Usability Testing in the National Information Processing Institute, Poland

Katarzyna Turczyn
National Information Processing Institute
Al. Niepodległości 188 b, 00-608 Warsaw, Poland
email: katarzyna.turczyn@opi.org.pl

Agnieszka Lepianka
National Information Processing Institute
Al. Niepodległości 188 b, 00-608 Warsaw, Poland
email: agnieszka.lepianka@opi.org.pl

Abstract—This article describes the specifics of usability research in public institutions using as example the work at the National Information Processing Institute in Poland. It describes the challenges faced by system creators, designers and researchers. It presents methods of preparing and conducting usability tests (preparation, execution and further steps after researches). The characterization of the systems created in the Institute shows the specifics of working with public systems and shares the insights from these researches. Based on the gathered information, the article proposes changes to improve the experience of systems’ users. It presents good practices that the creators may follow during the design process, such as: naming, icons, charts, cohesion, searching, text editing and information architecture.

Keywords - UX research; usability testing; information system; system design; National Information Processing Institute.

I. INTRODUCTION

The National Information Processing Institute (OPI PIB) is a public institution whose tasks include development of Information Technology (IT) systems for the Ministry of Science and Higher Education in Poland [1]. An inseparable element of system design and development at the Institute is research and testing, in particular usability tests, which facilitate the detection of errors on websites and applications as well as shortcomings in their architecture.

A. Method

The article is the result of qualitative research, the main element of which was the participant observation, which began when we started working at the Institute in 2018 as researchers in the User Experience (UX) research team at the Institute’s Laboratory of Interactive Technologies. Data and information on which this article has been created was also collected thanks to many conversations with other OPI PIB employees, mostly other researchers working in the Institute several years longer, but also designers, analysts, developers and product owners. Working at the Institute and fulfilling our responsibilities also allowed us to analyse documentation and other available materials. In this case, Hastrup’s sentence: "real life is lived, not talked or written" [2] is true. Thanks to the possibility of analysing our experiences, we are a little closer to reality than we would be by analysing only what we have heard or read.

The article is intended to approach the form of Geertz’s thick description [3], trying to describe the broadest context of the topic.

B. Research Question

The aim of taking a closer look at the system development process at OPI PIB is to show the specificity of the job of a UX researcher in the public sector. An analysis of what takes place at the Institute in terms of usability testing has allowed us to indicate the pros and cons of the working environment in comparison to the ideal process of system design and development, as required by the principles and guidelines of User Experience. By looking at the current experiences, we will be able to identify some areas which could bring the process closer to the ideal if they are altered and optimised. The Institute's example allows us to take a closer look at how working on the software looks and, more precisely, how usability testing of this software looks in the public sector. The article can be a contribution to further work and reflection about the specifics of work in the public sector and how it differs from the private sector.

Section II introduces the research area, describing the systems produced at OPI PIB. Section III presents the way of conducting usability research in the Institute: preparation, execution and further steps after research. Section IV describes challenges faced by system developers and the results of usability tests. In Section V, we draw the conclusions from our qualitative research and present in-depth observations of research implementation at the National Information Processing Institute.

II. RESEARCH AREA: A DESCRIPTION OF THE SYSTEMS AND CHALLENGES FOR THEIR DEVELOPERS

Our work as researchers in the UX research team at the Institute’s Laboratory of Interactive Technologies has allowed us to take a closer look at the following system development projects:

- Polish Graduate Tracking System (ELA) – a system for secondary school graduates, students and university employees; the system contains statistical data on graduates’ earnings and employment obtained from the Social Security Institution (ZUS).
- POL-on – a database system on institutions related to higher education and science in Poland, designed for employees from the public sector, in particular university employees.
Integrated System of Services for Science / Streams of Financing (ZSUN / OSF) – a system which facilitates submission of applications for funding in the science sector (for students, doctoral students, research workers) and the subsequent handling of these applications by public administration entities.

Navoica – a free-of-charge educational platform with Massive Open Online Courses (MOOC) courses.

Uniform Anti-Plagiarism System (JSA) – an anti-plagiarism platform for verification of dissertation and thesis content.

As the project descriptions above suggest, the direct recipient of the systems is the Ministry, and the systems’ users are predominantly university employees (both administrative and research staff), as well as scientists, students, secondary school graduates and university graduates. The direct recipients of the system testing output, in turn, are employees of the Institute – the developers and creators of the systems.

The systems developed by OPI PIB are mostly database systems, mainly used by employees of the science sector. To a large extent, the systems reflect the processes that had been taking place at universities and their dean’s offices before computers appeared, when paper forms and files played the key role. The development of the existing systems required the digitisation of data and the establishment of software to reflect the previously applied "paper-based" procedures. That is how the ZSUN / OSF grant application filing system was developed (among other systems). The current systems, for the most part, have not only grown out of paper procedures – they have also retained a lot of the legacy features. In Poland, electronic documents are not yet regarded as equal to paper documents because of, among other things, the attitude of employees of public institutions [4], not just due to the existing legal framework. Therefore, in one of the phases of system use (usually the final phase), the user is often required to print out a document to close out the process. This is the case of the JSA anti-plagiarism system, where the final scan report must be printed out. It seems fair to say that paper documents continue to determine the interface of the existing systems, at least to a certain extent.

The formats of documents and processes, including digital ones, depend on the legal conditions, laws and regulations. An example of a system determined by legal acts is the POL-on system. Each of its modules is conditioned by different legal sources, e.g., the "Employees" module is based on several acts of law [5]. We can therefore say that non-intuitive information architecture of some system elements, a lack of certain functionalities or the presence of illogical requirements within the system are sometimes not the fault of system developers, but a consequence of the legal framework.

Users of the systems developed by the Institute, despite their often similar motivations to use the system, differ from each other on many levels. The differences impact the final interface of the system. First of all, users have different levels of digital competence, varying even within one group such as the group of researchers applying for grants. Another element which contributes to the diversity is the variety of the fields of science represented by the researchers using the system: for example, some people find it easier to understand a legal text, while others find legal texts challenging.

The heterogeneity of system users also results from the disabilities they may have. The Institute creates public systems, which, according to the law in Poland, must comply with the requirements of WCAG 2.0 [6]. The WCAG guidelines are the relevant benchmark, and the websites, as well as applications developed by the Institute, are designed to be as responsive as possible to the needs of people with disabilities. Some of the systems are intended mainly for researchers. The JSA system is used primarily by supervisors and reviewers of dissertations, i.e., persons who hold a PhD degree or higher degrees. In the years 2000-2010, the average age of university-nominated professors in Poland was about 55 [7]. System design standards are changing, and elderly users often transfer their experiences from other media (newspapers, books, paper forms) to the portals and systems they are expected to work with [8]. Younger users are impacted by the website services they use, too. Additionally, at an advanced age people are more likely to experience problems with vision and motor skills. If systems are not adapted to the requirements of this group of users, the hardware barrier may be the consequence, for example, when the buttons and fonts are too small and when the user interface is too complicated [8].

Much of the work at OPI PIB consists in introducing changes, transformations, extensions to the systems which were developed when the standards and requirements were different from the current ones. As a result, the developers are facing limitations from the very start of their efforts. In order to maintain the coherence of the systems and to stay within the budgetary constraints of the project, they sometimes have to give up some ideas. It works similarly in all other companies on the market.

The systems developed by the Institute are mainly aimed at supporting the Ministry of Science and Higher Education, universities and academics in collecting information, managing projects, acquiring funds and broadening their competences. These system development projects are not market endeavours, in which the most important result is for the customer to buy a product or service. The users of the systems developed by OPI PIB sometimes have to use them for work (the POL-on system), and sometimes in order to acquire a grant (the ZSUN / OSF system). In this environment, some system developers may feel they are monopolists, which may trigger the risk of disregarding the needs and requirements of the users as the level of competition and motivation to continuously improve decreases. Regardless of what the system will look like, its users will still simply have to use it. For this reason, it is necessary to carry out research and analysis based on data from the end users of the systems [9].
III. Usability Testing in the National Information Processing Institute

Since 2014, the number of usability studies and research carried out by the Institute has increased significantly. The evolution in research has not only changed the number of studies, but also increased their diversity. New techniques, such as usability tests, focus group tests and in-depth interviews, have been added to the previously used workshop method. Additional tools, such as co and card sorting were also used. An important element that became a permanent part of the work of researchers at OPI PIB was the UX audit of systems for designers’ and developers’ needs.

However, the most frequently used research method at the Institute is still task testing, which takes place at the Institute’s headquarters in a specially adapted testing room. Tests with one invited respondent allow the testing team to see how a potential user will use the product. Such tests [10] show how comprehensible the system is and where the critical points are that need to be modified in the first place. Moreover, this type of testing provides information on how intuitive the application and its system is, and whether it satisfies the needs of the users [10][11]. A big advantage of the tests is that the developers of the system are able to see live reactions of the respondents as they are interacting with the system. Furthermore, during task testing it is possible to ask in-depth questions which may have arisen during the test (this is a big advantage over tests conducted remotely). However, it is important to inform the recipients of test output that a single test will not answer all their questions, and that the number of tasks that can be performed during one session is limited. Consequently, test objectives and questions should be prioritised. As regards respondents, it is very important to ensure that they feel comfortable during the test, in particular if the respondent is a university employee who may feel that their knowledge and skills are being put to a test by an institution that supervises their work.

Focus groups are a less common data gathering technique at OPI PIB. It is used in the early stages of system design and in the redesign of existing systems. Thanks to a focus group interview with invited users or prospective users, it is possible to collect a large amount of information, insights, and translate them into conclusions and recommendations in a very short period of time (compared to other techniques) [11]. Focus groups often give direction to changes, provide information about users, their patterns of behaviour and expectations, enabling the researchers to use projection techniques and collaborative design [11]. The greatest risk in focus groups is associated with the role of the facilitator. Incorrectly facilitated tests may distort the results. If the facilitator is too withdrawn in the testing situation, one of the respondents may take over the role of the leader. Shy persons with little leadership energy may choose to avoid active participation in the conversation. Moreover, the Groupthink Syndrome can also occur. It is also a mistake to assume that focus groups can be the source of opinions about the entire population of users – in fact in only offers information about a segment of the population.

A. Preparation

The process of data gathering preparation at OPI PIB is presented in Figure 1. It concerns the implementation of the most frequently applied type of testing at the Institute, i.e., task-based usability tests in a test environment. The points on the vertical axis represent the degree of control over the process by the investigators and the probability of complications. The horizontal axis shows the course of the data gathering process in time. In the following section, the next steps of the research process will be discussed.

After learning about the needs and questions of the system’s designers, the next step is the recruitment process. Due to the specific nature of the public institution in question, the recruitment of respondents is implemented by an external company selected in a tender. The complicated nature of the recruitment procedure further complicates the selection of the best external companies which, additionally, need to meet strict tender criteria. The time needed to prepare the tests makes it difficult to integrate them into different phases of project development, so sometimes the only solution is to carry out guerrilla research. The process takes time, allowing the researchers get to know the system, ask research questions and create test materials. In creating a scenario, apart from the golden rules presented by Iga Mościchowska in her book [11], two more rules are applied:

1) The test scenario is not only for the researchers – everyone should be able to understand the tasks, questions and their purpose, so that people not involved in the scenario’s development have the opportunity to comment on it.

2) The respondent can look at the scenario during the test; he/she should not be able to find any hints or the facilitator’s expectations in the scenario.

The form of the scenario itself and its layout depend on the type of test and the facilitator’s preferences. However, the rigid rules of public institutions reduce any leeway: test materials are supposed to follow the established principles, and patience is required if researchers wish to introduce any changes.

As system developers are usually very busy, it can be challenging to ensure that all of them watch the focus group. Over the course of two years we have noticed a change in the approach to tests and test attendance at the Institute.

Information passed on to respondents covers two areas: first of all, practical information that will enable the respondents to arrive at the test location, e.g., a map with the location and the transportation suggestions. The second information area is any material and documents that the respondents will have to sign before starting the test procedure (e.g., respondent’s consent).

OPI PIB has its own focus group room and one observation room. In the focus room, in the case of usability tests, two spaces are arranged. In the first one, the test is introduced, the house rules are presented, and a short introductory interview with the respondents takes place. The second area is where the computer workstation is located.
“Freezing” test versions can be a tricky point, which must be kept in mind, not only when reporting the test need to the project manager prior to the test. Unfortunately, errors can occur during testing, as can system malfunctions. Postponing tests and not delivering test versions on time happen relatively often. It works similarly in the public and private sectors.

Based on the above can claim that the highest probability of complications is in situations where researchers depend on other people, not technology. This is why the researcher's soft skills and good cooperation with project team members are so important. Unfortunately, despite the fact that these skills are highly relevant, it is not easy for entities from the public sector to ensure and provide employee training in this area.

B. Implementation

Usability tests lasting more than an hour and a half can be tiring not only for the respondent, but also for the observers and the facilitator. Unfortunately, the need and capacity to perform tests usually materialises when an advanced version of the system is ready (often a production-ready version). Stakeholders then want to test the entire system. Individual user sessions almost never take only half an hour – they are usually 60 to 90 minutes long. Unfortunately, it is challenging for the respondents to remain active and attentive for a long time. Although the optimal length of the test is of crucial importance, it is often subject to negotiations with the respondents at OPI PIB. The three pillars on which good research results are based are appropriately designed tasks, conversations and the test’s overall atmosphere.

While creating the tasks, it is of key importance for all tasks to be natural and logical – they should minimize unnatural actions like logging out and logging in to another account. Tasks should not be interrupted with questions. There should be no suggestions to abandon a task before the respondent has expressed his/her wish to abandon the task. Importantly, the perception of the passing time is different for the observers and different for the person who is actually performing the task. Since the tested systems are often very comprehensive and have many functionalities, requests for reflection and questions to respondents are asked after a task or series of tasks, rather than after the completion of the whole test. Such conduct may yield more information than just the results of observation of the tasks performed. Employees of the science sector (users of previous versions of the system) may have many valuable reflections. This type of testing is no longer a classic usability test, but a hybrid with an in-depth interview, although in-depth interviews should typically take place before solution design.

The second foundation of good testing is the conversation. In the introduction, it is always a good idea to inform the respondents what the tests will look like and how long they will take [12], as well as telling them tell about the possibility of task interruption. The investigator should also allow the respondent to ask questions in order to make the respondent feel more confident in the new situation and speak more freely. The most important information to be conveyed to the respondent is that he or she is not going to be evaluated – this seems particularly important when working with employees of the science sector as they sometimes perceive our institution as superior and affiliated with the Ministry of Higher Education and Science.

As part of the third pillar of testing, the facilitator should radiate positive energy and develop a friendly and open test atmosphere. It is essential for the facilitator to be empathetic. The results of the test largely depend on the facilitator's involvement in building a positive, relaxed atmosphere conducive to the respondent’s cooperation and information-sharing.

Insights from the tests and interpretations based on user feedback are provided in Section IV.

C. After the Tests

Before the test report is produced and after completion of testing, researchers at OPI PIB typically organise two summary meetings. The first one is informal and aims at discussing the results with the researchers involved in the project. The second one involves the stakeholders and is aimed at discussing the most important observations, and, if possible, should be organised within a short period of time after the end of the testing procedure. The meetings are also associated with the need to build relationships between members of different teams, which would allow the project teams to work in a more agile and dynamic way, abandoning some procedures from the waterfall/cascade model and improving some standards of work within creative/UX teams [13][14]. After the report has been created, the first thing to do is to establish the date of its presentation, before the report is sent out to stakeholders. If there is no set date, the stakeholders may find it challenging to find time to meet later. It is a good idea to remember to send the report out to the stakeholders (mainly designers) before the presentation. This is sometimes due to the fact that the designers may be slightly anxious as to whether the report will show their work in bad light. They may also feel that their contribution is being evaluated.

It is good practice to determine the progress in introducing changes sometime after the test, as well as determining if the designers have all the information necessary to implement the necessary modifications. Their continuous interest in the subject increases the probability that the proposed changes will actually be implemented. The abundance of responsibilities and, to a larger extent, the formal procedures in place at the Institute, make it difficult for the stakeholders to meet regularly. As people work in different teams, official appointments for every meeting need to be made, in most cases involving leaders and managers whose availability is very limited. Although managerial presence at the meeting is not always necessary for the quality and efficiency of the meeting, meetings often cannot be held without them. These formal requirements are also responsible for a formal meeting atmosphere that hinders an unrestrained and free exchange of ideas. It is also likely that the physical work environment influences communication between teams. Open space office
arrangements encourage conversations between employees, while at the OPI PIB people work in four different buildings.

IV. CHALLENGES AND USABILITY TESTING RESULTS

A. Challenges for Researchers and Designers

At OPI PIB, teams of UX researchers, UX designers and developers work across different departments. This specificity of work organization at the Institute means that people involved in the development of systems are separated from each other, and this undoubtedly hinders cooperation. As a result, researchers have a limited capability to monitor the further development of the product after the testing is completed. Therefore, it is important that the project team members cooperate closely and have frequent contacts to create product concepts together, co-design, co-develop and monitor further product use.

At OPI PIB, like in many other institutions and companies in Poland, UX testing is introduced at a fairly advanced stage of product development, triggering the risk of much higher costs of implementation of changes [15]. The benefits of early-stage UX testing include the ability to verify the identified target group, define the real needs of users, and investigate the initial concept of the system and its architecture.

The systems developed at OPI PIB are mostly commissioned by the Ministry of Science and Higher Education and are intended primarily for users from HEIs [16]. The systems are mainly based on regulations and laws, which directly affects the structure and functionality offered by these solutions, as well as imposing design constraints on developers and limiting the available options.

Due to the specificity and intended use of the systems designed and developed at OPI PIB, they are used mainly on desktop or laptop computers, so there has so far been no need to create mobile versions. However, the situation is now changing and new systems are being developed, designed also for users from beyond the academic community – therefore it is becoming necessary to develop mobile versions of the system, too. In this respect, the public sector is beginning to operate like the private sector – “Mobile first” is beginning to apply.

Some of the systems created by the OPI PIB are digital versions of various types of paper forms and application forms. The designers face the challenge of creating functional and user-friendly forms ensuring an easy fill-out process. Unfortunately, despite the availability of digital versions, users still have to print out paper versions of forms as well.

Increasingly, stakeholders appear as observers in tests because they recognize tests as a great opportunity to see how users interact with the systems and what problems they encounter. Thanks to participation in such processes, stakeholders can count on receiving prompt feedback on their effort, without having to wait for the final report. Behind-the-scenes conversations in the viewing room also offer a good opportunity to discuss and exchange ideas on how to design systems, or how to redesign them. Thanks to such conversations, it is possible to learn about the limitations of both designers and users which for some reason did not surface at project meetings preceding the testing phase.

Remarks and comments from the post-test reports should ideally be introduced into subsequent versions of the system, with the critical points repaired of necessary. Unfortunately, this scenario is only implemented in 50% of the cases – not due to lack of involvement or bad will, but to a large extent due to the binding legal constraints.

B. Usability Testing Results

The numerous usability studies and conversations with end users we have conducted at the Institute have allowed us to outline the principles that designers should bear in mind when creating systems. Here are some of the recommendations:

1) Names, headings and keywords. Users quickly browse the website for specific keywords, sentences or paragraphs and skip most of the text. It is therefore important to organize the content, group the elements and assign headings, titles and labels to them in an appropriate way. The terminology used should be simple, clear and understandable to all system users.

In state institutions, some of the terminology that can be found on websites is borrowed from acts of law, regulations or technical documentation. This leads to lack of terminological clarity for users and difficulties in understanding the content. Some system designers have recognised this problem and the need to introduce plain language so that users with different levels of education and knowledge can understand the text. Consequently, public administration entities now employ a growing number of experts in their UX teams (UX writers).

2) Icons. They help users remember content more easily and quickly, making the message more interesting. It is important to remember that icons should be adequate to what the system is supposed to communicate to its users. In the case of database systems, icons can help users understand content more easily.

3) Diagrams and graphs. They should be understandable and legible. Remember to include explanations and legends. It is worth noting that the graph and its description should be visible on one screen at the same time so that the user does not have to scroll between the graph and its description. System developers know what information and data they want to present on charts and assume in advance that their preferred way of presenting data will be clear and legible for the users as well.

4) Short texts. Large blocks of lengthy text are not attractive and discouraging to users. Text should consist of short or medium-length sentences grouped into paragraphs. The content can also be split into bullet points. Furthermore, the users who would like to find out more need to be taken into account as well – include a link to a page with extended information.
5) **White space.** System developers often misinterpret system legibility as a lot of white space on the screen. Such an approach to design is often counterproductive, since users who interact with such a site assess it as poorly designed. Designers believe that by giving up illustrations and graphics, they can avoid the superfluous content characteristic of commercial websites overloaded with advertisements, pop-ups and banners. This misguided ascetic approach may cause the system to be perceived, on the one hand, as clear and transparent, but on the other hand – as overly rigid and official due to the excessive amount of white space. During the tests referred to in this paper, the respondents pointed out that in many systems designed at OPI PIB there is too much "vacant" space. They believed that the blank area could have been better utilized to accommodate more text and condensed content. A large amount of white space may mean that there is little meaningful content on the screen, and the user has to scroll down to find out more.

6) **System coherence.** It is important that all elements of the system should fit together and the construction of the site should be coherent. The design should be tailored to the needs and expectations of the user. Consistency of the components makes the design intuitive, easy to navigate, and easy to use. The systems developed at the Institute are comprehensive and complex. Due to changes in legal provisions and for other reasons they must be updated from time to time.

7) **Searching, querying, sorting and display of results.** User queries should be as easy as possible. It would be good if the search results covered the whole system, not only its selected part or category. Filters should be designed so that the user can select several variants of the same feature. Users should also be given to the possibility to enter keywords with spelling mistakes, typos and incorrect conjugation. It is very important to present the search results properly, displaying the searched information or its fragments in the format expected by the user. Unfortunately, public institutions do not always want to rely on the industry’s best practice and solutions – instead, own solutions are created which are frequently neither proven nor tested. Perhaps such an approach is associated with the misconception that looking for own system solutions is a way to avoid being accused of plagiarism.

V. **Conclusion**

The aim of our research was to show the specificity of the UX research in the public sector in Poland and identify areas which could be changed in order to bring the process of developing and designing systems closer to the ideal. After analysing the presented experiences and data and with a view to facilitate a further development of research activity at OPI PIB, the authors’ aim is to ensure that the UX research of systems developed at the Institute are conducted at various stages of product development. During the tests and research conducted for this paper we noticed how important it is to apply different research techniques, appropriately matched to the given development phase of the system. An important task is also to change the attitude of designers and developers of systems to users. It is essential that system designers focus mainly on users and their needs, and that they take into account users’ limitations.

In order to achieve these goals, internal seminars can help to present the work of UX researchers and the entire testing process. It is also beneficial to indicate how both system developers and users can profit from properly delivered testing. Effective communication of said benefits can be facilitated by issuing reports to provide stakeholders with more information about users and the testing procedure. It is also necessary to organize meetings with stakeholders as often as possible in order to talk about their needs and indicate possible solutions. It is important to give users a sense of security, support and space for their creativity. Communication is the foundation for creating systems that will match users’ expectations. Furthermore, relations with other teams and the ability to communicate effectively and efficiently are important elements of the work of researchers. Additionally, researchers ought to continuously improve their competences and acquire new knowledge. It is of key importance that they continue to develop by attending training courses, reading professional literature or taking part in conferences where they can exchange their ideas and observations with other researchers.

**References**


Figure 1. Research preparation process in the National Information Processing Institute in Poland.