

USABILITY OF DIGITIZED CITIZENS' SERVICES – A HEURISTIC EVALUATION BASED ON EXPERIENCES WITH USABILITY LABS WITHIN THE IMPLEMENTATION OF THE GERMAN ONLINE ACCESS ACT

Laura Stasch and Anna Steidle¹

DOI: 10.24989/ocg.v338.8

Abstract

Germany's public administrations must go digital by law till 2022. The German Online Access Act (our translation for "Onlinezugangsgesetz", OZG) forces most services offered by public administration on federal, federal state and municipal level to become digitized. For most of these services still being paper-based and Germany not being one of the leaders in e- government according to many sources, the question of user acceptance arises. For answering the question whether the approach used in the digitalization labs leads to the development of digital public services that are accepted by future users, we conducted a heuristic evaluation of a prototype that was developed within the implementation of the OZG. The paper describes the setting, the test undertaken and the outcome and concludes with an estimate, whether the huge paradigm change towards the development of digital public services that are accepted by future users will be successful or not.

Keywords: *Onlinezugangsgesetz (OZG), digitalization labs, usability evaluation*

1. Introduction

E-Government is not state of the art in Germany. According to the Digital Economy and Society Index 25 of 27 European countries have advanced faster than Germany. [1] Currently, most public services in Germany are offered in a paper-based form only. Hence, the number of public services still to be digitalized is enormous and presents a major challenge for the public sector in the near future. With the "Onlinezugangsgesetz" (OZG) the national government has set the goal that Germany's public administration has to offer public services covered by the OZG digitally by 2022. [3]

For the implementation of the OZG the public services covered by the OZG have to be identified first, then prioritized and clustered. The basis therefore is the so-called OZG implementation catalogue. In summary, the OZG implementation catalog identifies about 575 public services that are covered by the OZG. These public services covered by the OZG are clustered into 35 life situations and 17 business situations. Each life and business situation is assigned to one of 14 topics. [3]

¹ Hochschule für öffentliche Verwaltung und Finanzen, Ludwigsburg, eMail: {first name}.{name}@hs-ludwigsburg.de

A federal department and one or more federal states are jointly in charge for each topic. For the processing of each topic they develop an individual plan for how to digitalize each public service that is assigned to the topic. Public services that have high priority according to the OZG implementation catalogue get digitalized in digitalization labs. [4] In order to ensure the best possible transformation of the public services the team belonging to the digitalization lab consists of interdisciplinary team members such as professional experts, design experts and future users. [6] In the digitalization lab the interdisciplinary team develops a prototype of each digital public service using user-centered methods such as design thinking. [5] After the digitalization of the digital public service the prototype is handed over to the IT service provider of each federal state who then can implement the digital public service in their systems using the prototype. [4] Note that a single digital public service has up to 16 different implementations for the 16 federal states respectively.

From a quantitative perspective, the goal set within the implementation of the OZG is to digitalize the 575 public services covered by the OZG until 2022. [3] From a qualitative perspective the goal is to develop digital public services that are accepted and used by future users. This qualitative goal implies a paradigm change in the development of public services from a restriction-centered to a user-centered perspective. Hence, in the design process of digital public services, user's expectations should be in the focus of attention, while the previously more important restrictions are treated as framework conditions. [4] In this context, the question is if the approach used in the digitalization labs leads to the development of digital public services that are accepted by the future user.

Generally, a high degree of usability is a key factor for the acceptance of digital public services. [10] Due the different usability standards introduced by different experts and due to the different perspectives of different stakeholders the definitions of usability vary considerably. [7] We refer to the definition of usability by Jakob Nielsen who defines usability as "quality attribute that assesses how easy user interfaces are to use." According to Nielsen usability consists of five different quality components: learnability, efficiency, memorability, errors and satisfaction. Learnability considers "how easy it is for users to accomplish basic tasks the first time they encounter the design". Efficiency considers how quickly the users can perform tasks once they have got to know the design. Memorability regards how readily the users can restore their skill after they "return to the design after a period of not using it". Errors considers how many errors the users make, how severe these errors are and how readily the users can recover from these errors. Satisfaction regards how pleasant the use of the design is. [10]

In order to answer the question if the approach used in the digitalization labs leads to the development of digital public services that are accepted by future users, the lead author conducted a case study in her bachelor thesis. [12] In the case study, we conducted a usability evaluation of a prototype that was developed in a digitalization lab called "special use of public space".

2. Method

The prototype regarded within the usability evaluation shows how an application for the special use of public streets, such as e.g. setting up a container on a public street, could look like. According to the OZG implementation catalogue the digitalization lab "special use of public spaces" belongs to the topic "management and company development". The federal state of The Free and Hanseatic City of Hamburg and the German Federal Ministry of the Interior are responsible for the processing of this topic. All rights to the prototype are held by the Federal Ministry of the Interior. [5] For selecting a suitable usability evaluation method we firstly set up selection criteria. Since the evaluation is conducted by a single person with limited resources, the method should be feasible for one person

and be time and cost efficient. Nevertheless, the method should show the prototype's potential for optimization. Furthermore, when selecting the method the prototype's characteristic, that the way the evaluator clicks through the prototype is predetermined, has to be taken into account. Secondly, it must be differentiated between the usability evaluation methods in user-based and expert-based usability evaluation methods. After opposing each area's characteristics, advantages and disadvantages to the selecting criteria, the team decided to select an expert-based usability evaluation method. [2]

Considering the selection criteria, the team thirdly selected the heuristic evaluation as expert-based usability evaluation method. A heuristic evaluation is a usability evaluation method that is used to identify a user interface design's usability problems. Therefore, the user interface design is inspected if it complies with acknowledged usability principles, the so called "heuristics".

The outcome of a heuristic evaluation is a list of usability problems caused by the user interface design's violation of an acknowledged usability principle. [9]

According to Nielsen a single evaluator is able to find only 35 % of a user interface design's usability problems. Due to that findings Nielsen recommends to conduct a heuristic evaluation by a small group of 3 to 5 evaluators. [9] With respect to the limited resources the team had the evaluation conducted by a single evaluator. Therefore, the evaluator firstly familiarizes him- or herself with the prototype by clicking through the prototype several times. Secondly the evaluator conducted the evaluation in a two-hour session by inspecting whether any element of the prototype violates one of the acknowledged principles.

The acknowledged principles that were used for conducting the heuristic evaluation are seven out of ten of Nielsen's usability heuristics, these are #1: Visibility of system status, #2: match between system and the real world, #3: User control and freedom, #4: Consistency and standards, #5: Error prevention, #6: Recognition rather than recall and #8: Aesthetic and minimalist design. Moreover, the team included everything that doesn't violate one of the heuristics but will irritate the user while using the prototype under the heading "other usability problems". [8] Due to the predetermined way the evaluator clicks through the prototype the three heuristics #7: Flexibility and efficiency of use, #9: Help users recognize, diagnose, and recover from errors and #10: Help and documentation can't be considered within this evaluation. But since Nielsen's usability heuristics are rather a thumb rule than stationary usability guidelines [8], this doesn't detract the evaluations quality.

In addition, to the number and nature of usability problems, the team also assessed their severity. In order to find out to which extent each usability problem reduces the prototype's usability, they categorized the identified usability problems according to the estimated extent in which the usability problems will disturb the user while using the prototype and how easily the user can overcome the disturbance. They categorized the usability problems in three categories: severe, moderate and slight usability problems.

Severe usability problems disturb the user in a severe way. Due to the severe disturbance, users have to invest a high cognitive effort to overcome these usability problems. This is why overcoming severe usability problems is extremely difficult for the user. These severe usability problems disturb the user in such a severe way that it will prospectively lead to the demolition of the filing of application.

Moderate usability problems rather irritate than disturb the user. The cognitive effort the user has to practice for overcoming the moderate usability problem is lower than the cognitive effort that the user

has to practice for overcoming the severe usability problem. Accordingly, moderate usability problems are easier to overcome than severe usability problems and prospectively will not lead to the termination of the filing of an application.

Slight usability problems violate one of the inspected heuristics. They are usability problems by definition so they reduce the prototype's usability. Despite the reduction of the prototype's usability these problems do not disturb the user while using the prototype. Since slight usability problems do not disturb the user while using the prototype, the user can overcome slight usability problems without applying much cognitive effort for overcoming. Hence, slight usability problems are merely cosmetic and prospectively will not lead to the termination of the filing of application.

Depending on this categorization the team deflects the prototype's degree of usability. In this context, the team defines that the less severe usability problems are found the higher is the degree of usability.

3. Results

Altogether 28 usability problems were discovered.

Regarded Heuristic	Number of violations	Number of severe usability problems	Number of moderate usability problems	Number of slight usability problems
#1: Visibility of system status	0			
#2: match between system and the real world	6			6
#3: User control and freedom	1		1	
#4: Consistency and standards	3	1		2
#5: Error prevention	3	1		2
#6: Recognition rather than recall	2		1	1
#8: Aesthetic and minimalist design	9			9
other usability problems	4		1	3

Table 1: Results of the conducted evaluation

With nine violations the eight heuristic "Aesthetic and minimalist design" was most frequently violated. This means, that the prototype often shows information which is rarely needed or irrelevant. This rarely needed or irrelevant information competes with the relevant information and reduces the relative visibility of the relevant information.

The second heuristic "match between system and the real world" was violated six times. Three of these six violations mean that the prototype uses terms that are unfamiliar to the user, so the prototype does not speak the user's language. The other three of these six violations mean that the information appear in an order that is unnatural or illogical for the user.

The fourth heuristic "Consistency and standards" and the fifth heuristic "Error prevention" were violated three times. The three violations of the fourth heuristic mean that the user must think at three points of the prototype if different actions or words mean the same. The three violations of the fifth heuristic mean that at three point of the prototype there are error-prone conditions that are not eliminated or presented to users before they pledge themselves.

The sixth heuristic “Recognition rather than recall” was violated two times. This means that actions and objects are not made visible so the users have to remember information between different parts of the Prototype.

The third heuristic “User control and freedom” was violated one time. This means that if the user used the prototype in an unwanted way he would have to run through an extended dialogue to correct the unwanted use.

The first heuristic “visibility of system status” was not violated at all. This means that the prototype always informs the user about what is happening.

Under the heading “other usability problems” we assigned four issues that will irritate the user while using the prototype. These were issues such as information overloads or showing conflicting information.

A detailed explanation of all usability problems can be found in the bachelor thesis. [12] As result of the severity analyses, the team found out that two of the 28 identified usability problems are severe usability problems, three of 28 usability problems are moderate usability problems and 23 identified usability problems are minimal usability problems. Although these usability problems reduce the prototypes usability, they do not disturb the user while using the prototype so the user also does not even have to overcome these usability problems. These 23 usability problems were merely cosmetic. According to this evaluation the team found that despite existing usability problems the prototype has still a high degree of usability.

4. Discussion

The findings of our usability evaluation show that the evaluated prototype has a high degree of usability and therefore can be judged as a useful tool to facilitate developing digital public services of a high quality (i.e. high user acceptance).

However, there is a number of limitations regarding the evaluation process. First, the above described heuristic evaluation was conducted by a single person. According to Nielsen a single evaluator is able to find on average 35 % of a user interface design’s usability problems. Due to this fact Nielsen recommends conducting the heuristic evaluation by a small group of 3 to 5 evaluators. [9] Hence, a larger group of evaluators might have detected more usability problems or judged their severity differently. If for instance further evaluators detected more usability problems and judge their severity rather as severe the prototype’s overall usability could be lower than according to our conducted evaluation. If further evaluators detected more usability problems and judged their severity rather as slight the prototype’s overall degree of usability could still be high. Second, the chosen method was an expert-based method. Hence, the evaluator was a usability expert and not a “real” user. This raises the question to which extent the usability problems identified by the evaluator correspond to usability problems an user would experience. [11]

Third, the three heuristics “#7: Flexibility and efficiency of use”, “#9: Help users recognize, diagnose, and recover from errors” and “#10: Help and documentation” by Nielsen weren’t considered within the conducted evaluation. Included these heuristics into the analyses might have led to different results.

Finally, it has to be taken into account that the degree of usability is only one of numerous key factors for the acceptance of digital public services. Other key factors for the acceptance of digital public services such as the user's trust in information security should also be considered.

The following implications for future research and evaluation may help to overcome the limitations listed above. The prototype should be evaluated by 2 - 4 additional usability experts. Furthermore, we recommend conducting another heuristic evaluation using the three not considered heuristics and conducting a user-based evaluation as soon as possible. Since the degree of usability is not the only key factor for the acceptance of digital public services, we also recommend regarding other key factors for the acceptance of digital public services.

In consideration of these limitations this heuristic evaluation is a first but very important step in the inspection of the paradigm change in the development process of digital public services.

Regarding only the degree of the prototype's usability, we still conclude that due to the prototypes high degree of usability the digital public services developed in the digitalization lab "special use of public spaces" will prospectively be accepted by future users. Regarding the qualitative goal set within the implementation of the OZG this means that the qualitative goal will prospectively be fulfilled.

Comparing the previous development of digital public services in Germany before the implementation of the OZG with the development of public services now within the implementation of the OZG, there is a fundamental difference. Within the implementation of the OZG the development of digital public services changed fundamentally from a restriction-centered to a user-centered perspective.

We assume that the reason for the fulfilling of the qualitative goal is the paradigm change in the development of digital public services from a restriction-centered to a user-centered perspective. Regarding the digitalization lab "special use of public space" this means, that the paradigm change led to the development of digital public services that have a high degree of usability and will accordingly be accepted by future users. In summary, this means that regarding the digitalization lab "special use of public space" the paradigm change towards the development of digital public services that are accepted by future users was successful.

Regarding the overall development of digital public services within the framework of the OZG we conclude that if the approach used in other digitalization labs for digitalizing public services will also be less restriction-centered and more user-centered like in the digitalization lab "special use of public spaces", the other digitalization labs also would prospectively fulfill the qualitative goal set within the implementation of the OZG. Within the bigger picture this would mean that all prototypes developed in the digitalization labs would have a high degree of usability. Translating these assumptions into the inspection of the huge paradigm change this means that the huge paradigm change towards the development of digital public services that are accepted by future users will prospectively be successful.

5. References

- [1] FOLEY, P., SUTTON, D. & u.a. (2018). Index für digitale Wirtschaft und Gesellschaft (DESI) 2018 Länderbericht Deutschland. Brüssel.

-
- [2] GRÜNWIED, G., (2017). Usability von Produkten und Anleitungen im digitalen Zeitalter. Erlangen.
- [3] HUNNIUS, S. and STOCKSMEIER, D., (2018). OZG-Umsetzungskatalog. Digitale Verwaltungsleistung im Sinne des Onlinezugangsgesetzes. Berlin.
- [4] IT Planungsrat "Digitalisierungsprogramm des IT-Planungsrats (Phase II): Start in die Verwaltungsdigitalisierung". https://www.it-planungsrat.de/DE/Projekte/Koordinierungsprojekte/Digitalisierungsprogramm/DigPro_node.html [03.07.2019].
- [5] IT-Planungsrat "Digitalisierungslabor zur Sondernutzung von Straßen - im Sprint zum digitalen Prototypen" https://www.it-planungsrat.de/DE/ITPlanungsrat/OZGUmsetzung/Digitalisierungsprogramm/05_DigPro_DigLabore/11_DigPro_DigLabore_DesignSprint_20190215/DigPro_DigLabore_DesignSprint_20190215_node.html [04.07.2019].
- [6] IT-Planungsrat "Digitalisierungslabore". https://www.it-planungsrat.de/DE/ITPlanungsrat/OZG-Umsetzung/Digitalisierungsprogramm/05_DigPro_DigLabore/DigPro_DigLabore_node.html [04.07.2019].
- [7] METZKER, E. and SEFFAH, A., (2004). The Obstacles and Myths of Usability and Software Engineering. Communications of ACM, S. 71 - 76.
- [8] NIELSEN, J., (1994). Heuristic Evaluation - Ten Usability Heuristics. <https://www.nngroup.com/articles/ten-usability-heuristics/>.
- [9] NIELSEN, J., (1994). How to Conduct a Heuristic Evaluation. <https://www.nngroup.com/articles/how-to-conduct-a-heuristic-evaluation/>.
- [10] NIELSEN, J., (2012). Usability 101: Introduction to Usability. <https://www.nngroup.com/articles/usability-101-introduction-to-usability/>.
- [11] SCHOLTZ, J., (2004). Usability Evaluation. National Institute of Standards and Technology.
- [12] STASCH, L., (2019). Die Verwendung von Digitalisierungslaboren zur Umsetzung des Onlinezugangsgesetzes - Eine Evaluation der Benutzerfreundlichkeit des Prototyps zur Sondernutzung im öffentlichen Raum. Ludwigsburg.