HEURISTIC AND THINK-ALOUD APPROACH FOR EVALUATE USABILITY IN ACADEMIC INFORMATION SYSTEM

\textsuperscript{a}Satrio Agung Wicaksono, \textsuperscript{b}Retno Indah Rokhmawati, \textsuperscript{c}Mochamad Chandra Saputra, \textsuperscript{d}Raden Arif Setiawan
\textsuperscript{a,b,c,d}Information System Department, Faculty of Computer Science, Universitas Brawijaya, 8 Veteran Road, Malang, 65145, Indonesia
E-mail: satrio@ub.ac.id

Abstract

Academic Information System is an essential system for performing academic activities. Universitas Brawijaya is using Academic Information System named SIAKAD-UB. SIAKAD-UB is an information system that deal with all kinds of student details and academic-related reports. SIAKAD-UB encountered many problems along with the growth of data. This study aims to discover factor impact on problems of interface usability found in existing SIAKAD-UB using Heuristic Evaluation and Think-Aloud method. This study involving 3 experts and 3 operators with the purpose of the evaluation received input from the experts and users. The result of experiment are found 3 problem heuristic with score 0 which mean no usability problem, 7 problem heuristic with score 1 which mean medium priority refinement, 7 problem heuristic with score 2 which mean low priority refinement, 7 problem heuristic with score 3 which mean high priority refinement. Heuristic evaluation and Think-Aloud find 7 aspect refinement are Visibility of system status, Match between system and the real world, User control and freedom, Consistency and standards, Recognition rather than recall, Flexibility and efficiency of use, Help and documentation.

Keywords: academic, heuristic evaluation, information system, think-aloud.
INTRODUCTION

Academic Information Systems are computer information systems that use for advising, course catalogs, student information and data, including financial aid, educational administration process, supporting the institution's research and educational endeavors [1]. Universitas Brawijaya is using Academic Information System named SIAKAD-UB. SIAKAD-UB is an information system that serve with all kinds of student details and academic-related reports on desktop based application. The features of SIAKAD-UB are interface are (1) user interface designed especially for academic officers; (2) facilitating academic data management; (3) Facilitating for academic courses schedule; (4) facilitating for curriculum transformation. SIAKAD-UB is in line with information technology service principle about efficiency and effectiveness of the work of all fields and improvement of work productivity.

The problem of SIAKAD-UB is there are many user complaints about the user interface. User complaint about the effectiveness of SIAKAD-UB that have a problem with accuracy and completeness to achieve specified goals like show the report of academic course. The report does not show the result as user needs. The problem shows that the users needed some development to solve the problem about easier to use, and matching SIAKAD-UB more closely to user needs and requirements. The last problem, based on the information from stakeholders there are no evaluation before about usability in SIAKAD-UB. One of the characteristic successful of software is usability.

Usability evaluation used for identifying the problem of using the SIAKAD-UB and then measuring aspect of usability such as problem that identified by the user. The advantages of great usability are increase productivity, decrease error from user, increase the acceptance of the system, improve the performance of the system.

The phase of usability evaluation in this research discussed with evaluating usability variable using heuristic evaluation and then using think-aloud we gather the information about user interaction between user and SIAKAD-UB with some task and the relation both of them.

HUMAN COMPUTER INTERACTION (HCI)

Human-Computer Interaction (HCI) is a multi-disciplinary focus on the design, evaluation and implementation of the interaction of computer systems used by humans and the environment [2]. Usability is a crucial issue in HCI, it’s become an important aspect to assessing the quality of the user interface [3]. The focus on HCI is about humans interaction to use computers as a tool to perform, simplify and support their tasks [4].

Usability is a quality attribute that is used to assess how easily the interface is used by the user. Usability is defined by 5 quality components [5]:

1. Learnability: measure how easy is it for users to accomplish basic tasks the first time they encounter the design
2. Efficiency: Once users have learned the design, how quickly can they perform the tasks.
3. Memorability: When users return to the design after a period of not using it, how easily can they reestablish proficiency.
4. Errors: How many errors do users make, how severe are these errors, and how easily can they recover from the errors.
5. Satisfaction: How pleasant is it to use the design.

There are many other important quality attributes. A key one is utility, which refers to the design’s functionality.

HEURISTIC EVALUATION

The purpose of a heuristic evaluation is to improve the design effectively. Evaluators perform evaluations through the performance of a series of tasks by design and customized to the criteria of each level. If a detected error can be reviewed for improvement before the implementation stage [6].

Evaluation of heuristic versions of Molich and Nielsen [6] as an approach in evaluating a machine maneuver system using the use (usability). The 10 principles of heuristic evaluation are visibility of system status, match between system and the real world, user control and freedom, consistency and standards, error prevention, recognition rather than recall, flexibility and efficiency of use, aesthetic and
minimalist design, help users recognize, diagnose, and recover from errors, help and documentation.

Identifying usability problems through a heuristic evaluation is the first step towards eliminating problems and improving the interface [7]. Ranking of usability problems by severity rating helps to determine those that should be addressed. The ratings also help in the allocation of resources for addressing the user interface problems. Nielsen [8] arranges a scale of 0 to 4 to represent the level of severity occurring in a system in terms of usability, as shown in Table 1.

According to Nielsen, severity is considered to be a combination of three factors: frequency, impact, and persistence. Severity ratings for all the usability problems can be found by sending a questionnaire to each inspector once the evaluations process has been completed.

Table 1. Severity Rating

<table>
<thead>
<tr>
<th>Rating</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>I don't agree that this is a usability problem at all</td>
</tr>
<tr>
<td>1</td>
<td>Cosmetic problem only: need not be fixed unless extra time is available on project</td>
</tr>
<tr>
<td>2</td>
<td>Minor usability problem: fixing this should be given low priority</td>
</tr>
<tr>
<td>3</td>
<td>Major usability problem: important to fix, so should be given high priority</td>
</tr>
<tr>
<td>4</td>
<td>Usability catastrophe: imperative to fix this before product can be released</td>
</tr>
</tbody>
</table>

THINK-ALOUD

Simple usability tests where users think out loud are cheap, robust, flexible, and easy to learn. Thinking aloud should be the first tool in your UX toolbox, even though it entails some risks and doesn't solve all problems. The method has a host of advantages. In particular, you hear their misconceptions, which usually turn into actionable redesign recommendations: when users misinterpret design elements, you need to change them. The term ‘think-aloud protocol’ refers to a type of research data used in empirical translation process research. The data elicitation method is known as ‘thinking aloud’ or ‘concurrent verbalization’, which means that subjects are asked to perform a task and to verbalize whatever crosses their mind during the task performance [9].

Think-aloud protocols involve participants thinking aloud as they are performing a set of specified tasks. Participants are asked to say whatever comes into their mind as they complete the task. This might include what they are looking at, thinking, doing, and feeling.

MEASURING USABILITY

Most computer software in use today is unnecessarily difficult to understand, hard to learn, and complicated to use. The problem is related to theory of usability. Usability are the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use [10]. Usability issues should be monitored during development, and the usability achieved should be evaluated.

The ideal way to specify and measure usability would be to specify the features and attributes required to make a product usable, and measure whether they are present in the implemented product [11]. Usability is a property of the overall system. It is the quality of use in a context. The existing methods for predicting usability are limited in their accuracy as they only model limited aspects of the users, the tasks and environments.

This research applies two methods, such as heuristic evaluation and think-aloud, because the results of this evaluation need to be validated from both of the expert and the user point of view. In think-aloud, user verbally expresses thoughts, feelings, actions, and experiences during interaction with the system [9]. Heuristic evaluation is easy to perform, cheap and able to find many usability problems (both major and minor problems). That is why the use of appropriate heuristics is highly significant [12].
GENERATION OF EVALUATION HEURISTIC

As mentioned before, this evaluation applies heuristic evaluation and think-aloud protocol. Both methods are performed simultaneously in different places but with the same scenario, accessing the main features of SIAKAD-UB. Here the steps in conducting heuristic evaluation in this study:

1. Establish an appropriate list of heuristics.
   We have carried out information on what usability aspects can be measured by applying heuristic evaluation and think-aloud observation techniques. Based on observation and using the approach suggested by Nielsen and Molich [1], these 10 rules on Heuristic Evaluation: (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; (7) flexibility and efficiency of use; (8) aesthetic and minimalist design; (9) help users recognize, diagnose and recover from errors; and (10) help and documentation.

2. Select the evaluators. In this study, the heuristic evaluation is used to evaluate SIAKAD-UB’s interface for operator as a user. In order to evaluate the problem using heuristic evaluation method, three experts were involved. The criteria of evaluator is a expert or developer of desktop and web-based information systems [5]. Many of the "problems" identified by heuristic evaluators were not problems at all [2].

3. Brief the evaluators so they know exactly what they are meant to do and cover during their evaluation [3]. Scenarios performed by the evaluator focus on accessing login, academic page, and curriculum page.

4. Evaluation phase. The evaluators will use the product freely to gain a feel for the methods of interaction and the scope [3]. They will find out the specific elements that they want to evaluate. The evaluators will carry out another run-through, whilst applying the chosen heuristics to the elements identified during the evaluation phase [3]. The evaluators would focus on specific or individual elements.

5. Record problems. The evaluators record problems themselves and be as detailed and specific as possible when recording problems. Evaluator writes down the evaluation on the questionnaire that has been provided.

6. Debriefing session. This session involves collaboration between evaluators and researchers to do the validation and suggest potential solutions for these problems on the basis of the heuristics. In this phase the problems were classified based on the severity rating and 10 principles of heuristic evaluation.

This study try to determine what heuristic evaluators thought the usability problems would be, and compare their responses with the problems that users actually had. Users’ problem collected by using think-aloud method. Here the steps of Think-Aloud Protocol:

1. Recruit representative users. Three users involved are operators who manage the data using SIAKAD-UB in their daily job.

2. This research applies Think-Aloud model Coaching Condition type, because it is considered more accurate to describe the perceived condition of the user. This model also known as Think-Aloud with active intervention [13] (i.e., more verbal feedback and probes where test administrator asks direct questions about different areas of Web site, such as areas where user is having difficulty/is pausing/or is describing area as confusing or frustrating; gives help or assists when participant is struggling; includes practice session before testing begins).

3. Give users representative tasks to perform. Users are required to access the main features of SIAKAD-UB, such as input course journals, data of student attendance, etc. When stuck, a participant was encouraged to continue and was given some direction on how to do the task.

4. The moderator who is a researcher recording entire user response during the Think-Aloud process.
CONTEXT OF THE STUDY

SIKAD-UB is an information systems that serve with all kinds of student details and academic-related reports on desktop based application that was developed by the Information and Communication Technology Unit of Universitas Brawijaya. Case study of this research was taken at the Faculty of Computer Science (FILKOM) Universitas Brawijaya. Even though SIKAD-UB has been implemented in FILKOM about 6 years, operators still find obstacles when interacting with SIKAD-UB. Hence, this research was conducted to investigate the problems SIKAD-UB usability.

SIKAD-UB is an application for academic operators. SIKAD-UB contains student academic transactions. In addition, it contains lecture data (lecture schedule, student attendance, course name, lecturer etc.). SIKAD-UB can also support curriculum change (merging subjects, change of credits, name change etc).

RESULT AND DISCUSSION

The results of the heuristic evaluation and think-aloud have confirmed each other, so that improvements are based on the needs of both the expert and the user. To simplify the analysis, the result is categorized by 10 rules on Heuristic Evaluation and explained the following:

1. Visibility of system status (1st rule)

Visibility of system status is a heuristic evaluation aspect focused on how the system provides a visual interaction of the processes that are running on the system. In the process of heuristic evaluation found one kind of usability problem with a scale of 2 based on severity rating, which is included in the minor usability problem. The problem is no information about system changes that should be included on every new release. Based on the Think-Aloud result, this problem causes the user does not understand the exact changes experienced SIKAD-UB from each version.

Based on CGAP Technical Guide Information System [14], for easing the transition, developers need to do several things such as training, user guides, or staff feedback. Training is appropriate if SIKAD-UB involves new staff in its implementation. The training aims to ensure that staff receives proper and comprehensive training, not just “as needed” information from one of their colleagues.

In this case user guide is required to minimize the possibility to make a guess and a subsequent mistake. Developer need to ensure these guides readily available to all staff and regularly update. The easier it is for them to gain access to easy-to-follow instructions, the less likely they are to call on the IT Department for support.

2. Match between system and the real world (2nd rule)

Applications should be able to provide comfort by displaying language and display that is easy to understand and in accordance with the user environment. Most problems in this aspect of the feature does not match the needs of the user or there are still features that are not used properly. The problems related to this dimension are further explained in Table 2.

Based on the results of Think-Aloud, operators want a simpler submenus, and buttons with related functions should be grouped. The potential solutions for the problem is rearrange submenu. Designers of web pages for multi-page content should place submenu navigation immediately adjacent to the content’s text and that all other site navigation should be located in more distant portions of the page that are visually unassociated with the article [15].

<table>
<thead>
<tr>
<th>No</th>
<th>Problem</th>
<th>Severity Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The placement of “Timetable Menu” is less precise and leads to errors</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>The process of calculating and storing values takes a long time.</td>
<td>3</td>
</tr>
</tbody>
</table>
The display of real-time calculations is very important in this case, as it affects the operator's efficiency. When users experience an online inspiring, involving and real-time result experience, they participate more intensely [16]. It is recommended to the developer to apply the optimal calculation using 3-tier architecture. This architecture, the client only processes the logic, and does not require much resources and a bit of coding required by the client [17]. This may improve the responsiveness of the display.

3. User control and freedom (3rd rule)

User often need a freedom of control for using and exploring in-app features. Sometimes, user choose a navigation by mistake and need a "emergency exit" to leave the unexpected state [18]. One of the problems that often arise in this aspect is navigation. The hidden navigation placement causes the user to be unaware of the existence of the navigation. The problems related to 3rd rule are further explained in Table 3.

Considered on the severity rating, this problem is not a priority, but it is interesting to discuss. Based on Think-Aloud results, the operator wants additional navigation should not be placed close to the main navigation, such as the academic calendar. It can distract operator's focus when interacting with SIAKAD-UB.

Navigations should not be tedious and should be consistent with other systems [19]. Navigation should allows a user to access for a specific goal as quickly and easily. It is recommended to rearrange the sitemap based on user needs to avoid the incoming submenu in the irrelevant navigation category.

Table 3. Problem and Severity Rating in Third Rule

<table>
<thead>
<tr>
<th>No</th>
<th>Problem</th>
<th>Severity Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Student presences are included in the submenu causing some users to have trouble finding the menu.</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 4. Problem and Severity Rating in Fourth Rule

<table>
<thead>
<tr>
<th>No</th>
<th>Problem</th>
<th>Severity Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The window size is too small for large data inputs</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>There is some navigation hidden in the sub menu</td>
<td>1</td>
</tr>
</tbody>
</table>

4. Consistency and standards (4th rule)

Consistency and standards is a heuristic aspect that requires an app to have clear consistency and standards, so it's not confusing to the user. The problems that often arise is navigation. The hidden navigation placement causes the user to be unaware of the existence of the navigation. The problems related to 4th rule are further explained in Table 4.

When Think-Aloud is conducted, the operator complains that the window size is too small for large data input. In one day, an operator have to input student attendance data with an average of 40 students per class. In one day, operator have to input the course journal with an average of 15-20 class per day. It indicates that the operator requires an interface that can support the operator to stay focused during input process.

Navigations, simple content area, proper button placement, and simple goal is to get user involved. But when the screen size is too small, it can distract the work of the operator/user. Main navigation should be intuitive dan easily identifiable. Navigation labels should clear, concise, and not hidden in the sub menu [20]. Good navigation is when the user is able to realize where they are now, where they should go next, and where exactly they were before.

The difficulty of reading on small screen size or window size, which makes less content or information visible at one time [21]. In the case of a seemingly claustrophobic display screen, web designers need to try and divide information into clearly defined sections. This section should be organized the way of visitors looking for information, not how the units itself is organized [20]. Web designer also need to ensure that content display appropriately for the device.
5. Recognition rather than recall (6th rule)

An application ideally meet the recognition rather than recall aspect by directing the user not to always remember something important and let the user recognize the system and information in the application. This aspect is identical to the characteristic of the application, so users who have often used the application are accustomed to using the app. The main problem in this aspect is that academic information cannot appear if the test scores entered in the system is incomplete. The experts and operators assess this issue including the second level of severity. From Think-Aloud result, researchers identified the impact of this problem is that the operator takes a long time to search for student academic information and needs to remember some past information to complete the task.

Users should not have to remember a lot of information to carry out the tasks, because memory load reduces users’ capacity to complete the main task. In this case, systems should provide information, objects, options, and instructions clear and visible, so that user need not remember unimportant information [22]. The speed of users can learn how to use something or to complete the task depend on how much they can remember/memorize information and how efficient when they using it [23], and early research proves that ease of use and memorability affect each other [24]. Whether the data is complete or not, the system should display the existing student’s test scores.

6. Flexibility and efficiency of user (7th rule)

The problem in this aspect is generally about how the application provides easy settings and does not make users confused in using the application. Ideally, the settings can be changed according to user needs easily. The problems related to 7th rule are further explained in Table 5.

From Think-Aloud result, researchers identified the impact of this problem is that the operator must calculate the student’s GPA manually if the information is needed immediately. Reports should be information that can be easily understood and can even be used to make decisions.

Table 5. Problem and Severity Rating in Seventh Rule

<table>
<thead>
<tr>
<th>No</th>
<th>Problem</th>
<th>Severity Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The student's GPA is not automatically updated when the study period has ended.</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>Reports are still in the form of data, not yet easily understood information</td>
<td>2</td>
</tr>
</tbody>
</table>

Students’ GPA data should be automatically calculated by the system and displayed informatively. It is proper that the GPA can even be predictive of student achievement [25]. For future research, it is suggested to develop students’ academic data to predict the potential crisis point, so it can be known that student performance who will be successful, raising, or failing. The visualization of the students data can be seen in [21]. This research classifies student performance to be successful, raising, and failing. If it can be known earlier, then student failure can be prevented immediately. The dashboard visualization is getting better when defining the needs of the system, first identified key performance indicators (KPI) and critical success factors (CSF) [26].

7. Help and documentation (10th rule)

The issues in this aspect are generally about whether or not support documentation and Help features are available for users who have difficulty or have questions when using the app. The main problem in this aspect is there is no feature FAQ or Help. So that new users or operators who have not received training will have difficulty using this application. The experts (Heuristic Evaluation) assess this issue including the second level of severity. Through Think-Aloud, the operator specifically requires the support of information and Help features if the operator has difficulty when using SIAKAD-UB. Good FAQ pages use legible typography, chunking, appropriate spacing, easy navigation to individual questions, and reflect the current questions of the site users [27].
Evaluations from both the expert and the user do not find problems in the following rules: Error prevention (5th rule), Aesthetic and minimalist design (8th rule), and Help users recognize, diagnose and recover from error (9th rule).

Based on interviews with operators (users), error prevention is not necessary because users have interacted with the system in a long time, so the possibility of fatal error can be avoided. It is suggested in future research to involve potential users or new users to better explore the model of error prevention as required.

Related aspects of aesthetics, experts assess the appearance of the system is appropriate as the academic information system. Operators also do not expect any design improvements. Because the simple design according to them just make it easier to do the job.

Related with help users recognize, diagnose and recover from error rules, the operators does not find the problem. But the operators suggests if there is an error message better delivered with a plain language (not codes).

Suggestions for future research is combining between heuristic evaluation and persona to make the results more user-centric[3]. The future research could compare the findings of Heuristic Evaluation with Think-Aloud and with personas. Most of studies stated that heuristic evaluation was most successful when carried out by expert, but they had a tendency to report issues that were not listed among the given heuristics.

CONCLUSION

Problems that arise during use may be completed with the proposed solution above. System development should actively involve potential users. Particularly, if the system will be used for work every day. The important thing is the selection of respondents who represent the character of the user. The academic information systems accommodate the needs of all faculties with a wide range of disciplines. This leads to varying degrees of difficulty and type of report. Usability plays an important role in measuring the level of ease of an application used by the user to complete the task. With proper user selection, two-party evaluation, for example, Heuristic Evaluations and Think-Aloud can objectively explore what needs to be improved and how the priority of improvement.

As a precaution of usability problems, users can be involved when extracting needs and prototyping. It starts from an open question about the main activity and background. System developers are recommended to identify the triggers and conditions in the activity. The prototype can actually be used as a way of usability testing. Users should be given control to determine what kind of design that fits and supports their tasks, so it needs a prototype that is capable to visualize the user’s needs. From these activities will generate a user goal that describes the stages of the user in completing the task. User goal can be used as the fundamental to create a user task to evaluate usability level when the system is ready to use.

REFERENCES


