Web-based knowledge assessment platform with eye tracking

Ivana Tomić
Computing and automation
Faculty of Technical Sciences, University of Novi Sad
ivana1996tomic@gmail.com

***Abstract*— This paper presents a software system for monitoring and analyzing eye movements during the creation of an online test. This software system enables the creation of tests with the marking of regions of interest, the creation of a test with tracking of eye movements, the collection of recorded data and their processing and analysis. The test creator can view the collected data in two ways: as raw, tabular data and visually, as a flow graph of views through regions of interest. In this software system, during testing, the coordinates are recorded in the database using the Eyeloop software. This software uses the computer's camera. After collection, the data is processed, that is, the coordinates belonging to the regions of interest are extracted. The processed data is stored in the database. This paper presents the basic solution and analysis, as well as some possibilities for expansion and improvement.**

***Keywords*— modern educational technologies and standards, tracking eye movements, online testing, data analysis and visualization**

I. Introduction

Advances in technology bring relief in almost all fields. As a consequence, most traditional processes have been replaced in part or in whole by processes performed by computers. This paper describes a software system that provides the application of technology in the process of testing respondents. The complete replacement of the traditional way of assessment (paper-based test work) with the use of computers for the testing process has brought relief to both examiners and examinees. In addition to the automation of evaluation (if questions with offered answers are used), simpler storage of tests (without the need to print), easier sharing of tests and display of results, etc., opportunities are also provided for the collection of a larger number of data, and therefore their analysis.

With traditional assessment, from the data, only answers that are part of the respondent are collected, on the basis of which the assessment is made. Using a computer for testing purposes in combination with various devices and software leads to a much larger amount of information. The obtained data are available for further processing and analysis, on the basis of which a large amount of information can be obtained. During testing on the computer, we can see how long the student stayed on which question, whether he went back to a previous question and changed the answers, etc. With the use of eye tracking devices, this range of information is even wider. Space is also opened for insight into how long the respondent looked at the question itself, then how long he read the answers, whether he stopped when he read the correct answer or continued on and read everything, etc. This paper describes the software system that deals with the collection and analysis of these parameters. The motivation for implementing such a system is precisely in the possibilities of processing and visualizing the collected data. Through adequate processing and analysis of the obtained data, the examiner can obtain information about the quality of his test, but also a large amount of information about each individual question. For example, it can tell how clear the questions and the answers were to the respondents by whether, say, most students read the text of the question or the answers several times. Also, it is possible to evaluate whether students react better when they have offered answers or essay questions, etc.

Gaze tracking provides a highly accurate record of the subject's visual behavior. The data obtained in this way can be used in other areas as well, which is another motive for implementing such a system.

The information obtained can be used to tailor tests to students. In addition, the information can be used to adjust the course and the way the material is presented, because thanks to eye tracking, we have an insight into what attracts the respondents' attention.

After the introduction, in the second section, an overview of related research is given. Then, in the third section, an explanation of the methodology used in this research is provided. The next section provides concluding remarks and directions for future research.

II. Related research

In the following, works dealing with online assessment, as well as assessment using eye tracking devices, will be described.

Paper [2] describes an online rating system developed to collect, process and return ratings to users. It discusses the design requirements, features and implementation of online assessment systems, as well as user (rater and student) reactions. The course called 'Introduction to Design and Communication' is specifically described. Students are divided into teams of four to six people. Each team works on a different project. They have a course page where they can receive general information, download assignments and ask questions in public forums.

In addition, each section has its own page where teams submit assignments and where they also have specific information available. Finally, there is a page where the courses are combined, which represents the faculty page. Depending on the role, the participants in the system have different needs, but the common need is for safe and reliable information. The HTTP protocol was used for communication between the centralized server and the client machine. The PHP programming language was used to implement the system. MySQL is used for data storage. The main disadvantage of this system is that it is not compatible with a mobile phone. It was concluded that even small details can hinder the usability and thus the satisfaction of the users of such systems.

This paper covers the very important segments that an online rating system should have. The method of implementation and use of the system itself is described in detail.

The research [3] describes software aimed at developing and implementing an adaptive learning and assessment system in order to increase the efficiency and quality of feedback. An automated evaluation system was developed. One of its advantages is that it can be used repeatedly because it remembers which answers were marked as correct. The process used for grading in this system is based on results. This means that the system evaluates the actual answers students give, not the process or mouse clicks to achieve them.

This system contains two lists. One of them is a list with correct answers, and the other is a list with incorrect answers and a feedback message for the student. It consists of a correct answer section, a wrong answer section, and a custom feedback component. The paper describes one of the ideas of implementing a system that automatically recognizes correct answers.

The grading process begins when the student's answer is compared to the list of correct answers. If no match is found, the item is compared to the list of incorrect answers. If the answer is not previously marked as true or false, the instructor is asked to identify the student's answer. The evaluator determines how many points each answer carries. This system also has a plagiarism detection plugin, which is another of its advantages. It was concluded that the system has a positive effect on the amount of feedback and reduces the time required for grading tasks. This confirms the effectiveness of such systems. As no impact on the quality of the feedback was shown, it is listed as a subject for further study.

Research [4] uses eye tracking to examine students' visual attention when solving multiple choice problems. Six students participated in the research, whose task was to solve the problem of predicting the occurrence of landslide hazards. They solved the problem using four images representing four combinations of four factors (temperature, rain, slope and cracks).

Specifically, this study investigated how students spend time reviewing the options offered as a solution. Participants' responses and visual attention were recorded using an eye tracker. The participants were also asked to think aloud while solving the task. It was concluded that there is a significant difference in the visual attention of students who successfully and unsuccessfully solved the problem. Successful problem solvers tend to shift their visual attention from irrelevant factors to relevant factors, while unsuccessful problem solvers do the opposite. It is evident that successful problem solvers can recognize and concentrate on relevant learning cues that lead to problem solving.

As one of the future researches, the examination of the influence of the student's prior knowledge on the allocation of attention during solving a scientific problem with multiple choices is indicated. The eye tracker used in this study has a sampling frequency of 60 Hz.

Thanks to this work, we get another of the possible primordial systems for grading with gaze tracking. Comparing views among students who successfully and unsuccessfully answered the questions expands the aforementioned set of possibilities that we get from data processing, visualization and analysis.

A study [5] investigated eye-tracking ability to determine differences in problem-solving behavior in biology, chemistry, and physics. Six participants with different levels of expertise and knowledge from the above fields participated. The recorded eye movement information includes the location of the gaze fixation on the screen, the duration of the fixation, the path of the eye movements (saccades), and the duration between fixations. Each participant has 18 multiple-choice questions, i.e. 6 questions from each area.

In this study, fixation duration was considered as the total time within each scientific topic. The research did not produce clear relationships between an individual's area of expertise and total time spent within each scientific topic. A large degree of variation was found among participants in terms of the time they spent on each topic and within each science topic. Comparison of fixation duration data did not produce clear and consistent differences corresponding to known levels of expertise between participants. In this research, zones, that is, areas of interest with a rectangular shape, were also set.

Previous research, reported in this paper, suggests that individuals who score the same on a test may differ in their subject knowledge or level of expertise. As a topic of future research, the possibility of examining the degree to which these protocols can be applied as an aid in the detection of ineffective educational techniques and intervention in the teaching process in order to better adapt the teaching methodology to the needs of students is mentioned.

This system, similar to our system, has areas of interest. In it, the gaze fixation time is taken as additional relevant data, which did not bring clear differences between students and fields.

In the study [6] eye movement tracking was used to record the student's visual attention while solving a test on understanding graphs in kinematics. 115 students participated in this research. The aim of the study was to investigate how a student's understanding of kinematic graphs affects directing attention to questions and to different answer options. Eye movements were recorded using a Tobii X3-120 gaze tracking system.

It is an interesting conclusion that students spent less time on the question when they answered correctly compared to students who answered the same question incorrectly, which was linked to self-confidence. There was no correlation between time spent on the test and student scores, overall.

This paper, in addition to the already mentioned comparison of students who answered correctly and incorrectly, also compares the time taken to create the test itself and to stay on each individual question.

III. Methodology

In this section, a description of the methodology of the software system follows. The implementation of the application will be described in turn, followed by the method of processing and visualization of the collected data.

The user application can be divided into a part related to the creation of the test itself, then a part for creating the test and finally for the obtained results. Testers are allowed to create questions first. There is a possibility of creating a question with suggested answers or with a text field, where respondents are supposed to enter their answer. The picture (picture 1) shows the form for entering questions.

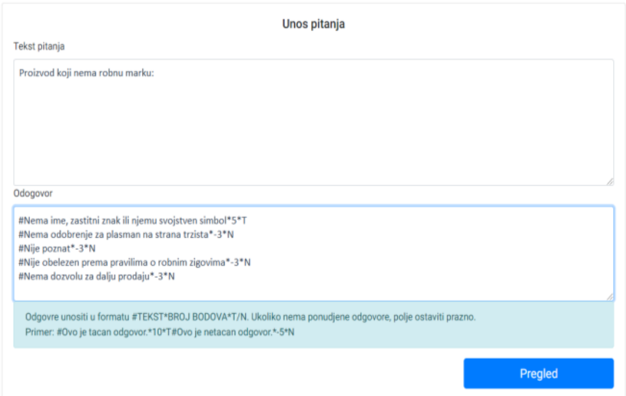

Figure 1. Question input form

In the next step, the examiner who created the question is allowed to mark in it the regions that will be recorded as regions of interest. The number of regions is not limited. Figure 2 shows the entry of the region of interest. The window is displayed in the same way as it will be seen by the examinee during the preparation of the test. After labeling, the regions of interest are stored in the database, individually for each question.

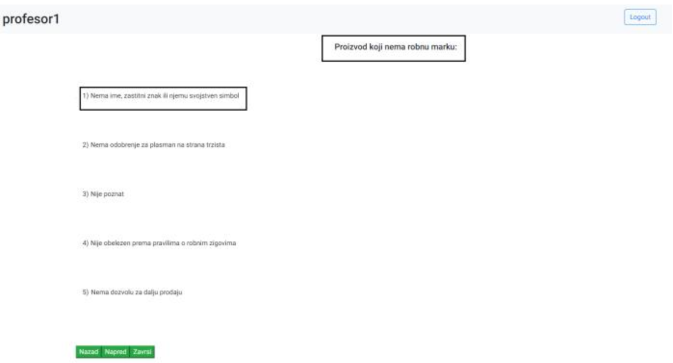

Figure 2. Entering the region of interest

After that, the examiner can create his own test by selecting the questions he has previously created. It is possible to change the order of the questions. Figure 3 shows these two steps. If the examiner has no created questions, instead of the question selection page, he will be redirected to the page for adding questions. The system was created so that it is possible to use one question in only one test.

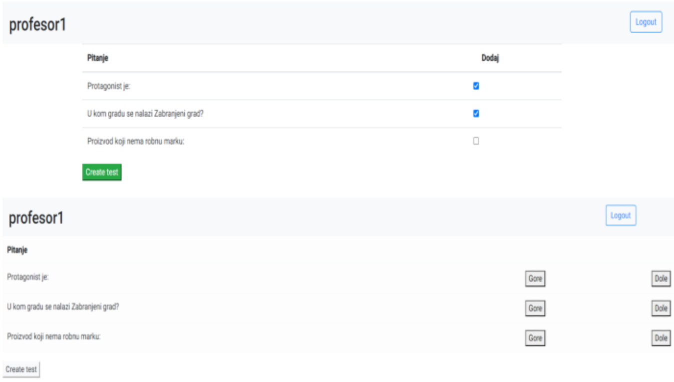

Figure 3. Selection of questions and their sequence in the test

The examiner can send respondents a link that will lead them to the test they want to test them with. The moment the examinee enters the link, the time of the start of the test is recorded. The respondent has only one question displayed on the screen at a time. He can move back and forth and change his answers (Figure 4). He can end his testing on any question.

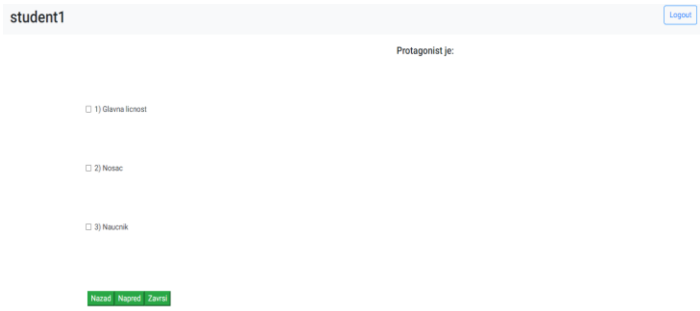

Figure 4. Presentation of the question from the point of view of the respondents

After completing the test, the examiner can see all solved tests on his profile (picture 5). Points are automatically calculated for questions with suggested answers, while for essay questions the number of points is entered by the examiner.

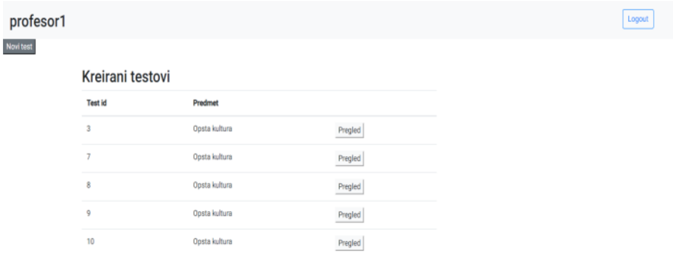

Figure 5. Examiner profile

Clicking on the "View" button (picture 5) opens a new window for the examiner (picture 6). There, the examiner is shown the name of the student who took the test, as well as options to open the questions and answers of the examinee by clicking the "View" button. On this page, it is possible to enter the number of points obtained for questions that had essay answers. By clicking on the "Download file" button, the examiner can download the raw eye tracking data recorded during the test. If the examiner chooses the "Analysis" button, he will be redirected to a page where the question will be displayed with a graph of the flow of views through the regions of interest.

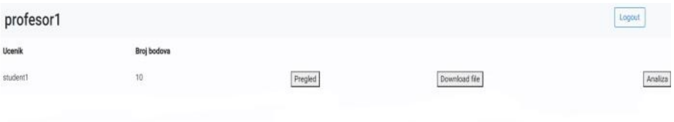

Figure 6. Overview of solved tests

On their profile, respondents can see all the tests they have taken up to that point, as well as the number of points achieved (picture 7). That number of points is updated when the examiner enters the points for the essay questions, if there were any in the test.

Figure 7. Profile of respondents

The application was created using*Java Spring*- acombined with Angular. A MySQL database was used for data storage.

The software used for eye tracking is*Eyeloop*[1]. This software is open-source and written in the Python programming language. It works with the calculator's camera. It can be used on all platforms. There is a possibility to change the parameters depending on the user's needs. This software detects the pupils and then tracks their movements and records the coordinates. The disadvantage of such software is that their accuracy is proportional to the quality of the camera.

The software used to track eye movements records the collected data in*Excel*the file. We get the time and coordinates from the data. At the end of the test, after saving the solved test, the system loads the collected data from the file and saves them in the database. The raw data is compared to labeled regions of interest, which are recorded for each question individually. The system stores data on the beginning and end of solving each question. If the respondent returns to one of the previous questions, that time interval is also recorded. Thanks to the time data recorded by the gaze tracking software and the recorded time of solving the question, we can connect the time moments and through them the coordinates to the regions of interest of the desired question.

Next to each solved test, the examiner has the option to download the raw data into*Excel*file or to show him a view flow graph through the regions of interest. The graph is displayed on a page identical to the one the respondent saw during the test (Figure 8). The examiner can move back and forth through the questions and will be shown a graph on each. If the system did not recognize that the respondent was looking at one of the regions of interest, a notification that there is no data is displayed.

Figure 8. View flow graph through regions of interest

For better visibility, each change of the region of interest to the next region of interest implies a change of color. Changes are shown chronologically.

IV. CONCLUSION

This paper presents a solution that enables online testing while monitoring the examinee's gaze flow. The system allows examiners to create a test and mark regions of interest on questions. During the preparation of the test, the coordinates of the subject's gaze are recorded. Finally, the examiner can download the raw data or see it displayed as a view flow graph (Figure 8).

Using eye tracking devices during testing can improve the quality of tests. Creating software that enables the use of eye-tracking devices can be useful for research in other fields as well. This paper presents one of the applications of data obtained by eye tracking.

For the needs of this solution, an application was created that enabled the creation and testing, as well as the collection, processing and analysis of data. During processing, it was necessary to map the corresponding coordinates recorded by the gaze tracking software with the regions of interest.

In future research, data should be collected using gaze-tracking devices, which would provide more reliable data. Visualization is of great importance in systems like this. In future work, the data should be presented in several ways, e.g. through heat maps.

LITERATURE

[1] Eyeloop,https://github.com/simonarvin/eyeloop[Accessed 8 July 2021].

[2] Thompson, Mary Kathryn; Ahn, Beunguk, "The Development of an Online Grading System for Distributed Grading in a Large First Year Project-Based Design Course", Proceedings of the 119th ASEE Annual Conference & Exposition, 2012.

[3] Kevin Matthews, Thomas Janicki, Ling He, Laurie Patterson, "Implementation of an Automated Grading System with an Adaptive Learning Component to Affect Student Feedback and Response Time", Journal of Information Systems Education, Vol. 23(1)

[4] Meng-Jung Tsaia, Huei-Tse Houb, Meng-Lung Laic, Wan-Yi Liua, Fang-Ying Yang, "Visual attention for solving multiple-choice science problem: An eye-tracking analysis", Elsevier, 2011.

[5] Robert H. Tai, John F. Loehrb and Frederick J. Brigham, "An exploration of the use of eye-gaze tracking to study problem-solving on standardized science assessments", International Journal of Research & Method in EducationVol. 29, No. 2, October 2006.

[6] P. Klein, A. Lichtenberger, S. K¨uchemann, S. Becker, M. Kekule, J. Viiri, C. Baadte, A. Waterlaus and J. Kuhn, "Visual attention while solving the test of understanding graphs in kinematics: Aneye-tracking analysis", Eye Tracking in Physics Education Research, August 2019