

Q-ONLINE

Integrating a questionnaire system in an organization

Cláudio Sapateiro, Hugo Gamboa, Nuno Pina Gonçalves

*Escola Superior de Tecnologia de Setúbal,
Rua Vale de Chaves – Campus do IPS, Estefanilha,
2910-761 Setúbal,
Portugal*

Keywords: Internet, Questionnaire, Platform, Human Computer Interaction, Data Mining

Abstract: Organizations are increasingly using questionnaires as a form of collecting data. Our work¹ focus in a creation of a web based questionnaire platform, Q-Online, that managed a multi-questionnaires projects on a multi-users environment. The project goal is to provide a standard structure to collect data in several organization situations, particularly answering the needs of our organization: a school of technology. Examples of application of the platform are the collection of data from students or teachers and the usage inside of an Elearning system. The system was tested in a major school questionnaire focusing the entire school population. We present preliminary results from this questionnaire. The user interaction during the answering of the questionnaire was monitored in order to enable future retrieve of behavioural information. The data analysis developed permits a first overview of the questionnaire answers while Data Mining techniques will be provided to identify relevant information in the answers data.

1 INTRODUCTION

In the last years we have noticed that organisations have paid a great attention to collect and organise useful information to their activity areas. Present state of Web applications and technologies allows the easy distribution of questionnaires, becoming a current way of collecting information used in several departments and organisations (Nichols et al, 1998).

Analysis, processing and consolidation of received information between departments could not be easy, losing this way (or at least becoming more difficult) the chance to obtain information that can represent a real profit of the total information got. Also users are confronted with different formats, structures and possible repeated questions. All this factors doesn't promote the right filling of the questionnaires and information reliability (Robert D. et al., 1998; Don A.Dilman et al., 1999).

This paper reports the creation of the Platform Q-Online, a web application that allows creating and generating questionnaires in an environment that is multi-user, multi-questionnaire and with security layer having the internet as support. This platform will allow the reutilisation of already developed projects, as well as the profit of the compatibility of

the information, also allowing the eventual crossing of inter-projects information.

The interaction behaviour of the user, while answering a questionnaire is monitored by the platform Q-Online. This is done by integrating a system called Web Interaction Display and Monitoring (WIDAM) that remotely monitors the human computer interaction of a user while visiting the several questionnaire pages. Our research goal is to obtain a richer understanding of the several user answers by comparing his interaction behaviour while he is exposed and give answers to the questions.

Another objective of the Q-Online platform will be, considering the uniformization of the information structure, to give an improved analyse of the recovered information. Immediate basic results of the obtained information are given. Based on these first analyses the information will be filtrated and sent to future consolidation, using Data Mining techniques.

In reality we have nowadays several systems to create and generate questionnaires, some of them free of charge and at our disposal in a web-based format, and others are commercial products. In a general way all of them present almost complete

¹Part of the implementation of Q-Online was possible due to João Daniel graduate student work, from ESTSetúbal, to whom, we would like to express our thanks, for his work and commitment.

functionalities and similar among them. However the Q-Online aims to add some extra functionalities to the ones we are used to. First of all developing our own product will easily allow the integration we intend with other systems like the E-learning platform (RoBling, G. et al. 2000; Gamboa, H. et al. 2001), which exists in ESTSetúbal, and with the Intelligent Miner to the Data Mining analysis (IBM Redbooks 2003). In fact this last point is also an innovation to the usual results processing of the existing tools. The major test to the system was done with a questionnaire, involving all the ESTSetúbal 2500 student population. At last the monitoring of the questionnaires filling integrated in a tool with these characteristics is not usual too.

2 SYSTEM ARCHITECTURE

The Q-Online platform architecture is based on a questionnaire management tool that enables users edit, test, visualize, and fill the questionnaire by the authenticated users. Other questionnaires, from the platform can be totally or partially reused, as well as sections, or simple questions. Authors can also permit other authenticated users to access and use their existing questionnaires.

Figure 1 presents a use case for Q-Online platform. There are several actors that interact with the system. Administrators, with full access to the platform, manage new users and configure their privileges. Authors create and manage their questionnaires. Users fill questionnaires created by authors, and visitors visualize obtained results from the questionnaires.

Authors, after their authentication, start questionnaire creation with sections definition and their associated questions. For each question it is possible to define some attributes: type, properties, possible answers and dependencies.

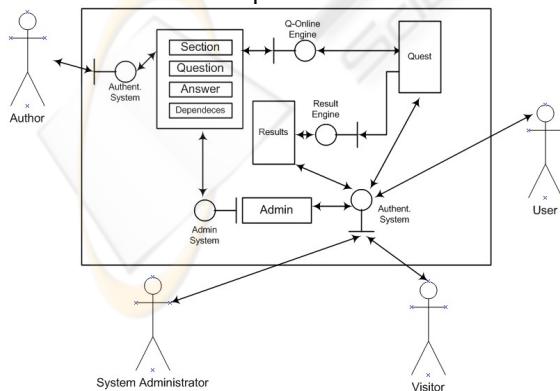


Figure 1: A use case for the Q-Online Platform

When users Access the platform, Q-Online engine dynamically generates the questionnaire to fill, from the structure created by authors. Visitors can access questionnaires results created by the Result Engine, based on the obtained answers.

2.1 Users and authentication

Q-Online platform has an initial authentication system for administrators, authors, users and visitors. This system is managed by the administrator.

Questionnaire may be public or private depending on the author's choice. If it is public, the filling will not need the authentication system management.

As well as filling, results visualization can also be managed by the author, and can be public or private. If it is public, any visitor may have access to results without authentication.

2.2 Q-Online structure

Q-Online is based on a structure with questionnaires and their contents. From these contents the Q-Online engine generates the wanted questionnaire. This structure has, for each questionnaire, one or many sections, and inside each sections their correspondent questions.

Conceptually, as it is possible to observe in figure 2, questions can only exist if sections exist. Each question has possible answers, dependencies and other properties. These properties can be managed by the author and are related to question types and if they are mandatory or not.

Each question has many possible types, from multiple answers, single or multiple classification answers and text answers.

Questionnaire results are obtained and stored for each question, inside each section, for a given questionnaire.

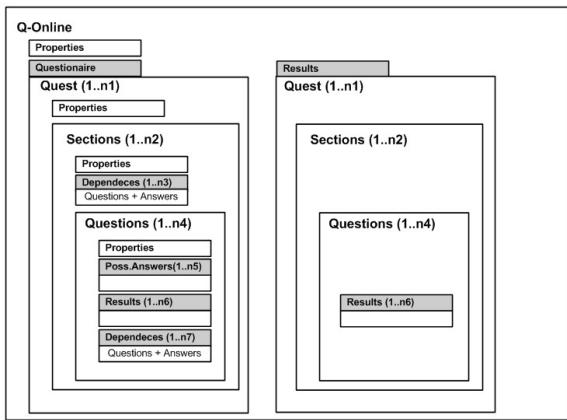


Figure 2: Q-Online structure

2.3 Dependencies

Dependencies are an important issue among questions and sections because, many times, questionnaires have a large number of questions and only a subset of them are applied to a specific user, due to answers given in other questions. For example, if a questionnaire has 100 questions, maybe only 60% of them are to be asked to the user. The best way to avoid having the user to see all the questions is to create dependencies among questions and sections, and showing him only the questions he need to fill.

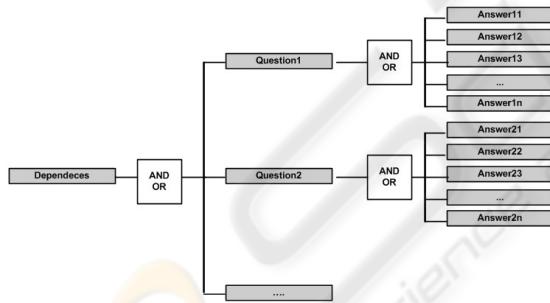


Figure 3: Dependencies structure

Having in mind the platform's concept, a section may depend from one or many answers to a question. In this context, a section may only appear if a certain condition – a specific answer to a given question - is true. This condition can be a conjunction or a disjunction or both. Questions may have the same kind of dependencies.

Figure 3 shows created dependencies structure.

2.4 WIDAM system

The Web Interaction Display and Monitoring system (WIDAM) (Gamboa, H. et al., 2003) is a web based

client-server application that offers the capability to record the user interaction. The WIDAM system allows the usage of an interaction monitoring system directly over a web page, without the need of any installation, requiring low bandwidth.

The client script collects several interaction events and evokes the remote procedure in the server via XML-RPC sending the interaction data. For each event the data passed is:

Event Id; X position; Y position; Object-id; Extended data; Timestamp. The X-Y positions are the mouse coordinates where the event occurred. The Object-ID identifies the DOM object in the web page where the event was called, used to identify the text or picture in the page related to the generation of the event. The extended data carries information about the pressed key. The time stamp identifies the time when the event occurred.

The WIDAM system was also used in a biometric verification system, where the user interaction with a web page was monitored and analyzed to validate the identity of the user. The results were comparable to other behavioural biometric techniques like voiceprint and signature verification. The work is reported in (Gamboa, H. et al., 2003a).

2.5 Results analysis

Users obtain questionnaire results after their filling. Users may view the most and less answered questions, ignored questions, as well as obtained percentages for each possible answer and for each question with no answers at all.

The system has three types of possible answers: categorical, numerical and free text. In case of numerical type, the system offers simple statistics like average, standard deviation, maximum value, minimum value, modal value, mean value.

This statistical results offered by Q-Online platform help visitors making a first selection to data, giving a higher confidence level to the selected answers.

Questionnaire results will be later imported to a data mining tool, where some mathematical analysis will be done to help validating the results, and to give some patterns and associations related to the obtained data.

Data Mining which is also referred to as knowledge discovery in databases means a process as non trivial extraction of implicit, previous unknown and potential useful information (such as knowledge rules, constraints, regularities) from data in databases (Piatetsky-Shapiro et al., 1996).

With this kind of studies, visitors, authors and authenticated users will be able to identify some

clusters where data have the same characteristics. Visitors may also identify associations between data, and may also associate all this information with statistical information obtained from Q-Online platform. The Data Mining tool can also help users to identify more complex statistical information, such as bivariate statistics and correlation analysis of the given answers.

3 EXAMPLE OF APPLICATION

The High School of Technology of Setúbal (ESTSetúbal) part of the Polytechnic Institute of Setúbal is an example of organisation in which the several departments produce questionnaires to retrieve information, using each one his own structure and technology.

Every year, in the beginning of the scholar year the students fill in a questionnaire which, later, allows the school to take new measures to improve the school conditions and promote the students success.

About two years ago, ESTSetúbal have started a project of receiving and analysing the student's information using Data mining tools.

With the platform Q-Online we constructed the questionnaire scholar success 2003 to be filled online. This questionnaire consisted of a list of questions used in previous years combined with a list of questions that the work group of Data Mining has considered important.

The scholar success questionnaire target, was a total of approximately 2500 students. We obtained 1800 questionnaire fills. After checking the questions, based on the elimination of the non correct answered questionnaires or not answered at all we obtained 1653 questionnaires to analyse.

To each asked question Q-Online provide two kinds of answer, the categorical and the numerical. In the categorical is possible to obtain the quantity of given answers for each category as well as the associated proportion. To the numerical ones is possible to obtain some simple statistics, like the average, the highest and the lowest value. It is also possible to create a top-ten and a bottom-less with the given answers to each question.



Figure 4: Q-Online results screenshot

In the figure 4 is shown a snapshot of Q-Online results analysis. The text of the picture is in Portuguese since the organization is a Portuguese school.

4 CONCLUSIONS AND FUTURE WORK

The project of the platform Q-Online has been validated with the above mentioned practical case.

In reality the described questionnaire has been obtained with an inter-departments effort from ESTSetúbal that ended with the isolated creation of questionnaires and the repeated information collection. In fact the obtained result serves the required specifications from the different involved departments, since each of them can analyse the questions they have elaborated. The use of a common structure makes possible the analysis of eventual crossing of all the obtained information.

The developed system proved that the application has always been stable, even when the number of users has been around 500.

The collected interaction will be analyzed using pattern recognition methods in order to detect a set of behavioural characteristics. We have the goal of producing probabilistic statements about the user behaviour while answering a question. One of the first tasks is to construct the set of the behavioural states we will be able to detect. The results of the study could be extended to different questionnaires and testing environments with a particular interest in

Intelligent Learning Environments where the detection of behavioural states would be used to give feedback to the learner and better conduce the user activity and experience.

The future work will be the integration with a data mining server, allowing that results of a data mining analysis be available to users that access the platform from their machines.

REFERENCES

- Nichols E.; Sedivi, B..1998. In press "Economic data collection via Web: A census bureau case study". Proceedings of Survey Methods Section. American Statistical association annual meetings, Dallas, Texas.
- RoBling, G.; Freisleben, Bernd. 2000. "Topkapi: A tool for performing Knowledge tests over the WWW". Interdisciplinary Conference on Electrical, Electronics and computer Engineering in the third millennium (ICECCE)2000.
- Robert D.; Toratora, Jon Conradt; Don A.Dilman.1998."Influence of plain vs fancy design on response rates". Proceedings of Surveys methods section. American Statistical Association Annual Meetings. Dallas, Texas.
- Don A.Dilman, Robert D. Toratora, Jon Conradt, Dennis Bowker.1999. "Principals for constructing Web surveys". Technical report. Washington State University.
- Integrating IBM DB2 Intelligent miner visualization into your web application. IBM Redbooks.2003.
- Gamboa, Hugo; Ribeiro, Nuno V.; Holt Anatol, Fred, Ana. 2001. An e-learning Platform Concept, 2a Conferência Internacional em Tecnologias de Informação e Comunicação na Educação: Chalanges 2001.
- Gamboa, Hugo; Ferreira, Vasco. 2003.WIDAM - Web Interaction Display and Monitoring. 5th International Conference on Enterprise Information Systems, ICEIS'2003.
- Gamboa, Hugo; Fred, Ana. 2003a. An Identity Authentication System Based on Human Computer Interaction Behaviour. Proceeding of the 3rd Workshop on Pattern Recognition in Information Systems PRIS 2003.
- Piatetsky-Shapiro, G.; Fayyad, U.; Smith P..1996. From data mining to knowledge discovery: An overview. In U.M. Fayyad, G. Piatetsky-Shapiro, P. Smith, and R. Uthurusamy, editors, Advances in knowledge Discovery and Data Mining, pages 1-35. AAAI/MIT Press, 1996