 1-4, doi: 10.1109/IT51528.2021.9390111

[9] T. Pfeifer, R. Schmitt (Eds.), *Masing Handbuch Qualitätsmanagement*, ISBN 3446462309, München, Carl Hanser Verlag GmbH Co KG, 2021

[10] *Kairntech*, https://kairntech.com, accessed on 2022-05-10

[11] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," arXiv preprint arXiv:1810.04805, 2018

[12] J. Greitemann, *Methodik für die systematische Identifikation von Produktionstechnologien*, Dissertation, TU München, 2016

[13] *Expert.AI*, https://community.expert.ai/articles-showcase-56/precision-and-recall-f-score-and-accuracy-measuring-nlp-performance-191, accessed on 2022-05-10

[14] I. Kokkinos, *Pushing the boundaries of boundary detection using deep learning*, https://arxiv.org/pdf/1511.07386

[15] D.M.W. Powers, *Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation*, Preprint, https://arxiv.org/abs/2010.16061

[16] G. Hripcsak, A.S. Rothschild, *Agreement, the F-Measure, and Reliability in Information Retrieval,* Journal of the American Medical Informatics Association, Volume 12, Issue 3, May 2005, 296–298, https://doi.org/10.1197/jamia.M1733

[17] C. Gibbons et al., *Supervised Machine Learning Algorithms Can Classify Open-Text Feedback of Doctor Performance With Human-Level Accuracy*, In: Journal of Medical Internet Research Vol 19 (3), 2017

[18] *BDEW*, https://bdew-codes.de/Content/Files/Emobility/171129_Richtlinien_E-Mobility-Codenummern_final.pdf, accessed on 2022-05-02

Author

Felix Roeckle holds a B.Eng. in Business Engineering from the Cooperative State University Baden-Wuerttemberg as well as an M.Sc. in "Innovation Management & Entrepreneurship" and "Business Administration" from TU Berlin and the University of Twente respectively. After two years of work experience in the industry, he joined Fraunhofer IAO as research fellow in July 2016. His research is focused on the efficient design of future mobility systems.