Baklavas, G., Economides, A.A. & Roumeliotis, M.: **Evaluation and comparison of Web-based testing tools.** *Proceedings WebNet-99, World Conference on WWW and Internet*, pp. 81-86, AACE 1999.

Evaluation And Comparison Of Web-Based Testing Tools

Grigoris Baklavas, Anastasios A. Economides and Manos Roumeliotis
Applied Informatics Department, University of Macedonia
Thessaloniki, 54006, Greece
E-mail: grigoris@uom.gr, economid@uom.gr, manos@uom.gr

Abstract: In this paper, we evaluate popular Web-based tools for test design, implementation, delivery, automatic grading and result analysis. We compare these tools with respect to the variety of question types that support, the capabilities for multimedia use, the security, the easiness of development, maintenance and delivery of tests, the automatic grading and the statistical analysis of the results. Based on these criteria and practical experience with the tools, we concluded that the best ones were Cyber Exam and QuestionMark Perception.

1. Introduction

The growth of the Internet in the last decade has given rise to the demand for virtual classrooms and distant learning. The concept behind this was that, through the Intranet of a University or even the Internet, a group of people consisting of a teacher and his students can 'meet' and perform a lecture via their computers. Recently, significant progress has been made towards this direction. The integrated support of text, graphics, audio and video by the Web makes it an appropriate vehicle for learning and knowledge acquisition. The student is actively involved in the learning process using interactive tools and collaborating with the teachers and other students [Hall 97, Brooks 97, Khan 97, Economides 99, Economides 97, Mamoukaris 99].
An important parameter regarding distant learning is the taking of the examinations of a course via a computer network. To facilitate the teachers in authoring, delivering, grading and analyzing the exam tests, special software tools had to be designed. Recently many tools that facilitate the exam taking process have appeared. These tools can be installed in a server and then used by teachers and students, replacing the classical form of written examinations or helping the self-evaluation of the student progress.
Although all of them share a basic template, they have major differences in their characteristics. In this paper, we evaluate the most popular Web-based testing tools, which are:

- Question Mark Perception of Question Mark Corporation [QuestionMark Perception]
- Hot potatoes of Half-baked Software [Hot potatoes]
- Test Maestro and School Maestro Internet Publisher of Russ & Ryan EdWare [Test Maestro]
- Examiner and FastTEST of Assessment Systems Corporation [Examiner]
- LXR * TEST of Logic Extension Resources [LXR*TEST]
- WebCT of World Wide Web Course Tools [WebCT]
- Cyber Exam of Virtual Learning Technologies [Cyber Exam]
- C-Quest of Cogent Computing Corporation [C-Quest]

We compare these tools with respect to the following criteria:

- Variety of question types
- Multimedia use
- Security
- Easy development, maintenance and delivery of tests
- Grading
- Statistical analysis of results

In the following section, we briefly present each one of these software tools. In section 3, we compare them with respect to the most important criteria. Finally, in section 4, we conclude and suggest directions for improvement of these tools.

2. Testing Tools Presentation

In this section, we present the evaluated testing tools.

2.1 QuestionMark Perception:

This is the best product of QuestionMark Corporation, a company that specializes in the design of testing tools. It helps authors to create questions easily with no need for HTML knowledge. The author follows on-screen instructions or uses a question wizard for even easier question generation. The question types supported are: Multiple choice, Multiple response, Explanation, Numeric, Selection, Text Match. Every question or test can have a time limit. Multimedia use is quite easy, and therefore any type of file (graphics, pictures, diagrams) can be inserted in a particular question. These files should be in .gif or .jpg format. Students use unique usernames and (if needed) passwords to access the test.
Perception generates tables of the grades achieved by the students that took the test, with the percentages of their right or wrong answers. This feature is remarkably useful as it allows the instructor to justify the correctness of a test or even a particular question of the test.
A disadvantage of Perception is the lack of support for Unix, requiring Windows NT environment only.
Generally, Perception is one of the top testing tools, and scores high in all the major criteria. It combines easiness, security, variety and statistical analysis in a very satisfactory degree.

2.2 Hot Potatoes:

This program is a collection of six utilities, each generating a different type of question, namely JBC for multiple choice, JCloze for filling blanks in a text, JQuiz for text insertion, JCross for crossword generation, JMix for ordering mixed words of a phrase and JMatch for text match questions. All questions can be time-limited. Hot Potatoes constructs the question Web pages automatically, immediately after the instructor inserts the questions and answers. These pages can then be stored in any Web server in order to be used by the students.
The major disadvantages of Hot Potatoes stand in the statistical analysis and in the security of the tests. It does not provide any kind of access control. Hot Potatoes can be used in unofficial testing and student self-evaluation.

2.3 Test Maestro & School Maestro Internet Publisher:

Test Maestro and its complementary program have a major difference to the other packages. They generate written papers for exams and not Internet-based tests. Questions supported by Test Maestro are: True/False, Fill in Blanks, Multiple Choice, Short Answer and Text Matching.
It is fully compatible with MS Office programs, and therefore Excel plots, Word documents or even mathematical equations can be easily inserted in a question. The generation of a test is password protected and therefore access to the question bank is protected.
Statistical analysis as well as automatic grading is totally absent from Test Maestro. These functions are left for the instructor to perform, since the exams generated are in written form.

2.4 Examiner:

Examiner is generating tests from a bank of previously developed questions. It can provide instant feedback and explanations to the student after he answers a question. Examiner is very efficient in the use of multimedia, as video or audio files can be inserted in a question. So, the instructor can develop questions about the recognition of a sound or a word.
However, the instructor can only construct Multiple Choice questions, which may have a time limit. Furthermore, the program can show the result to the student, if required, or present a detailed analysis to the instructor. Then, it is possible to store the results in a database and produce critical statistical values as mean, median and standard deviation.

2.5 FastTEST:

The main feature of FastTEST is the development of a question bank from which the instructor can pick the ones needed for his test. The program supports word processor features (Bold, Italic, Underlined, Colored Text, variable font size, left, right, centered alignment) and is fully compatible with MS Office. Its questions can be Multiple Choice, True/False, and Open Answer.
Although FastTEST is quite efficient in the statistical analysis of the test grades, it does not produce any results automatically. It only provides windows for the instructor to insert the numbers of his personal analysis of the results. Access control is accomplished by passwords on the instructor end, but not on the student end.
The main disadvantage of the program is the limited use of multimedia. Only pictures of type .bmp and .wmf can be attached to a question.
FastTEST is designed mainly for generation of papers for written exams. Its supplement, FastTEST On-Line, can publish the tests on a server, and have them time-limited, but it can only record the answers without doing any extra analysis on them. Therefore, it is in a significantly lower level than the other competitors.

2.6 LXR*TEST:

This program is designed for both written and computer-based examinations and can be definitely regarded as one of the best programs of its kind as it is very efficient in almost all the aspects of interest. The question types supported are: Multiple Choice, True/False, Text Matching, Numeric and Open Answer. Time-limited questions are also allowed. A major advantage of LXR*TEST is the multimedia use. Apart from the typical file types that are supported, a user can attach QuickTime movies in a question, a feature only present in this program.
The instructor is again able to develop a test using a previously built question bank. The individual questions are set up in a word processor-like environment with all the advantages that such a function provides. It can provide special grading for each answer so that the instructor can give higher marks in some questions. Standard security is provided with passwords not only for the instructor but also for every individual student.
The LXR*TEST is one of the best in the statistical analysis of the results. The program can record and store all the answers in a database and automatically create a table with the grades of all participating students. Then, the instructor can inspect the answers in every question of a particular student. Also, LXR*TEST generates a printable list of the grades of all students that can be attached to a notice board.

2.7 WebCT:

WebCT is not just a test authoring tool, but a teaching environment with functions like storing of lecture notes, support of mailing list for students, discussion area, chat area, glossary, index, syllabus, timetable, on-line exams and much more. Specifically, the exam generation section of the program is quite simple and the instructor can easily develop tests. It supports the following question types: Multiple Choice, Text Matching, Short Answer and Paragraph, with a choice of time limits. WebCT's special feature is the ability of giving negative values to special answers in order for the students to loose points when selecting that particular answer. Multimedia use is supported very efficiently. Access control is achieved by unique usernames and passwords that are needed not only to access the exams but also to log on to the server where WebCT is stored. The student answers and grades are stored in a database for further analysis by the instructor.
Generally, WebCT provides a very efficient teaching environment but it is not specialized in testing.

2.8 Cyber Exam:

The Cyber Exam provides all characteristics of a testing tool at very high quality. Its working environment is a web browser like Netscape Navigator and MS Internet Explorer. It supports many question types: Multiple Choice, Multiple Response, True/False, Short Answer, Fill in Blanks, Text Matching and Essay. Time-limited questions are also present. Multimedia can be used very efficiently. Apart from the usual type of pictures, audio and video, JAVA applets are supported and can be easily inserted in a question. Security measures are standard, providing password protection in both the instructor and the student ends. It provides a very high quality of statistical analysis. Immediately after a student completes a test, the program automatically grades it and is able to present the result, if wanted, to him. Cyber Exam also produces statistical reports, such as higher/lower grade, mean and standard deviation of all the grades. The results can be easily imported in a spreadsheet package like Excel or even a statistical package like SPSS. Cyber Exam can be easily regarded as one of the top programs in the area. It achieves a high level of functionality and quality for all criteria.

2.9 C-Quest:

This program is a collection of tools helping test authoring and development. These tools are: C-Quest db for the question database construction, C-Quest Test for the written examinations generation, C-Quest Web for computer-based exams, and C-quest Echo for exams via e-mail.

The questions types supported are: Multiple Choice, True/False and Multiple Response. Time limits are allowed for every type. Many multimedia file types can be inserted in the questions. These can be pictures (.gif or .jpg), sound (.wav or .mid) and video (.avi). Access control is password protected with the instructor giving the usernames of the students. After the completion of the test by a student, the results are transported to the C-Quest Web Administrator. There, the generated tables can be sorted, printed or even be exported to statistical and spreadsheet packages.

Generally, C-Quest is a nice test development program, but there is nothing special to it in order to be considered as one of the best. In all aspects, it scores about average.

3. Testing Tools Comparison

In this paper, we compare Web-based testing tools with respect to some criteria. Our experience has shown that they have major differences in almost every aspect we examined. On every criterion, there were tools that scored very high, while other performed below average. In this section, there is a complete description and scoring of the testing tools. This description is made separately for each criterion and is followed by a table that gathers the given information.

Regarding the criterion of the different Question Types supported by the tools, we remark that the standard number of types is five. However, Hot Potatoes and Cyber Exam support more types. Specifically, Hot Potatoes is the only program that can generate crosswords and sorting out a phrase from its words. Cyber Exam, apart from six basic question types, supports essays that have to be e-mailed to the professor for manual grading. Therefore, Cyber Exam has to be considered as top in this criterion. Examiner, FastTEST and C-Quest score the lowest.

The criterion of Multimedia Use can be a very decisive factor as there are major differences between the software tools in this area. LXR*TEST, WebCT and Cyber Exam can be considered as being the top under this criterion. In these packages, almost every type of multimedia file is supported. Test Maestro and FastTEST, can be considered unsatisfactory.

The criterion of Security is a very important issue for a test delivery tool, since the exams have to be accessed by a very special group of people, the instructor and the students of a particular course. Most of the tools set up their own security parameters. However, Hot Potatoes does not consider the security issue and leave it to the instructor to decide what to do. Therefore, it is considered of a lower standard than its counterparts.

Simplicity of test authoring and test taking is another criterion we examine. QuestionMark Perception, WebCT and Cyber Exam provide very user-friendly environment for test generation using wizards or web browser interfaces.

Automatic Grading with special values for particular answers is a criterion supported by half of these tools. Perception, Examiner, LXR*TEST, WebCT, Cyber Exam and C-Quest can be considered as top since different scores can be assigned to different questions.

Finally, Statistical Analysis is a criterion for which there are great differences between the evaluated tools. Some of them, like Hot Potatoes and Test Maestro have no or very limited support of statistics of the results. FastTEST, WebCT and C-Quest perform standard statistical analysis. Perception, Examiner, LXR*TEST and Cyber Exam are superior to the others providing advanced and detailed statistical analysis.

After extensive investigation and experimentation with these testing tools, we came out with the following comparison table. In this table, we score the performance of these tools with respect to important criteria.

	Q.M	H.P.	T.M.	EX.	F.T.	LXR	Web	C.E.	C-Q.
Question Types	5	6	5	1	3	5	5	7	3
Multimedia Use	B	B	-	B	C	A	A	A	B
Security	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Simplicity	A	B	B	B	B	B	A	A	B
Automatic Grading	Yes	No	No	Yes	No	Yes	Yes	Yes	Yes
Statistical Analysis	A	-	-	A	B	A	B	A	B
Cost in USD	neg.	Free	35	875	199 OL 150	1199	neg.	neg.	db: 160 Test: 152 Web: 602

The shortcuts refer to the names of the testing tools and are the following:

Q.M. QuestionMark Perception
H.P. Hot Potatoes
T.M. Test Maestro
EX. Examiner
F.T. FastTEST
LXR LXR*Test
Web WebCT
C.E Cyber Exam
C-Q. C-Quest

We also use the following scores:
A= Excellent
B= Very Good
C= Good

The sign (neg.) means that the price is negotiable between the company and the buyer depending on the number of instructors/students that are going to use the software.

4. Conclusions

The analysis we made in the previous Section states clearly that, overall, Cyber Exam and QuestionMark Perception can be considered as the best choices for Web-based testing. They scored very high in all criteria and had no major drawbacks in every aspect. Therefore, Cyber Exam and QuestionMark Perception can be considered as an option.
On the other hand, Hot Potatoes can be used for easy creation of unofficial questions and tests. Test Maestro is better suited for the production of test papers for written examinations, while Examiner is a tool that scored well in almost every criterion except the one regarding the supported question types, which is a fact that makes it unattractive. FastTEST and C-Quest are considered to be of a lower standard than its counterparts and LXR*TEST can be marked as the third best choice after the two we mentioned above. Finally, WebCT is a great selection not only for creating Web-based tests, but also for the generation of complete university courses that are going to be taught using the Web instead of the classic lecture-based teaching.

References

[Brooks 97] Brooks D., (1997). *Web Teaching: a guide to designing interactive teaching for the world wide web*, Plenum Pub Corporation.

[Khan 97] Khan B., (1997). *Web-based Instruction,* Englewood Cliffs, N.J.

[Hall 97] Hall B., (1997). *Web-based Training Cookbook*, Wiley.

[Economides 99] Economides A.A. & Georgiou A.C. & Karagiannidis C. (1999). Acceptability Evaluation of Web Based Courses, *19^{th} World Conference on Open Learning and Distance Education,* Vienna, ICDE 1999

[Economides 97] Economides A.A. (1997), Training and Continuing Education on Internet, *3rd Conference on Informatics and Telematics Applications,* Thessaloniki, 1997

[Mamoukaris 99] Mamoukaris K. & Economides A.A. (1999), Web-based distance education systems, *17th Annual Conference AOM/IAOM,* San Diego, 1999

[QuestionMark Perception] http://www.questionmark.com/

[Hot Potatoes] http://web.uvic.ca/langcen/halfbaked

[Test Maestro] http://www.rredware.com

[Examiner] http://www.assess.com/examiner.html

[FastTEST] http://www.assess.com/fasttest.html

[LXR*TEST] http://www.lxrtest.com

[WebCT] http://www.webct.com

[Cyber Exam] http://www.vlearning.com/cyberexam/

[C-Quest] http://www.cogentcorp.com/ccc/cq.html