DESIGN AND IMPLEMENTATION OF WEB-BASED LECTURE EVALUATION SYSTEM

Winarno
Institut Agama Islam Negeri (IAIN) Salatiga
Jl. Tentara Pelajar No. 2, Salatiga, Jawa Tengah, Indonesia 50721
Email: winarno@iainsalatiga.ac.id

ABSTRACT

This study aimed to produce a web-based lecture evaluation system in the Post-Graduate Program of IAIN Salatiga. The developed system was called SEVADO (*Sistem Evaluasi Perkuliahan Dosen*/Evaluation System of the Lecturer's learning). The method of this research was research and development (R&D). The data collection techniques were documentations, interviews, focus group discussions (FGD), and questionnaires. The research subjects were 10 students during the pilot study and 46 students in the main research. The aspects which was tested includes functionality, usability, reliability, performance and supportability. To test the aspect of reliability and performance, WAPT tools (Web Application Performance Testing) software was applied. The researcher concludes that the web-based lecture evaluation system of the Post-Graduate Program of IAIN Salatiga was feasible. It met the five aspects of system testing: functionality (score 4.2/excellent category), usability (score 4.3/excellent category), supportability (score 4.3/excellent category), performance (97%), and reliability (92%).

Keywords: Information Systems, Evaluation System, Web-based Evaluation System.

ABSTRAK

Makalah ini memaparkan tentang pengembangan sistem evaluasi perkuliahan dosen berbasis web di program pascasarjana IAIN Salatiga. Sistem yang dihasilkan diberi nama SEVADO (Sistem Evaluasi Perkuliahan Dosen). Penelitian menggunakan metode penelitian dan pengembangan. Pengumpulan data menggunakan teknik dokumentasi, wawancara, diskusi terarah dan kuesioner. Subyek pada uji coba terbatas sebanyak 10 mahasiswa dan pada uji coba diperluas menjadi 46 mahasiswa. Untuk mengetahui kelayakan, sistem diuji berdasarkan lima aspek, yaitu functionality, usability, supportability, reliability dan performance. Untuk mengetahui reliabilitas dan kinerja, sistem diuji menggunakan prangkat lunak WAPT (Web Application Performance Testing). Hasil analisis data menunjukkan bahwa sistem evaluasi perkuliahan dosen berbasis web di Program Pascasarjana IAIN Salatiga dinyatakan layak, yaitu; aspek functionality dengan perolehan skor 4,2 (kategori sangat baik), aspek usability dengan perolehan skor 4,3 (kategori sangat baik), aspek supportability dengan perolehan skor 4,3 (kategori sangat baik), aspek performance dengan persentase keberhasilan sebesar 97%, dan aspek reliability dengan persentase keberhasilan sebesar 92%.

Kata Kunci: Sistem Informasi, Sistem Evaluasi, Sistem Evaluasi Berbasis Web.

INTRODUCTION

University lecture as a form of learning process involves planning, implementation, and regular evaluation. In terms of evaluation, it mostly focuses on the scope of learning outcome. On the other hand, the scope of lecture evaluation assessed by students still requires supports for implementation (Suherman, 2003, p. 106). Lecturers are obliged not only to evaluate learning outcome but also to accommodate students' evaluation on lecturers' performance during the learning process. It means that lecturers are responsible for both students' achievement and their own performance in teaching. The statements show that course subject evaluation actually applies to students and lecturers (Jabbarifar, 2009; Mardapi, 2008, p. 3).

Lecturers are capable of improving the teaching process through deep observation on the data of students' perception on the teaching performance, and understanding of students' learning style. It enables them to find students' perception about their methods, approaches, and teaching strategies. In order to do so, they need to populate data from their students each semester (Marmah, 2014; Matiru, Mwangi, & Schlette, 1995, p. 56).

Lecture evaluation benefits teaching staff, according to Mas'udi's (2014) arguments. They state that evaluation provides beneficial report concerning strength, weakness, and potential in learning process. In details, Arikunto (2008, p. 7) and Nuriyah (2014) explicate that by having evaluation, teaching staff will find out the material appropriateness for their students. Teaching performance will improve if teachers – or in this sense is lecturers – receive regular, particular, and conceptual reports. In the level of higher education, the response from student evaluation can be the data sources to improve lecturers' performances (Nuriyah, 2014).

The function of lecture evaluation is to monitor and control academic quality. The evaluation is required in order to maintain the quality of lectures. Matiru, Mwangi & Schlette (1995, p. 284) argue that after a lecture ends, regular monitoring and revision are needed to keep the academic quality standard. Both lecturers and universities are responsible for dealing with evaluation, so that the quality standard can be exposed to public.

A university with high quality standard will be superior to other universities. On the contrary, higher education with no monitoring and maintenance of such standard will have lower credibility and get negative perception from public in the long run.

To date, the application of lecture evaluation system in the Post-graduate Program of IAIN utilizes the paper-based format in which students assessed their lecturer based on the provided instruments. Considering the current technology development, the application of that evaluation system possibly uses computer technology on the basis of web development (Henderson, Selwyn, & Aston, 2015; Imai et al., 2017; Koehler, Correia, Alpay, & LeVally, 2017). The benefits of information technology are efficiency, and accuracy in grading, and speed in viewing the result (Negi, Negi, & Pandey, 2011).

The web-based lecture evaluation system is relevant to improve the quality of lecture evaluation which is so far paper-based (Hwang, 2011). This system is made to facilitate the storage of data evaluation results because the results are stored in soft files which are much easier to save and retrieve than the printed or paper files. Based on the above explanation, this research is necessary to do the design and implementation of web-based lecture evaluation system to facilitate the work in the lecture evaluation system.

METHOD

The research applied R&D which was used to develop SEVADO (*Sistem Evaluasi Perkuliahan Dosen*/Evaluation System of the Lecturer's Lectures). R&D method was a process that is utilized to validate and develop products (Sugiyono, 2015, p. 28). Its procedure model had been simplified in the Figure 1.
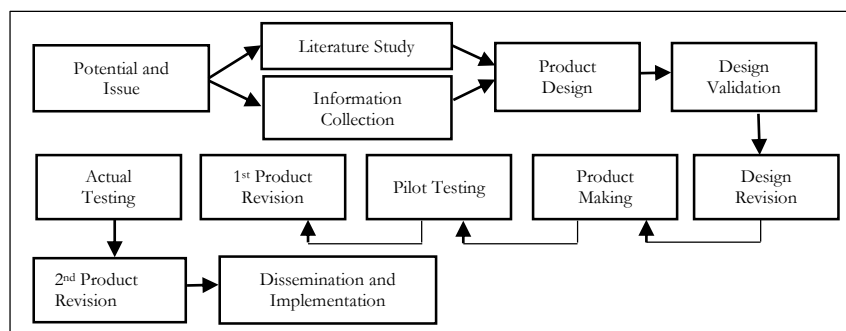

Figure 1. Research Procedure and Development of SEVADO (Sugiyono, 2015).

Product relevance became the variable in evaluation. The aspects of testing included functionality, usability, reliability, performance and supportability (FURPS). Functionality is information system capability to provide users' needs. Usability consists of human factors such as esthetic, consistency in user interface and online assistance. Reliability covers frequency and degree of failure, recovery, accuracy, prediction and the mean time between failures. Performance emphasizes on the functional required condition such as acceleration, efficiency, supply, and accuracy. Supportability relates to capability of the evaluation, development, adaptation, maintenance, and compatibility.

To test the aspect of reliability and performance, WAPT tools (Web Application Performance Testing) software was applied in this research. WAPT is a tool of pressure and stress test that allows a developer to analyze web performance. Meanwhile, to test the aspects of functionality, usability and supportability, this research utilized questionnaire to get users' perception toward those three aspects on the developed information system.

Research population in this study included people, properties, models, and characteristics of academic life in the Post-graduate Program of IAIN Salatiga. The sampling technique applied in this research was probability sampling. It means that each member of population had equal chance to be selected as the sample. Meanwhile, the type of probability sampling used was disproportionate stratified random sampling. This technique was taken when the population has non-proportional strata (Sugiyono, 2015, p. 82). Data collection in this study covered documentation, interview, FGD, and questionnaire. The quantitative data from the experiment were analyzed by using descriptive statistics. Respectively, it was converted into qualitative data on the scale of five to comprehend the product quality, as shown in Table 1 (Widoyoko, 2009, p. 238).

Table 1. The Conversion of Quantitative Data into Qualitative Ones.

Formula	Classification
$X > Xi + 1.8 \times sb_i$	Excellent
$Xi + 0.6 \times sb_i < X \leq Xi + 1.8 \times sb_i$	Good
$Xi - 0.6 \times sb_i < X \leq Xi + 0.6 \times sb_i$	Fair
$Xi - 1.8 \times sb_i < X \leq Xi + 0.6 \times sb_i$	Poor
$X \leq Xi - 1.8 \times sb_i$	Very poor

Description:

X = gain score (empirical score).

Xi = ideal mean.

= ½ (ideal highest score + ideal lowest score).

sbi = ideal standard deviation.

$= \dfrac{1}{6}$ (ideal highest score + ideal lowest score).

Ideal highest score = ∑criteria items x highest score x respondent number.

Ideal lowest score = ∑criteria items x lowest score x respondent number.

RESULTS AND DISCUSSION
Web-based Lecture Evaluation System

System is a collection of interrelated elements to achieve a certain purpose (Edström, 2008). A system describes an event or reality such as place, things, and real people. System is a collection of components that connect one to another (Indrajit, 2001, p. 2). System is a set of elements constructing collections or procedures or flowing diagrams in order to achieve individual or communal goals by operating data and/or things in a certain time to produce information and/or energy and/or things (Mulyadi, 2010, p. 5). Based on those arguments,

system can be defined as a collection of elements that operate simultaneously to complete a certain task.

Evaluation refers to a hierarchical measurement, assessment and examination. The comparison of observation with the criterion is a measurement. The description of the evidence is an assessment, and the judgment of the value of implication of the behavior is an evaluation. Based on Joint Committee on Evaluation Standard, evaluation is the systematic assessment of the worth or merit of some object (Faizah, 2014).

Evaluation is a process that is implemented to assess particular program. Evaluation is a grading process based on criteria or objective standard (Djaali & Ramly, 2000, p. 3). Evaluation is a part of efforts in improving institution quality, performance or productivity. The purpose of evaluation is to have insight and comprehend the on-going process in learning. By this process, information about what aspects have met the expectation or required consideration is gained (Mardapi, 2008, p. 19). Evaluation provides information to the class and teachers to improve teaching learning quality. Evaluation is a process of summing up the results of measurements or tests, giving them some meaning based on the value judgment (Faizah, 2014).

Evaluation is a key stage in all teaching-learning processes (Winter, Turner, Spowart, Muneer, & Kneale, 2017). Lecture evaluation is a process of populating information to assess the process and outcome of lectures in order to make policy to improve their quality. In this context, there are two types of evaluation including evaluation of the learning outcome and learning process. The evaluation of learning outcome is conducted by lecturers. On the contrary, the evaluation of learning process will be more effective if it is assessed by other people not the lecturers. Those people are the ones involved in the process, the students. It means that they are eligible to evaluate the courses taken and pay attention to the learning process. For evaluation, observers require skills to create intelligent grading considering the aspects of practitioners, program participants, and outsource professionals (Edström, 2008; Yusuf, 2017). Lecture can be evaluated from the aspect of: 1) learning design, 2) syllabus, 3) learning materials, and 4) teaching references and methods (Matiru et al., 1995, pp. 277–282). The objective of lecture evaluation is to set up the value, benefits, strength, effectiveness, impact of performance, process, and learning outcome.

In the meantime, a common element of the internet is World Wide Web (www). Technically, a web is an information system in the format of texts, graphics, voice, and others stored in server web internet using hypertext form (Janner, 2010, p. 47). In short, the definition of www is a population of interconnected multimedia documents using hypertexts link. By clicking hyperlink, users can have a leap from one document to another.

From the definition of system and evaluation above, it can be summarized that web-based lecture evaluation system is a system of data processor that is utilized to evaluate lectures as the support of managerial performance of an institution. It will ease an institution to provide relevant information by using web technology as the communication tool.

The Process of Designing the Web-Based Lecture Evaluation System
1. The stage of system design.

The current system of lecture evaluation in the Post-graduate Program in IAIN Salatiga is paper-based. It encounters issues of storage, process, and information transmission to lecturers as part of the assurance program of academic quality. Interview was conducted to lecturers and students. The results show that the implementation of lecture evaluation system during the use of paper-based lecture evaluation system should be evaluated using the computer application system. Following the interview, FGDs were conducted with information technology (IT) experts, the quality assurance unit of IAIN Salatiga, and the Head of Post-graduate Program of IAIN Salatiga to discuss the plan to build the web-based lecture evaluation system. The

developed system aims to support the evaluation of the lectures in the Post-graduate Program of IAIN Salatiga.

To develop the proposed system, this study initially conducted a need analysis based on the current lecture evaluation. From interviews and FGD, this study revealed that lecture evaluation system should be on the basis of internet. The follow up stage tried to design information system including system model, database design, and interface design.

2. The stage of validation design.

The overall design – system model, database design, and interface design – was tested by experts. The instrument for validation shown in Table 2 and Table 3. This validation aimed to find out the feasibility of the design with the particular needs. In this stage, the experts were from three fields: IT, professional practitioners, and the Quality Assurance Unit of IAIN Salatiga. The design validation by IT experts was to test whether the design runs well in several developed designs. Moreover, the design validation from the professional practitioners was to find out whether the developed designs have fulfilled the practitioners' expectations related to the need of supervision information. Data population in the validation stage occupies questionnaire.

Table 2. Software Test Instrument for Experts.

Aspect	Criteria item	Score				
		5	4	3	2	1
Functionality (F)	▪ The design of information system has provided the related information (F1). ▪ The design of information system has included data process procedure (save, edit, delete, and data interface) (F2). ▪ The design of information system has provided searching features (F3). ▪ The design of information system has provided username and password as the account security feature (F4).					
Usability (U)	▪ The design of information system is simple and easy to understand (U1). ▪ The design of information system has attractive looks on each page (U2). ▪ The layout of the information system design is orderly set up (U3).					
Reliability (R)	▪ Frequency and degree of failure. ▪ Recovery, accuracy, prediction and mean time between failures.					
Performance (P)	▪ Functional required condition such as acceleration, efficiency, supply, and accuracy.					
Supportability (S)	▪ The design of information system is possibly operated in different search engines: google chrome, mozilla firefox and internet explorer (S).					

Table 3. Software Instrument Test for User (Students).

Aspect	Criteria item	Score				
		5	4	3	2	1
Functionality	*Completeness*					
	- System information has loading data processing (save, edit, delete, and perform data).					
	- The information provided is very effective in helping to give the information I need.					
	Consistency					
	- This Information system has been consistently performing commands according to the desired buttons					
	Security					
	- The Information system has been completed with username and password as a security equipment of the					
Usability	*Accuracy*					
	- This system is able to perform data processing (save, edit, delete, perform data) appropriately.					
	Error Tolerance					
	- SEVADO system is still able to run when an error occurs, either in the login process, or data processing (save, edit, delete, and perform data).					
	Simplicity					
	- Information, menus, and buttons that exist in SEVADO system can be understood without any difficulty.					
	Operability					
	- SEVADO System can be operated easily.					
	- This system is very easy to learn.					
	View					
	- The appearance of each interface in this information system has been designed in an attractive way.					
	- The layout in the system is very well organized.					
	- The background and the text colors are clear and distinguishable.					
	- The information system uses easy-to-read font sizes.					
	- Color variations in information systems are sufficient in terms of brightness and color contrast.					
Supportability	- This system can be operated from various searching engines, such as google chrome, mozilla firefox and internet explorer.					

Alternatives: 5 - very good, 4 - good, 3 - enough, 2 - poor, 1 - very poor.

3. Determining the conversion standard of quantitative data to qualitative ones.

Prior to data analysis, a conversion standard of quantitative data to qualitative ones is required in order to measure product quality. The formula used in the conversion of scores into value was based on the former set up to define the criteria of the design feasibility. Based that calculation, the five scale data conversion can be simplified in the following Table 4.

Table 4. The Data Conversion Standard of Quantitative Data to 5 Scale Qualitative Ones.

Score	Score interval	Criteria
A	$X > 4.2$	Recommended
B	$3.4 < X \leq 4.2$	Feasible
C	$2.6 < X \leq 3.4$	Fair
D	$1.8 < X \leq 2.6$	Less feasible
E	$X \leq 1.8$	Inappropriate

4. Product making.

The next stage is actualizing the overall design and revising it into a product. The end product of supervision information model of Islamic education school is as in the followings.

a. Homepage.

All users (admin, supervisor, teacher, and guest) could access this SEVADO page. By typing the address (www.ppsiainsalatiga.ac.di/sevado) into the internet web browser, the users were automatically directed into the default address as shown in Figure 2. After inputting the address, users would be directed to homepage of SEVADO.

Figure 2. The Default Address of SEVADO.

b. Login form.

Registered users (admin, supervisors, and lecturers) could log into individual user accounts. By clicking the login button, the users would be directed to login form as depicted in the Figure 3.
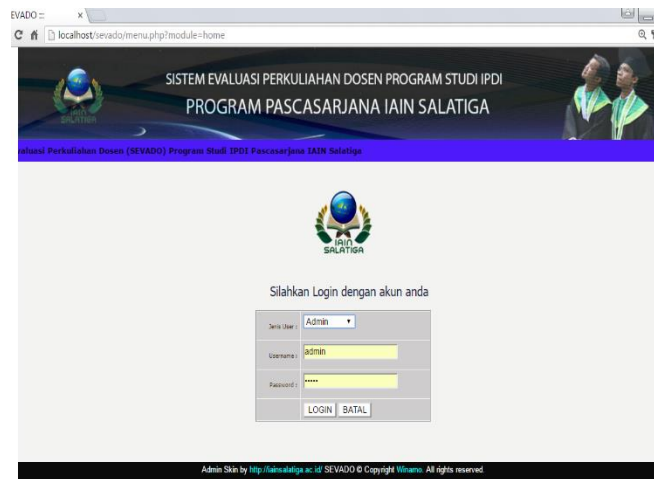

Figure 3. Login Form Page of Admin Accounts.

Figure 3 illustrates the login form and security system. In order to protect each user account, it was provided with three types of security systems: user type, username, and password. They were integrated with one another. After passing the login system with those three types of security system, users enter their personal account. If they violated the security system, they would be notified that the login failed.

c. Admin page.

An admin who passed the login system would enter the admin page as illustrated in the Figure 4.
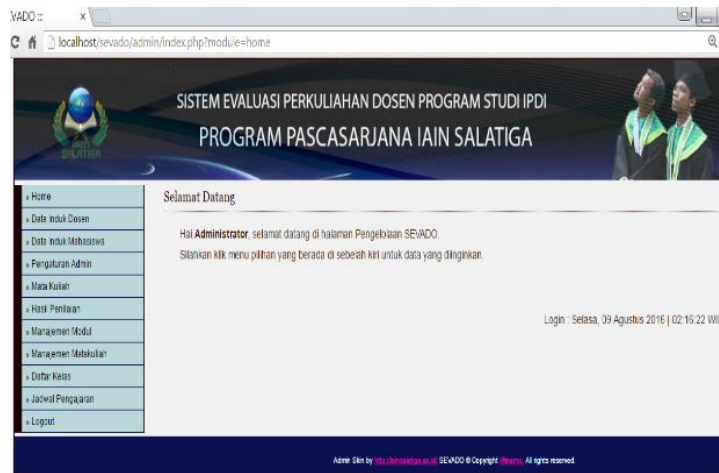

Figure 4. The Main Admin Page.

In this page, an admin was able to manage the system configuration including lecturer and student personal data, administration configuration, course subjects, learning outcomes, module configuration, course subject management, lists of classes, and class schedules.
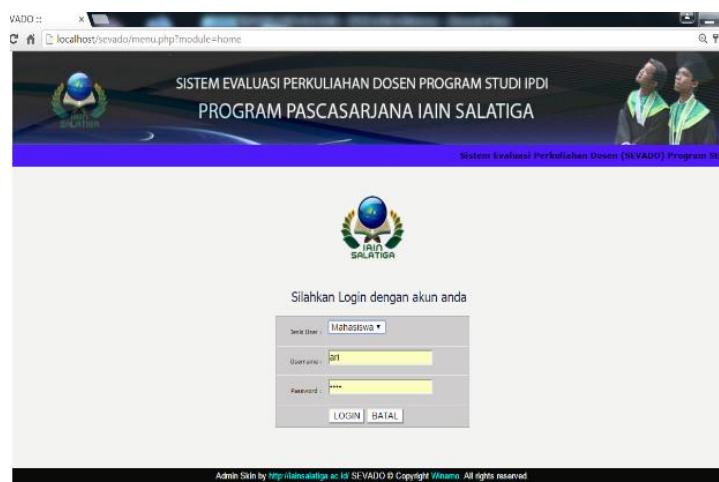
d. Student page.

The login appearance of student page is illustrated in the Figure 5.

Figure 5. The Login Form Page of Student Accounts.

After students completed the login information, they would be able to update their personal data, access scoring icons and grades as presented in the Figure 6.
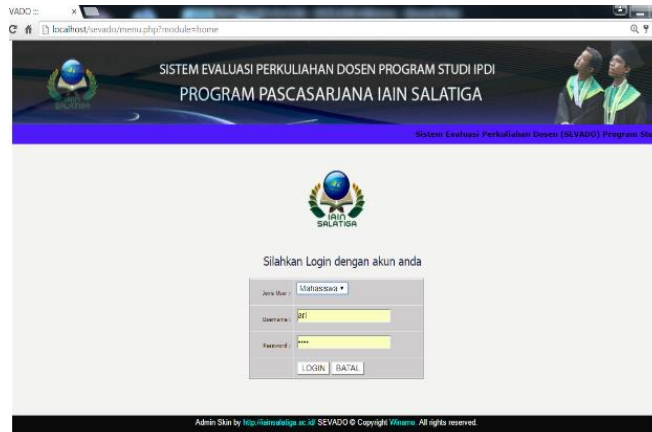
Figure 6. The Main Page of Student Accounts.

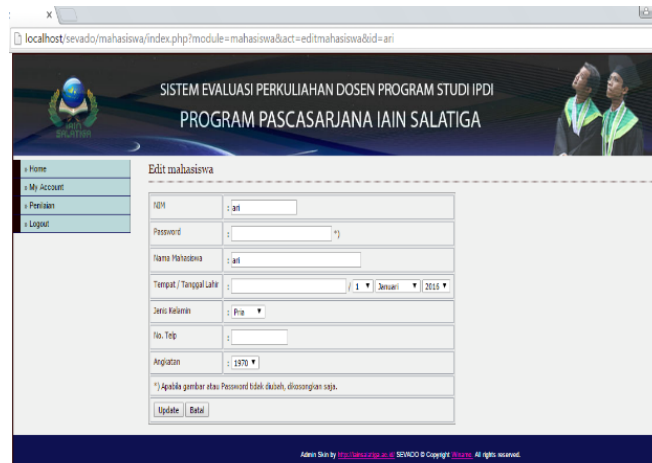To update student personal data, they have the following form to fill out.

Figure 7. The Student Account Page to Update Personal Data.

Students also had access to evaluate teaching learning process by choosing one of the subject courses as exposed in the Figure 8.
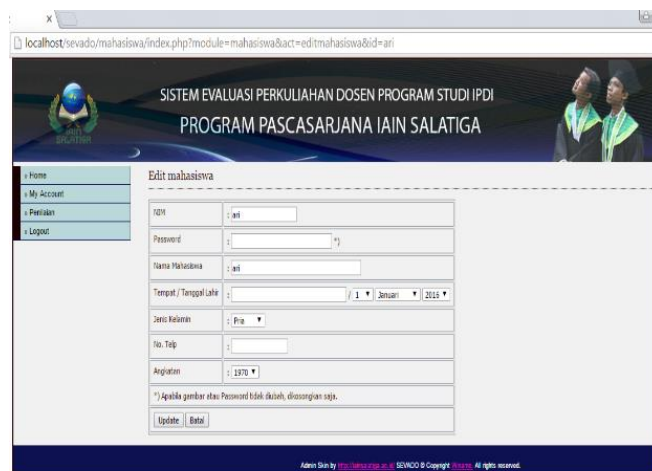

Figure 8. The Student Account Page to Evaluate a Subject Course.

In short, students would find main form of SEVADO by inputting related data to evaluate teaching and learning process of a particular subject course. The form of such evaluation is provided in the Figure 9.

Figure 9. The Student Account Page to Fill Out the Instrument of Lecture Evaluation.

5. Pilot research.

All system had been set up including user's interface, buttons, and database. Respectively, the system was tested with small scale of samples. In this pilot research, ten students took part as volunteers. The data population in this validation stage applied two different techniques: questionnaire and WAPT tools software. The software was utilized to test the aspects of reliability and performance from the developed system. On the other hand, the questionnaire aimed to test the aspect of functionality, usability and supportability. The questionnaire was in the format that had been set up in the beginning of research.

The result of this pilot research was analyzed to conclude the feasibility of the developed system. The steps of data analysis of the pilot research are outlined as follows.

a. The data analysis of the aspects of functionality, usability and supportability.

The analysis of these aspects was conducted by annotating the data, calculating the scores and mean of each aspect, and then converting them to qualitative values in accordance with the data conversion table. The result is displayed in Table 5.
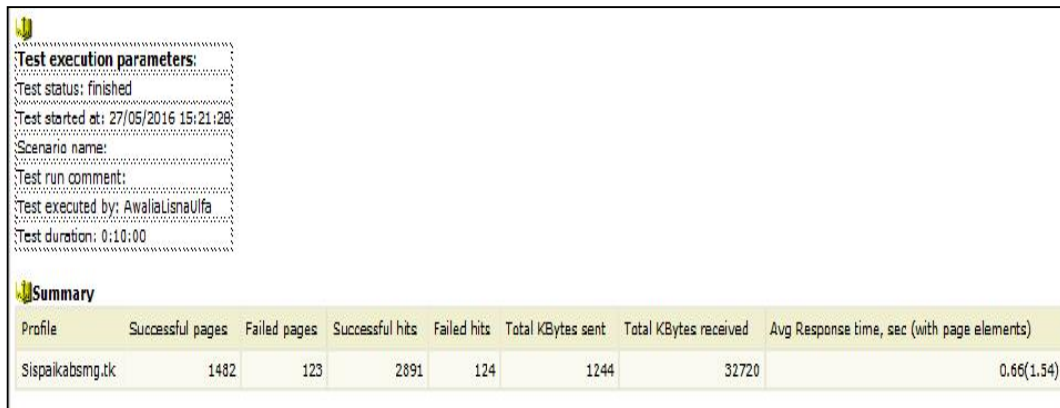
Table 5. The Analysis of Data Conversion of Each Indicator in the Pilot Research.

Item	Item score	Criteria	Aspect	Score	Criteria
F1	2.6	Less feasible			
F2	3.1	Fair	Functionality	3.4	Fair
F3	4.1	Feasible			
F4	3.7	Feasible			
U1	3.3	Fair			
U2	4.3	Recommended			
U3	4.2	Feasible			
U4	4.3	Recommended			
U5	4.1	Feasible	Usability	4.2	Feasible
U6	4.3	Recommended			
U7	4.3	Recommended			
U8	4.0	Feasible			
U9	4.2	Feasible			
U10	4.5	Recommended			
R1	4.1	Feasible	Reliability (R)	4.2	Feasible
R2	4.3	Recommended			
P1	4.2	Feasible	Performance (P)	4.2	Feasible
S	4.5	Recommended	Supportability	4.5	Recommended

Based on the above data, the criterion items of F1, F2, and U1 need revision. Simultaneously, the data was compared with the comments gathered in the pilot research. Thus, it is concluded that the aspects for revision include homepage provided with a brief data of supervisors and security system.

b. The data analysis of the aspects of reliability and performance.

To get the data of system's reliability and performance, the authors used the help of WAPT applications. This test was carried out in 10 minutes by employing 20 users to enter every 20 seconds each. The test result is obtained at the Figure 11.

Figure 11. Reliability and Performance Aspect Test Results Using WAPT Application.

In the testing of the system's performance, there were 102 successful sessions and 123 failed sessions; and then there were 1482 pages opened successfully and 123 pages failed to open. Therefore, the total number of pages to be opened was 1605 (1482 + 123). The number of successful hits was 2891 times and the failed ones were 12 hits. So, the total hits done were 2903 times (2891 + 12). Total data sent was 1244 Kbytes and 32730 Kbytes of data was successfully received. Average time needed to respond was 0.66 seconds with the number of page elements was 154.

1). The performance error is as follows:
   = (Number of failed pages: Total Pages opened) *100%
   = (123:1605) *100%
   = 7.66%
   So the percentage of system performance success is:
   = 100% - error performance
   = 100% - 7.66%
   = 92.34% = 92%

2). The percentage of error hits is as follows:
   = (Number of failed hits: Total hits done) *100%
   = (12: 5370) *100%
   = 4.41%
   So the percentage of successful hits with open pages is:
   = 100% - error hits
   = 100% - 4.41%
   = 96.69% = 97%

By applying the WAPT software, the result shows that the system performance is 92%, and system reliability is 97%. Therefore, it can be concluded that the aspects of developed system reliability and performance is outstanding.

6. Data analysis of the main research.

Having tested in the pilot research, the system was then tried out in the main research by involving 46 students.

a. The data analysis of the aspects of functionality, usability and supportability.

The analysis of the aspects of functionality, usability and supportability was conducted by annotating the data, calculating the scores and mean of each aspect, and then converting them to qualitative values in accordance with the data conversion table. Based on the data in Table 6, overall aspects are considered recommended.

Tabel 6. The Analysis of Data Conversion of Each Indicator in The Main Research.

Criteria item	Item score	Criteria	Aspect	Per-Aspect Score	Criteria
F1	4,2	Feasible			
F2	4,3	Recommended	*Functionality*	4,2	Recommended
F3	4,3	Recommended			
F4	4,2	Feasible			
U1	4,2	Recommended			
U2	4,2	Feasible			
U3	4,2	Feasible			
U4	4,2	Recommended			
U5	4,3	Recommended	*Usability*	4,3	Recommended
U6	4,4	Recommended			
U7	4,3	Recommended			
U8	4,2	Feasible			
U9	4,2	Recommended			
U10	4,3	Recommended			
S	4,3	Recommended	*Supportability*	4,3	Recommended

b. The data analysis of the aspects of reliability and performance.

By applying the software of WAPT, the result shows that the system performance was 97%, and system hits is 92%. Therefore, it can be concluded that the aspects of the developed system's reliability and performance are outstanding.

CONCLUSION

The application of Web-Based Lecture Evaluation System in the Post-graduate Program of IAIN Salatiga has met the standard criteria in the aspects of functionality, usability, reliability, performance, and supportability. It means that the system is recommended to be implemented. To conclude, it is recommended that the Web-Based Lecture Evaluation System be implemented in Islamic higher education in the purpose of improving the effectiveness of lecture evaluation system and the quality of teaching learning process in each semester. In addition to those advantages, the number of visitors to institution website will significantly improve and improve the web metrics rank.

BIBLIOGRAPHY

Arikunto, S. (2008). *Penilaian program pendidikan*. Jakarta: Bina Aksara.

Djaali, P. M., & Ramly. (2000). *Pengukuran dalam bidang pendidikan*. Jakarta: PPs UNJ.

Edström, K. (2008). Doing course evaluation as if learning matters most. *Higher Education Research & Development*, *27*(2), 95–106. https://doi.org/10.1080/07294360701805234

Faizah, A. (2014). Evaluasi kebijakan peraturan majelis Dikdasmen Muhammadiyah tentang kinerja guru di SMA Muhammadiyah 1 Gresik. *Jurnal Kebijakan Dan Pengembangan Pendidikan*, *2*(1), 93–100. Retrieved from http://ejournal.umm.ac.id/index.php/jmkpp

/article/view/1904/2009

Henderson, M., Selwyn, N., & Aston, R. (2015). What works and why ? Student perceptions of 'useful' digital technology in university teaching and learning. *Studies in Higher Education*, *42*(8), 1–12. https://doi.org/10.1080/03075079.2015.1007946

Hwang, H. J. (2011). Design of an effective learning evaluation component in web-based instruction. *International Journal of Multimedia and Ubiquitous Engineering*, *6*(4), 1–12. Retrieved from http://www.sersc.org/journals/IJMUE/vol6_no4_2011/1.pdf

Imai, Higashikakiuchi, K., Nakayama, A., Kagawa, K., Ando, K., & Hattori, T. (2017). Development and evaluation of web-based computer simulator for e-learning of computer literacy. *IEEJ Transactions on Electronics, Information and Systems*, *137*(12), 1633–1640. doi: 10.1541/ieejeiss.137.1633

Indrajit. (2001). *Analisis dan perancangan sistem berorientasi object*. Bandung: Informatika.

Jabbarifar, T. (2009). The importance of classroom assesment and evvaluation in educational system. In *Proceedings of the 2nd International Conference of Teaching and Learning (ICTL 2009)* (pp. 1–9). Putra Nilai: INTI University College. Retrieved from http://ictl.intimal.edu.my/ictl2009/proceedings/index.html

Janner, S. (2010). *Rekayasa perangkat lunak*. Yogyakarta: Penerbit Andi.

Koehler, N., Correia, A.-P., Alpay, N., & LeVally, C. (2017). Formative Evaluation of a Web-based Multimedia Intervention to Support Learning of Statistics. In *Proceedings of the 9th International Conference on Computer Supported Education* (pp. 92–99). Retrieved from http://www.scitepress.org/digitallibrary/PublicationsDetail.aspx?ID=klJtpkHqTD0=&t=1

Mardapi, D. (2008). *Teknik penyusunan instrumen tes dan nontes*. Yogyakarta: Cendikia Press.

Marmah, A. A. (2014). Students ' perception about the lecture as a method of teaching in tertiary institutions . views of students from college of technology eduction , Kumasi (COLTEK). *International Journal of Education and Research*, *2*(6), 601–612. Retrieved from http://www.ijern.com/journal/June-2014/50.pdf

Mas'udi. (2014). Evaluasi pembelajaran (menelaah landasan filosofis evaluasi pembelajaran dalam perspektif pengembangan keilmuan di STAIN Kudus). *ThufuLA*, *2*(2), 317–331.

Matiru, B., Mwangi, A., & Schlette, R. (1995). *Teach your best: a handbook for university lecturers*. Bonn: DSE-Germany.

Mulyadi. (2010). *Kepemimpinan kepala madrasah dalam mengembangkan budaya mutu: studi multi kasus di madrasah terpadu MAN 3 Malang, MAN I, dan MA Hidayatul Mubtadi'in kota Malang*. Jakarta: Badan Litbang dan Diklat, Kementerian Agama RI.

Negi, P. S., Negi, V., & Pandey, A. C. (2011). Impact of information technology on learning, teaching and human resource management in educational sector. *International Journal of Computer Science and Telecommunications*, *2*(4). Retrieved from www.ijcst.org/Volume2/Issue4/p13_2_4.pdf

Nuriyah, N. (2014). Evaluasi pembelajaran: sebuah kajian teori. *Jurnal Edueksos*, *III*(1), 73–86. Retrieved from http://www.syekhnurjati.ac.id/jurnal/index.php/edueksos/article/view%0A/327

Sugiyono, P. (2015). *Metode Penelitian Kombinasi (Mixed Methods)*. Bandung: Alfabeta.

Suherman, E. (2003). *Evaluasi pembelajaran matematika*. Bandung: JICA, UPI.

Widoyoko, E. P. (2009). *Evaluasi program pembelajaran*. Yogyakarta: Pustaka Pelajar.

Winter, J., Turner, R., Spowart, L., Muneer, R., & Kneale, P. (2017). Evaluating academic development in the higher education sector: academic developers' reflections on using a toolkit resource. *Higher Education Research & Development*, *36*(7), 1503–1514. https://doi.org/10.1080/07294360.2017.1325351

Yusuf, W. B. (2017). *Evaluasi program pembelajaran dalam meningkatkan mutu pendidikan*. Salatiga. Retrieved from https://www.researchgate.net/profile/Wirastiani_Yusuf/ publication/ 316281230_Evaluasi_Program_Pembelajaran_dalam_Meningkatkan_Mutu_Pendidikan/l inks/58f92e9f0f7e Pendidikan.pdf?inViewer=1&pdf