Poster paper in the 21st Annual NACCQ, Mann, S. & Lopez, M. (eds)
Auckland, NZ, July 4-8 2008 www.naccq.ac.nz

SUMATO: an xml-based Survey Management System

Bevan Yot Chan, Alister Galpin, Jarrod Irwin, Dave Kennedy
Christchurch Polytechnic Institute of Technolgy
Christchurch, NZ
kennedyd@cpit.ac.nz

SUMATO (Survey Management Tool Online) is a web-based application for the creation, collection and analysis of surveys. The aim is to replace the traditional paper-based questionnaires used by researchers. A prototype was first developed that used a MySQL database. The core functionality was available via a very basic user interface. The latest version replaced the database with a collection of xml files and improved and extended the user interface.

In some situations it is easier, simpler and smaller to use xml files rather than a DBMS. SUMATO is one of those situations. Relational databases are good for storing highly structured data. They are not so good at managing semi-structured data. Questionnaires comprise a variable number of questions of varying types (short answer, long answer, multi-choice, etc). The required xml data structure proved to be much simpler than the equivalent MySQL database schema. The display of questionnaires was accomplished using xslt files. Using xml files meant the application is easier to install and run, easier to maintain and platform independent. For this application to be a usable in-house system further development is required to improve the functionality and address security and multi-user control issues.

Keywords: xml, database, survey management.

Many web-based applications require a DBMS to store and maintain data on products, customers, orders, etc. But for some situations a DBMS is an additional overhead that could be replaced by an xml data structure (Jouvernaux, 2004). Such a structure is an advantage for storing semi-structured data and can easily take advantage of related technologies such as xslt (Morrison, 2000). Situations where there is a small amount of data, few users and modest performance requirements may well be better suited to the use of xml files instead of a RDBMS (Bourret, 2005). The advantages of such an approach are a less complex 2-tier model (Jouvernaux, 2004), portable data, and a smaller application. The major advantages are seen when the data is semi-structured as the xml preserves the document structure, related technologies such as xslt can be used to display the documents and xpath can be used for searching. As (Bourret, 2005) notes semi-structured data is not only difficult to store in a RDBMS because it requires many different tables this in turn makes for slow retrieval times. Rebuilding a document for presentation is also much more complex than using an sxlt stylesheet attached to an xml document. There are also some disadvantages to be aware of such as a lack of security, no transaction handling, no integrity constraints and no multi-user control (Bourret, 2005).

1.1 SUMATO1

Many researchers use questionnaires to gather data. The first iteration of SUMATO was developed as a web-based application that would enable a researcher to create and edit surveys using a browser. Each survey has an associated unique keyCode. This keyCode is sent to those invited to answer the survey who can then login, enter the keyCode, and complete the survey. An initial prototype of this application was built using a MySQL database, php using classes and a model, view, controller structure, and javascript. The functionality required to create a survey that contained multiple questions of various types, to answer surveys, handle logins and survey ids, and store completed surveys was built but the user interface was very basic. The associated database contained thirteen tables.

The aim was to create an application that could be used by all researchers at Christchurch Polytechnic Institute of Technology (CPIT). This would involve installing the software on an IT department server. The IT department required that it should be platform independent, not dependent on a DBMS, and easy to install and maintain. They suggested data storage using xml files. A much more user-friendly interface was also required.

1.2 SUMATO2

The second iteration of SUMATO started with the php programs from the first iteration. The DBMS was replaced by a collection of xml files. The thirteen tables were replaced with five xml documents. The database handling classes were modified to interact with the xml files and Ajax was used instead of javascript. A menu system was created that allowed a researcher to create, edit, and delete surveys, and to display completed surveys. The presentation of surveys was handled using an associated xslt stylesheet. A simple survey summary display was also created.

1.3 SUMATO3

A report on SUMATO2 has been completed. Recommendations for the next development cycle are:

- Refactor the code using the CPIT standard Framework
- Change the xml structure to comply with the Triple-S Survey Interchange standard
- Address the security risks that have been identified
- Address the usability issues that have identified and carry out further in-depth testing
- Save each completed survey in a separate xml file to overcome concurrency problem

2 References

Bourret, R. (2005). XML and Databases. Retrieved 12th March 2008, from http://www.rpbourret.com/xml/XMLAndDatabases.htm

Jouvernaux, M. (2004). *XML in action: A data driven web site with XML*. Paper presented at the National Advisory Committee on Computing Qualifications, Christchurch, New Zealand.

Morrison, M. (2000). *XML Unleashed*: Sams.