Big Data Usage Can Be a Solution for User Behavior Evaluation: An Automotive Industry Example

Citation for the original published paper (version of record):

Orlovska, J., Wickman, C., Söderberg, R. (2018). Big Data Usage Can Be a Solution for User Behavior Evaluation: An Automotive Industry Example. Procedia CIRP, 72(1): 117-122.
http://dx.doi.org/10.1016/j.procir.2018.03.102

(article starts on next page)

Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 72 (2018) 117–122

www.elsevier.com/locate/procedia

51st CIRP Conference on Manufacturing Systems

Big Data Usage Can Be a Solution for User Behavior Evaluation: An Automotive Industry Example.

Julia Orlovska [a,*], Casper Wickman [a,b], Rikard Söderberg [a]

[a] Chalmers University of Technology, Department of Industrial and Materials Science, 412 96, Göteborg, Sweden
[b] Volvo Car Corporation, 91300 User Experience & Craftsmanship Centre, Torslanda PV3A, SE-405 31 Göteborg, Sweden

* Corresponding author. Tel.: +460731428333, E-mail address: orlovska@chalmers.se

Abstract

The current level of User Interface complexity leaves almost no space for subjective assessment of user interaction design. Successful Human-Machine Interface (HMI) assessment can be conveyed through big data analysis of user behavior in the real environment. A series of interviews with UI/UX engineers from the leading Swedish automotive manufacturer revealed the limitations of existing methods and deficiency of objective information sources. Design of the case regarding real user data analysis that is presented can bring a better understanding of different users' behavior patterns, which can lead to the improvement of future HMI systems. The data-driven approach can establish a foundation of robust methodologies regarding the objective evaluation of HMI design.

Peer-review under responsibility of the scientific committee of the 51st CIRP Conference on Manufacturing Systems.

Keywords: big data; real user data; data analysis; usability; HMI; automotive UI; user interaction; user behavior

1. Introduction

The current state of automotive interfaces can be described as a system of systems with a high level of complexity. It is hard to predict and evaluate the constantly changing, randomly induced scenarios of human behavior in the real environment. The major obstacle is an inability to predict individual reactions in different situations, since user reactions to a particular scenario depend on a variety of parameters such as the speed of the vehicle, road conditions, driving mode, and multiple functional tasks that needed to be executed simultaneously.

To be able to evaluate user behavior and system performance to recognize and predict potential usability issues of specific functions usability evaluation methods (UEM) are primarily used today. However, UEM methods for HMI assessment of launched products is mainly based on inquiry methods for usability or user experience evaluation, such as interviews, surveys or questionnaires [1]. These methods presume subjective ratings, with usability ranking scales, and focus on gathering subjective impressions of using the HMI, rather than focusing on specific tasks or measuring performance. The questions that usability experts are able to ask during these inquiries are very general because they refer to a variety of users at the same time, without knowing if a particular user encounters the discussed usability problem or not. To receive detailed information from users, usability experts need to set up a range of additional follow-up questions. However, this task is arduous due to the complexity of automotive HMI. Usability engineers have to evaluate hundreds of HMI functions for a single vehicle. Moreover, questionnaires covering many details often require additional time from respondents and consequently provide weak results. Excessive questionnaire design can affect the consistency of the responses and often suffers from lack of discrimination regarding commonly used

2212-8271 © 2018 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of the scientific committee of the 51st CIRP Conference on Manufacturing Systems.
10.1016/j.procir.2018.03.102

scales (e.g., semantic-differential) [2]. Another important issue is the decrement of human ability to recall the required information over time [3]. As a result, the average user can only report a few issues that are important to him/her. Often drivers have difficulties in estimation of the time they spent using a particular application or the timeframe the system needed to respond to the command. Moreover, users tend to estimate their own actions and choices inappropriately. This can be explained by the fact that a driver is often not aware of the most effective patterns regarding the use of the interface. Usually the driver is just experimenting with the HMI during the learning phase and develops his/her individual usage scenarios. Human attention has a sporadic nature and is easily distracted. It is difficult to recall all parameters that were involved in the usability evaluation process and detect their importance for the correct functional performance.

One way to improve the evaluation of user behavior is to consider available data during the evaluation process [4]. Automotive software systems nowadays provide an enormous stream of data linking people to their activities, vehicle's location, sequential tasks they perform and the driving conditions they meet. Big data analysis of these signals is able to comprehend these data flows and use these as a powerful tool for HMI evaluation.

In this paper we present the hypothesis that big data analysis of real user behavior can serve as a solution for enhancing the quality of HMI evaluation in the automotive industry.

Notably, the use of big data in this case is not aiming to replace subjective evaluation, but rather to enhance it. Measurement of user performance data alone cannot help us to answer questions such as “Why did different usability issues occur?” or “Why do users prefer certain operational patterns over another?” There is still plenty of space for subjective evaluation. However, big data analysis of user performance in the real environment in combination with system performance evaluation could help engineers to cluster users who encountered similar usability issues under equivalent usage conditions. The evaluation of users’ behavioral patterns can eventually help to understand the root cause of the encountered problem or establish a stronger basis for a more precisely subjective evaluation than ever before.

This approach can dynamically utilize implicit knowledge available on the data level, as well as generate new knowledge by detecting and including new signals that are beneficial for the understanding of user behavior. The main advantage to involving data in subjective evaluation is that the system does not tend to “forget” or “generalize” any issues. The gathered information is based on objective data and could easily be translated into a specific number, percentage or other objective value.

It is important to mention that observation of user behavior through performance data can detect the tasks or the parts of the interface that are problematic for the driver, but they cannot identify the root cause of the problem [5]. Therefore, although performance data can tell us about the occurrence of an event on the interaction level very effectively, it is not easy to answer the questions that refer to the reason for the occurrence of the event. As the result, we leave those types of questions for subjective evaluation at a later stage.

Despite the potential advantages, the domain of using big data for user-oriented studies in the automotive industry is poorly explored. This paper reveals main challenges that usability engineers describe today and discusses ways for improvement.

The performed study includes interviews with the UI/UX engineers from a leading Swedish automotive OEM (Original Equipment Manufacturer). Through the analysis of these interviews, we obtained a list of questions/inquiries for Driver Support system evaluation. The Driver Support system is a semiautonomous system that can control the vehicle by steering, braking and accelerating through various traffic conditions.

Furthermore, we investigated the possibility to use big data analysis for Driver Support system assessment by identifying the feasibility of required data retrieval. The obtained qualitative results suggest further research is warranted, and a framework is proposed to use real user data analysis for Driver Support system evaluation. Data mining and acquisition of test vehicles in this study was supported by the WICE system (flexible automotive platform providing access to measurement data from test vehicles) and provided by the OEM. The total number of measured vehicles was 97.

This paper structured as follows: Section 2 presents the background of relevant knowledge; Section 3 introduces the qualitative methodology used in this work and presents the case study design; Section 4 discusses early findings; Section 5 discusses advantages and the main challenges of this approach; Section 6 presents conclusions and recommendations for further research.

2. Background

Big data is a relatively new term that indicates an amount of data that is difficult to store, process and analyze using traditional database technologies [6]. The definition of big data can vary from a large volume of data for scientific visualization [7] to a large volume of data that is beyond the technical capability to store, manage and process efficiently

[8]. The most traditional way to describe big data is using the "three V's" - big data characteristics: Volume, Variety, and Velocity [9]. However, some researchers argue that Value, the fourth "V" characteristic, is the most important dimension of big data [10, 11]. Value extracting is the main purpose of big data processing. It highlights the importance of big data as a source of knowledge and refers to the process of discovering hidden values from large datasets [12].

The definition provided by Gantz and Reinsel [10] reflects the view on big data adopted in this study: "Big data describes a new generation of technologies and architectures, designed to extract value from very large volumes of a wide variety of data, by enabling high-velocity capture, discovery, and/or analysis." Big data can include transactional data, warehoused data, metadata, and other data that could be captured from information on the product consumption or utilization (e.g., GPS data, media or infotainment data).

2.1. Technical characteristics of the data mining system

The case study design described in this paper includes the help of a WICE system, which allows the field evaluation of the vehicle fleet involved in the study. The WICE system is a telematics platform providing data access from the test vehicles. It consists of two major parts: (i) the in-vehicle telematics data measurement system (ii) the back-end server infrastructure, together with a web-based front-end user interface that includes data storage units and database. Overall, the system provides metrology services including various signal types for collection and measurement. The WICE system is able to manage information from the vehicle fleet by keeping track of map-based positioning, mileage, and uptime or diagnostic codes. A more detailed description of the WICE system can be found in Johanson [13].

2.2. Functional description of Driver Support system

The DS (Driver Support) system is a semi-autonomous system that can steer, brake or accelerate the vehicle through various traffic conditions, as well as provide help in the control of the vehicle on the road. In the case study design, we evaluate two functions of the DS system: Adaptive Cruise Control and Pilot Assist.

2.2.1. Adaptive Cruise Control

The ACC (Adaptive Cruise Control) function helps the driver to maintain the vehicle's speed with a preselected time interval to the vehicle positioned in front. ACC is usually enabled during long journeys with steady traffic conditions. It is achieved using vehicle cameras and radar and automatically adjusts the vehicle's speed with regard to other objects moving in front.

2.2.2. Pilot Assist

The PA (Pilot Assist) function helps the driver to keep the car in the road lane and maintains the interval to the preceding vehicle, considering the speed control. This is achieved using cameras and the vehicle's radar unit. The driver can take back control of the system at any time since it is semi-autonomous.

3. Methodology

The hypothesis presented was based on the assumption that if big data could be used successfully at different stages of product development and production in the automotive industry, it would be possible to acquire and extract user-oriented data and utilize this in user behavior evaluation. An additional task was to understand how big data can be used for user behavior evaluation and what benefits the utilization of this data can bring.

For this purpose, a two-stage study was performed. During Stage 1 unstructured face-to-face interviews were conducted with nine UI/UX engineers involved in HMI evaluation processes. The primary goal of these interviews was to identify questions regarding user behavior evaluation that could be answered with the help of big data analysis and to detect measurable areas regarding the usage of DS system. As a result, we obtained a list of 18 evaluation questions/inquiries to the system that are difficult to grasp by existing evaluation methods. For example, "How many users deactivate the function after they have tried it once?" or "How much time does it take the user to learn a function?"

Throughout Stage 2 several workshops with UI/UX engineers, data engineers, and data analysts were organized. The major goal of organized discussions was to identify possible types of data needed to answer the 18 evaluation questions/inquires that were identified in the previous step. We verified the possibility for required data retrieval (from the side of the data engineers), and defined and described different usage scenarios and different metrics measuring scenarios. We also defined the user characteristics essential for the case study design (drivers should have previous experience regarding the DS system), and the measuring parameters, i.e., the number of trials, the time frame for measuring, etc. that needed to be set.

The case study for data-based subjective evaluation of user behavior for the DS system was subsequently designed.

4. Results

We propose an approach to data-based evaluation that integrates the real user data analysis into the overall assessment of HMI (see Fig.1)

The study procedure planned was as follows: (1) to perform the objective evaluation, the evaluation questions/inquiries are set; (2) the required data is identified and the data set defined; (3) measuring parameters are defined: i.e., the number of trials, the time frame for measuring, specific user parameters, if any; (4) data engineers collect the required data according to measuring parameters; (5) collected data is analyzed to answer the evaluation questions; (6) an evaluation is performed to determine if any hidden knowledge can be extracted (or confirm the known hypotheses); (7) if a usability issue was detected, we cluster users regarding their behavior and measure the magnitude of the clustered groups to understand if the detected issue is essential for a particular evaluation question. Otherwise, we redesign the evaluation questions/inquiries by including new measuring parameters and repeat steps (2-7); (8) we include and consider as many related data signals as we identify to learn more about the detected issues until the level of uncertainty is reached; (9) we conduct a subjective assessment, which can be based on methods of customer polling or data-based expert evaluation.

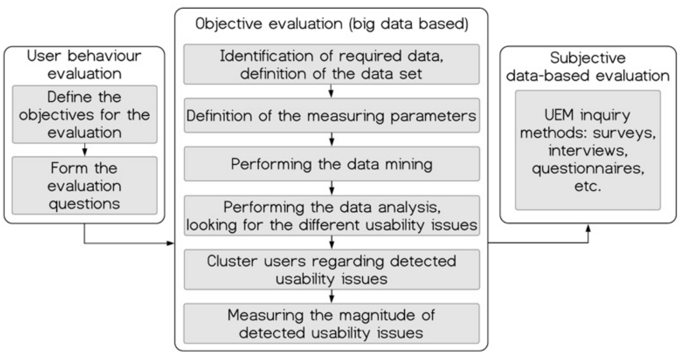

Fig.1. Case study design block scheme

In our case, subjective evaluation presumes data-based expert evaluation or employing inquiry methods such as surveys, questionnaires or interviews to gather supplementary data after the testing is completed. One example of a question for subjective evaluation is, "Why do users use ACC more than the PA function?" However, the questions that users will be asked in this case will be based on the data analytics and user clustering that was performed at the previous stage, resulting in questions that will be more precise and specific. Big data analysis brings us the possibility to ask the particular users who have been encountering an exact usability issue (one that we consider essential) and will enable us to receive a very specific answer that can extend our understanding of the previously unknown causes of the problem.

Additionally, to be able to trust data insights, we need to exclude the "guessing factor." For this purpose, we include and consider only those data signals around detected issues that allow us to make conclusions without any speculations around the data.

4.1. Prototyping Results / Proof of Concept

The results of the collected data analysis can be used for different purposes, such as predicting the future behavior of the user, current market needs, improvement of the future generation of interfaces or lookup of any deviations from the regular behavior pattern, as well as identifying users for further investigation.

In the first phase we explored user behavior to see if the functions performed as intended. Additionally, we wanted to understand the current needs of customers and verify if the requirements set is actually meeting user needs.

The preliminary results showed that ACC is used more often compared to PA. From 100% of all activations of the DS system, 94.72% reported the use of ACC and only 5.28% reported the use of PA.

Since the main purpose of PA is to provide a lane keeping function for the ACC functionality, the majority of drivers surprisingly avoided the option to delegate steering control to the system.

Consequently, two assumptions need to be investigated further: (i) users do not trust the system for unknown reasons (e.g., safety, novelty, fear, etc.); (ii) PA function does not provide sufficient quality (e.g., it is unavailable when needed, or deactivates automatically). The next logical step is to investigate the system performance to exclude the possibility of poor quality delivered by the PA function.

It is important to mention that the DS system is not suitable for all driving situations. The system performance depends on external factors, such as weather conditions, road type, driving style, traffic conditions, etc. In fact, we assume that during the summer period, when the weather conditions are sufficiently better, usage of the DS system is higher compared to the rest of the time. The data partially confirms this assumption (see Fig. 2). For the ACC function, usage increased from April 2017, and almost doubled during the summer period, starting to slowly decrease after August 2017. Regarding the PA, we cannot make any assumptions yet, since the measuring time for PA started half a year later in comparison to ACC.

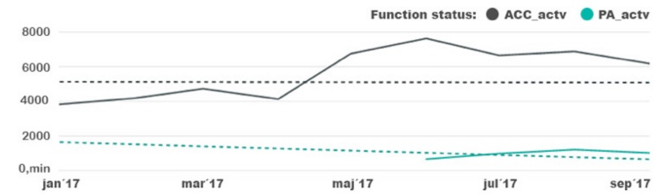

Fig.2. Change of functions status during the year

In this study we also wanted to clarify if the usage of the functions was dependent on the time of day. For this reason we measured the use of the functions at different hours of the day (Fig. 3). The data gathered shows that the majority of drivers use ACC and PA independently of the time of day. Therefore, we need to be certain that the quality of sensors for road condition detection must provide the same quality during night hours as well.

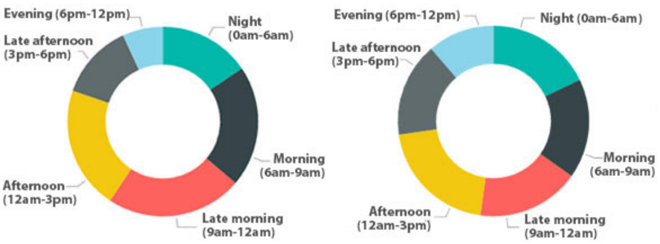

Fig.3. Usage of ACC (left) and PA (right) during the hours of the day

Combining different data parameters, such as activation duration time, distance of activation, vehicle speed, GPS data, road condition sensors data, etc., we will have a better understanding of how users use the DS system and what usability issues they encounter.

5. Discussion

The prototyping results showed that with the help of big data users are able to be divided according to the chosen parameter, such as usage of the function during day/night hours or summer/winter seasons, speed range and other factors. It helps in classifying drivers into different user groups for further investigation. The examples shown are simple, but the combination of different parameters together can provide more advanced results with regard to delineation of the different user groups. Subjective evaluation based on particular knowledge of a specific user group brings more accurate results in discovering the causes of the identified usability issue.

User behavior analysis in the real environment, with the help of big data analysis, is becoming more agile and a rival for automotive engineering teams. For example, analyzing users' behavior for semi- or fully automated functions that are embodied in the vehicles (PA, ACC, automated parking, etc.) we can discover the magnitude of appreciation and acceptance of the newly introduced features. Knowledge regarding the level of trust to these functions corresponds to whether the user needs more time to adapt to and trust this new technology functionality. This can significantly improve strategic planning concerning the assessment of technology readiness level and can contribute to the managerial implications regarding the adaptation of newly introduced functionality.

Besides the higher goals of strategic planning, big data analysis can be helpful for HMI evaluation:

- **It helps to deal with user interface complexity.** The more complex interfaces we have, the more difficult it becomes to grasp and evaluate them. As a result, we usually limit our goals to the investigation of specific parts of the interface, or even specific functions. In turn, this gives us fragmentary information that is often not useful for detection and estimation of usability issues. In this situation, using big data for HMI evaluation is more effective.
- **It allows understanding of user behavior in combination with system performance**. The common practices to evaluate user and system performance separately are not able to show the dependencies of how user behavior changes with alteration of system performance. Users themselves can hardly estimate what will happen if the system performed better/worse, faster/slower than it used to. The comparative data analysis of user and system performance can support these kinds of questions.
- **It enables more accurate feedback regarding user performance.** Depending on the real environment conditions, the user behaves differently in comparison to the quiet and focused environment of the lab where the tasks and their sequences are predefined. The level of concentration in performing lab-tasks is usually higher compared to the real environment and, therefore, may not correspond to the actual behavior.
- **It can capture geographic diversity of users.** This fact is very beneficial for a globally extended market. The possibility to collect data without being dependent on a physical location of the respondent provides evaluators an opportunity to evaluate and compare different markets.

- **It allows testing of an unlimited user range.** The number of participants does not significantly affect the time and the cost of evaluation. The more users we evaluate the cheaper it becomes per user testing.

However we also identified the following limitations regarding data identification and data processing:

- Over the years, OEMs have generated enormous volumes of data, yet there is a lack of structure and consistency. Software platforms are often not designed for the collection of all of the data required, or there may be an inability to process information from the sensors or measuring units currently in use.
- In this study user identification is important for cases where the dynamics of usage process need to be monitored. However, it is not an easy task to understand if possible intervention in the monitoring process by a person other than the car owner took place. We decided to identify the user through the sum of first level parameters for user identification: driver profile, the key used, seat memory activation, manual seat positioning, and driving route. If the measurement of those parameters shows that the user coincides with the user profile to a level of 75% or more, we identify the user as the same according to the user profile. Otherwise, we do not include the data in the data analysis, due to the high probability of user authentication error.
- Due to the nature of this study we were not able to measure user emotions or perceptions. However, gathering the subjective impressions of users and their preferences or opinions regarding various aspects of the HMI is particularly important for developing of the next generation of such systems. To obtain a holistic understanding of user perceptions, these should be assessed by traditional inquiry methods, such as surveys, questionnaires or interviews.
- There are stringent demands concerning the regulations of privacy and security issues in the information space where it is possible to identify individuals, their behavioral preferences, track location, and activities. To overcome these issues, new methods and techniques, such as information encryption, must be used to ensure that user confidentiality is not violated.

6. Conclusion

Big data and technological advances have an undeniably great value for the future of understanding user behavior. However, big data is not knowledge. We need to develop methods for classification and sophisticated extraction of the relevant information for successful product design. This cannot be done through one discipline alone. Several disciplines must be involved, where the applied approach has to be combined with the theoretical.

We presented the design of the method for data-based user behavior evaluation, validated by industry professionals. We also presented first results that could shape future studies regarding the evaluation of user behaviors. We have shown that a big data approach can increase the detection of usability issues, allowing their magnitude to be measured by clustering users with similar behavior.

Acknowledgements

We want to thank our industrial, automotive partner for providing us with the necessary data for this case study and giving us full support and constructive feedback on our investigation. That support is gratefully acknowledged.

The authors also wish to acknowledge the contribution of Sigma Technology AB for significant help in data identification, data analysis and the visualization of results.

References

[1] Ivory, M.Y. and Hearst, M.A., 2001. The state of the art in automating usability evaluation of user interfaces. ACM Computing Surveys (CSUR), 33(4), pp.470-516.

[2] Magidson, J., Thomas, D., & Vermunt, J. K. (2009). A new model for the fusion of maxdiff scaling and ratings data. In 2009 Sawtooth Software Conference Proceedings, pp. 83- 103. 2009.

[3] Schaeffer, J., & Carlsson, A. L. (2014). The method of photo-elicitation from a phenomenological perspective. In DS 77: Proceedings of the DESIGN 2014 13th International Design Conference.

[4] McAfee, A., Brynjolfsson, E., Davenport, T.H., Patil, D.J. and Barton, D., 2012. Big data: the management revolution. Harvard business review, 90(10), pp.60-68.

[5] Tullis, T. & Albert, B., 2013. Performance Metrics. Measuring the User Experience, pp.63–97.

[6] Hashem, I.A.T., Yaqoob, I., Anuar, N.B., Mokhtar, S., Gani, A. and Khan, S.U., 2015. The rise of "big data" on cloud computing: Review and open research issues. Information Systems, 47, pp.98-115.

[7] Cox, M. and Ellsworth, D., 1997, August. Managing big data for scientific visualization. In ACM Siggraph (Vol. 97, pp. 21-38).

[8] Brown, B., Bughin, J., Byers, A.H., Chui, M., Dobbs, R., Manyika, J. and Roxburgh, C., 2011. Big data: the next frontier for innovation, competition, and productivity. McKinsey Global Institute.

[9] Berman, J.J., 2013. Principles of big data: preparing, sharing, and analyzing complex information. Newnes.

[10] Gantz, J. and Reinsel, D., 2011. Extracting value from chaos. IDC iview, 1142(2011), pp.1-12.

[11] Chen, M., Mao, S. and Liu, Y., 2014. Big data: A survey. Mobile Networks and Applications, 19(2), pp.171-209.

[12] Alpaydin, E., 2010. Introduction to Machine Learning, 2nd edn. Adaptive Computation and Machine Learning.

[13] Johanson M., WICE: Automotive telematics, fleet management, rapid prototyping and software download for test vehicles. [online] Alkit Communications AB. Available at: http://www.alkit.se/doc/WICE-whitepaper.pdf [05 October 2017].