Journal of Software Engineering and Applications, 2014, 7, 713-719
Published Online July 2014 in SciRes. http://www.scirp.org/journal/jsea
http://dx.doi.org/10.4236/jsea.2014.78066

Security and Audit Trail Capabilities of a Facilitated Interface Used to Populate a Database System with Text and Graphical Data Using Widely Available Software

Kevin Beland, Kirk Larson, Thaine Rowley, Matt Mueller, Christopher Smith, Al Rizzo, Daniel Valandra, Marc Rendell

Creighton University and the Rose Salter Medical Research Foundation, Omaha, NE, USA
Email: rendell@asndi.com

Received 19 May 2014; revised 15 June 2014; accepted 10 July 2014

Abstract

In prior work we described techniques used to capture, store, and retrieve narrative data forms from a database using widely available off the shelf software. This concept makes use of the security and versioning features of database architecture together with those of image-supported electronic document information capture. We present herein the security and audit trail features of our construct. Unique usernames and passwords are used at the operating system level to enforce client validation and control access to the database system via an electronic signature using a Boolean logic algorithm. We used the database domain to capture original data and any changes to the data, developing an audit trail displaying identification information. Version controlled and change tracked electronic documents can be retrieved through the client side web interface to enable direct search operations, thus uncoupling the client from database query languages.

Keywords

InfoPath, Electronic Form, Data Storage

1. Introduction

The entry of information into structured databases allows the implementation of searching methods on the data to reveal relationships, infer knowledge, and provide information. Yet, the development of a database requires considerable planning to ensure correct storage and retrieval of desired items while maintaining correct rela-

How to cite this paper: Beland, K., *et al.* (2014) Security and Audit Trail Capabilities of a Facilitated Interface Used to Populate a Database System with Text and Graphical Data Using Widely Available Software. *Journal of Software Engineering and Applications*, **7**, 713-719. http://dx.doi.org/10.4236/jsea.2014.78066

tionships between the data. An alternative is to create electronic forms in the narrative domain by allowing free form template design and capturing all visible information where data schemas can be inferred and the plasticity of information capture is increased. Information captured in such digital repository mediums is inherently searchable using Structured Query Language (SQL) commands, and storage of electronic documents requires less overhead compared to physical media. In 1995, 80% - 90% of organizational information was claimed to be in documents rather than structured databases [1]. In a previous paper, we presented the features of a narrative domain interface used to populate a database system with both textual and graphical data constructed using widely available software. However, even when considering the storage and information capture benefits of electronic document storage, no less attention should be given to the methods of security necessary to ensure proper retrieval and editing of the information contained within the documents. Electronic documents must be properly secured and versioned to ensure that data are not compromised after indexing and storing the contents.

In this paper we present the security and audit capabilities of our electronic form repository implementation and the narrative data capture that creates it. In summary, our implementation pipeline begins with a narrative form interface (Microsoft InfoPath), where electronic form templates containing fields open for client input are created to capture both text and graphical data as designed for that particular template. The data elements are then parsed and exported to a SQL compliant database where document retrieval is facilitated through recall of the entire narrative view of the document. Using flexible, user-composed templates and allowing clients to enter data into forms remove the overhead, complexity, and time needed for designing proprietary database interfaces [2].

Once a client has finished entering or changing data, the client submits the electronic form to a web service hosted on a secured Local Area Network (LAN) along with a verifiable signature and unique password. The LAN web service that receives the form parses the data for key reference fields and identification metadata for indexing. Each form is given a unique primary identification key generated and monitored by the database which is used to identify documents with multiple existing versions edited by one or more clients. The entire form is retrieved from the database for data entry into the narrative interface when changes or updates need to be applied to the electronic form. Once retrieved and edited, the updated form is re-submitted to the database via the LAN web service allowing the revised form to be recorded as a new version for change detection and audit purposes. In the following sections the security and audit characteristics of our electronic form repository and our LAN client interface are discussed.

2. Methods

2.1. User Verification

Our implementation exists on a private LAN hosted by in-house servers using Windows Server 2008 R2. Users are authenticated into shared network spaces with an Active Directory Server requiring a username and password for access. Additionally the narrative form templates include username and password fields used to identify the client who submitted the form. Submission to the database through the LAN requires the username and password fields on a form to be verified against the Active Directory to ensure that access to the database is concordant with the identity and location of the user. This ensures that only authenticated users can add information to the database. Only users with administrator permissions can view or change data without using the client side narrative interface.

2.2. Submission Process and Electronic Form Storage

The transfer of data from InfoPath to the database server is facilitated through Hypertext Transfer Protocol (HTTP) requests as defined by RFC2616 [3] and recieved using Simple Object Access Protocol (SOAP) protocols [4]. When an electronic form is submitted to the receiving LAN hosted web server, the forms contents are parsed for identifying meta-data contained within the document. Important document identifying fields retrieved during the process include username, title, domain, date-time, and version information. This identifying data are collected and inserted into the database along with the form data. The database accepts the electronic form meta-data retrieved from the web interface and indexes the data along with storing the electronic form in Extensible Markup Language (XML) format. All images added to the form remain encoded in their original format within the XML document. One of the captured database fields contains the entire XML formatted electronic document providing a means of quick and thorough retrieval of the entire document as well as the individual data elements.

2.3. Version Control

Each unique document inserted into the database is given two identification numbers. The first ID number is the unique identifier for the document template. The second ID is simply incremented from zero for each unique submission of a form and is used to track the linear version of the document. The application is resubmitted every time a user has retrieved the document and finished entering changes, thus allowing the database to store every version of the document. Keeping and storing each version of the document ensures document integrity at each time point of submission while providing a complete, visually narrative view of the document's contents.

2.4. Audit Trail Functionality

Our electronic document storage architecture provides an audit trail of each document lifecycle by redundantly storing every document version. Every change made to a document that generates a new version is inherently tracked through the storage of that document. All changes found between versions are attributed to the user that submitted the document to the database. A view of a document's audit trail as it exists within the database is provided to all clients through the same web interface used to retrieve documents. Identity meta-data and time-stamp information for each document version is provided by the audit trail view.

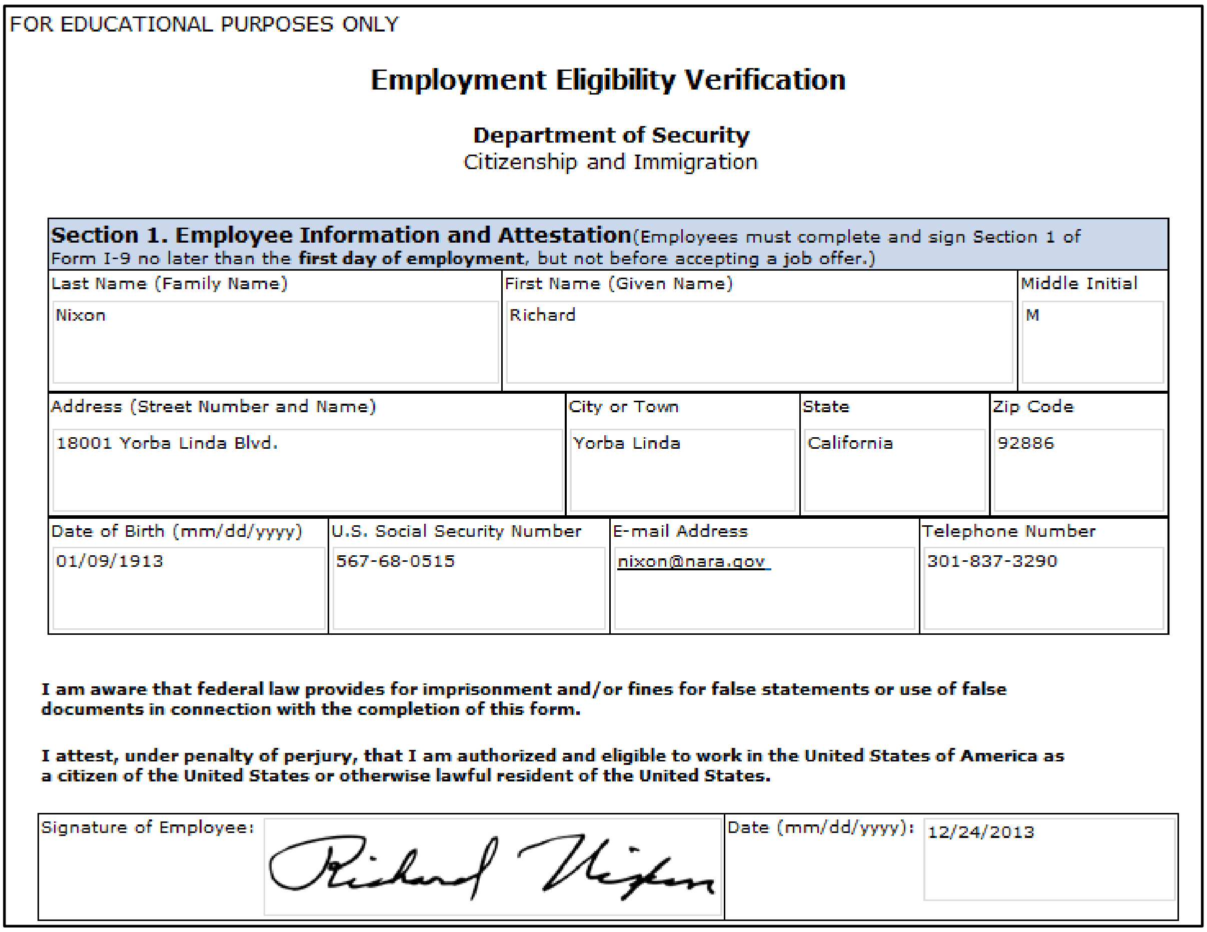
3. Results

We have created a fictitious example of an employer verification process to visualize how our system accepts documents, retrieves documents, secures entry, and tracks changes. A sample narrative interface provided for electronic document data collection can be seen in **Figure 1**. In our **Figure 1** sample Richard M. Nixon is provided a narrative format for entering personal data that needs to be verified by a series of agencies before he can be considered for federal employment. The form contains data connection information for the purpose of submission to the web server, which accepts a valid username and password. The form is submitted at each original data entry point such as insertion of identification documents including his drivers' license, fingerprints, and tax return in the form of images as well as his signature and those of reviewing officials.

In our employment verification example as seen in **Figure 2**, J. Edgar Hoover, head of the FBI, retrieved the document and verified Nixon's fingerprints by entering new information and inserting a signature to the document. The document was then submitted, capturing a new version of the document along with the changes made, the time and date the changes were made, and who made the changes. Older versions of the document remained stored in the database with identifying metadata to provide an audit trail of changes. An example of a captured audit trail has been summarized in **Figure 3**, which illustrates an original submission of an employment verification form followed by 4 validation events in which changes were made to the document by users registered to the LAN. The versioning events captured by the database in our example include verification of fingerprints by J. Edgar Hoover, former director of the Federal Bureau of Investigation, verification of prior employment by Dwight D. Eisenhower and Gerald Ford, formerly Presidents of the United States, and a final sign off by Barack Obama, the 44th President of the USA. The changes made to the document at every validated submission were captured and attributed to the user that submitted a changed version of the document. Audit teams reviewing the history of the example document and its revisions as seen in **Figure 3** would have access to the document at every time point where changes were submitted.

4. Discussion

We have designed a data pipeline using a flexible interface for entry of both textual and graphical data to a database using licensed software. Our purpose herein was to describe the security features and data-tracking capabilities of our construct. We use a web based internal LAN approach to communicate between a narrative form and the database. Although document transmission through HTTP without encryption is inherently insecure (see RFC2660 [5]), security is achieved through limiting LAN access to on-site connections using Active Directory provided username and password authentication and by firewalling all incoming and outgoing connections on our closed system. Furthermore, the electronic form database restricts query access to database administrators



FOR EDUCATIONAL PURPOSES ONLY

Employment Eligibility Verification

Department of Security
Citizenship and Immigration

Section 1. Employee Information and Attestation(Employees must complete and sign Section 1 of Form I-9 no later than the **first day of employment**, but not before accepting a job offer.)

Last Name (Family Name) Nixon	First Name (Given Name) Richard	Middle Initial M

Address (Street Number and Name) 18001 Yorba Linda Blvd.	City or Town Yorba Linda	State California	Zip Code 92886
Date of Birth (mm/dd/yyyy) 01/09/1913	U.S. Social Security Number 567-68-0515	E-mail Address nixon@nara.gov	Telephone Number 301-837-3290

I am aware that federal law provides for imprisonment and/or fines for false statements or use of false documents in connection with the completion of this form.

I attest, under penalty of perjury, that I am authorized and eligible to work in the United States of America as a citizen of the United States or otherwise lawful resident of the United States.

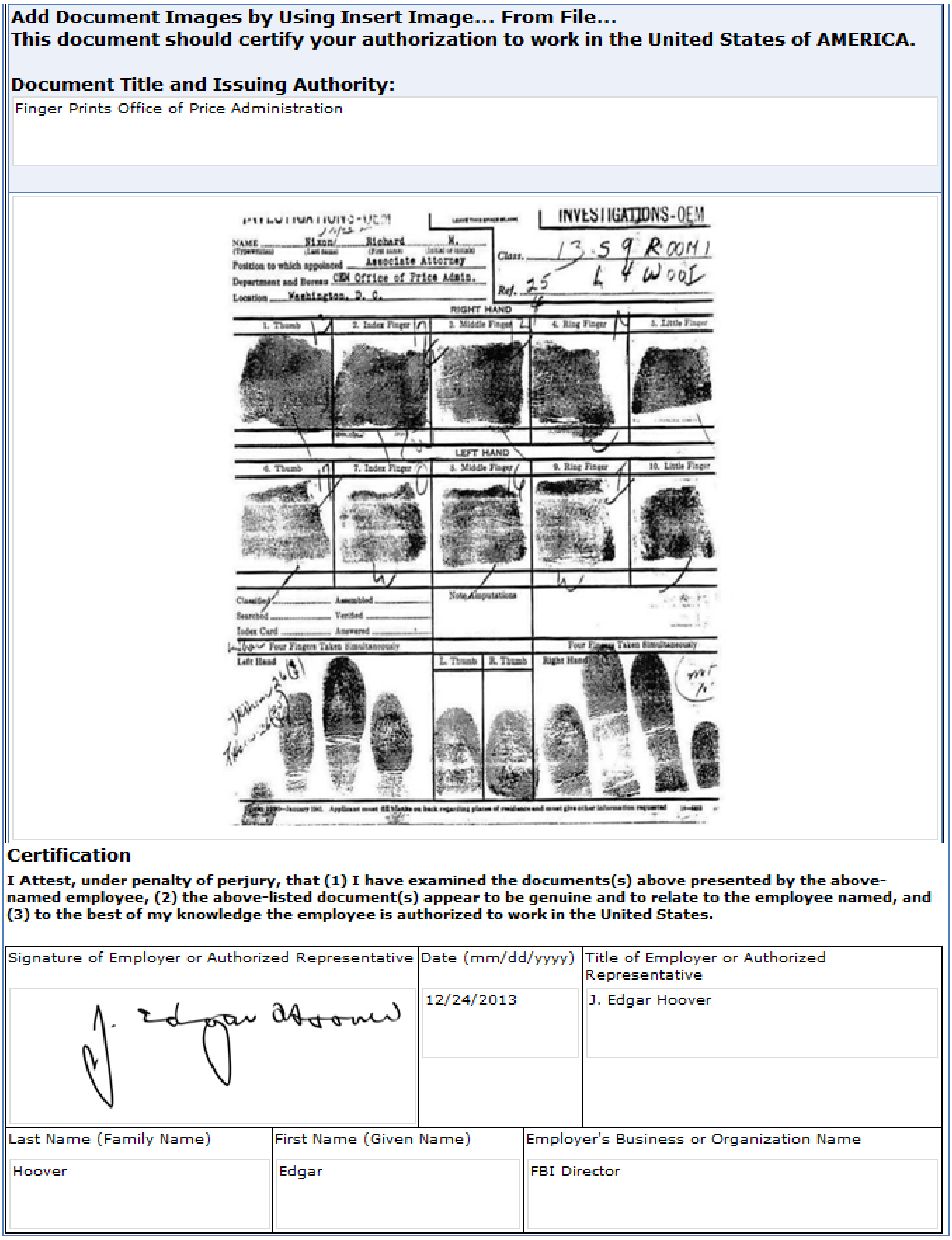
Signature of Employee: Richard Nixon	Date (mm/dd/yyyy): 12/24/2013

Figure 1. Sample employment verification electronic document.

since all data entry occurs in the narrative interface. In our presented example we have not elected to encrypt the data within the database. If encryption techniques are used, the encryption must not prevent the flow of data toward external parties due to poorly executed audit protocols [6]. It should be emphasized that an advantage of our approach is that only the submission interface is accessible to users, not the database itself. This limits the possibility of compromise to the individual database elements. In our presented electronic document storage system architecture, audit functionality is provided through redundant storage of the entire document as it existed during each submission event. By providing an audit team access to the document at every point of its life time, the audit team has access to undeniable proof of the state of a document after every possible modifying operation on the individual elements contained within the document. Additional change tracking detection within the electronic documents text fields can be deployed using text comparison tools by comparing each new document version with the prior version [7] [8]. Tracking changes of text fields within each document provides a quick method of detecting which document version received what changes and may detect small changes that would otherwise go unnoticed.

For some domains, only the most recent version of a document should be retrieved in a linear fashion thereby removing the possibility of version branching. Some domains may need certain documents to branch off into multiple versions as data are input into different fields from multiple clients, where ultimately the final document is created from a merge of all document branches. For the formerly mentioned domain requirement, a useful version control implementation may be a Lock-Modify-Unlock solution, also known as mutual exclusion, for its ability to prevent multiple clients from entering and submitting data at the same time to the same area of the document [9]. The latter domain example, where document versions may come from multiple concurrent changes in different areas, must use a version control scheme implementing a Copy-Modify-Merge solution [10].

There are varying requirements in providing an audit trail for electronic documents. Most audit requirements



Add Document Images by Using Insert Image… From File…
This document should certify your authorization to work in the United States of AMERICA.

Document Title and Issuing Authority:

Finger Prints Office of Price Administration

INVESTIGATIONS-OEM

NAME (Typewritten) Nixon/ (Last name) Richard (First name) M. (Initial or initials)
Position to which appointed Associate Attorney
Department and Bureau OEM Office of Price Admin.
Location Washington, D. C.

Class. 13 5 9 ROOM I
L 4 WOOI
Ref. 25

RIGHT HAND

Certification

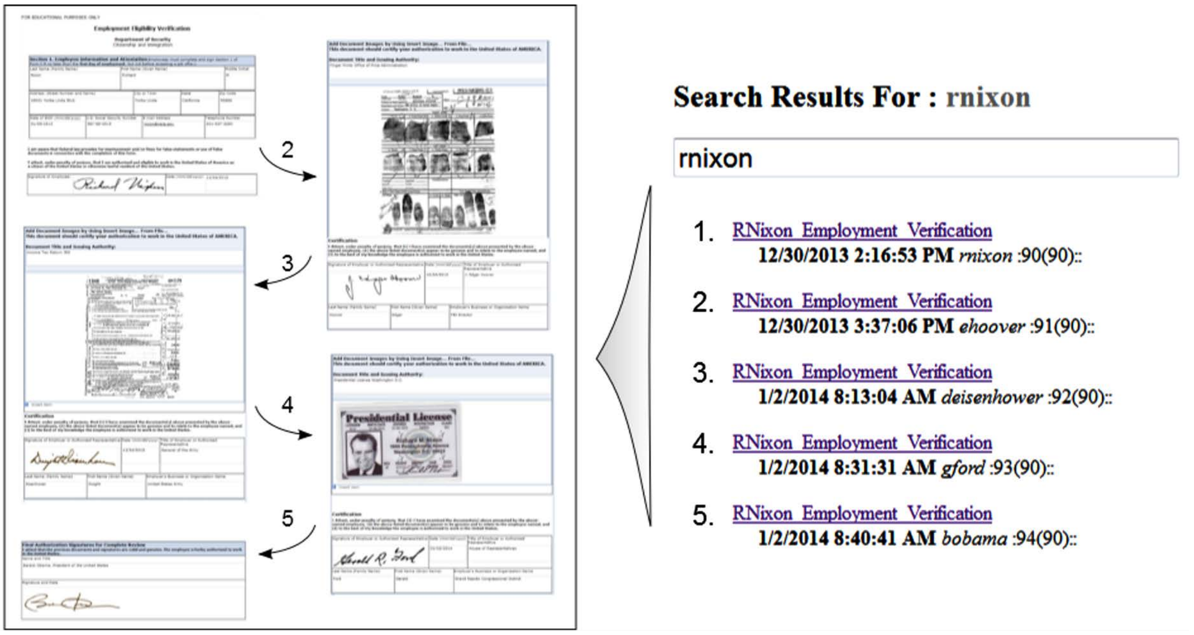
I Attest, under penalty of perjury, that (1) I have examined the documents(s) above presented by the above-named employee, (2) the above-listed document(s) appear to be genuine and to relate to the employee named, and (3) to the best of my knowledge the employee is authorized to work in the United States.

Signature of Employer or Authorized Representative	Date (mm/dd/yyyy)	Title of Employer or Authorized Representative
J. Edgar Hoover	12/24/2013	J. Edgar Hoover

Last Name (Family Name)	First Name (Given Name)	Employer's Business or Organization Name
Hoover	Edgar	FBI Director

Figure 2. Finger print verification form.

agree that all modifying operations on electronic documents must be tracked along with what modification occurred, when it occurred, and why. Using a public company accounting domain as an example, audited docu-



Search Results For : rnixon

rnixon

1. RNixon Employment Verification
   12/30/2013 2:16:53 PM *rnixon* :90(90)::
2. RNixon Employment Verification
   12/30/2013 3:37:06 PM *ehoover* :91(90)::
3. RNixon Employment Verification
   1/2/2014 8:13:04 AM *deisenhower* :92(90)::
4. RNixon Employment Verification
   1/2/2014 8:31:31 AM *gford* :93(90)::
5. RNixon Employment Verification
   1/2/2014 8:40:41 AM *bobama* :94(90)::

Figure 3. Linear audit trail verification form pipeline.

ments must be retained for 7 years and changes made to the documents must be recorded along with who executed the change, the date it happened, and why [11]. Audit requirements, as a result of an attempt to create authoritative proof of the proceedings of events and the sanctity of data through documentation, can be onerous [12]. A thoroughly monitored, reliable, and narrative-presented system is expected to provide proper documentation of operations that modify electronic documents.

5. Conclusion

In previous sections a general approach to the problem of enforcing electronic document security and change tracking was presented. This technique provides a client access to a secure, image-compatible electronic document database while tracking and storing all incoming modifications to those documents. Electronic document submissions are kept secure and available to only those clients with proper validations by limiting network access to clients with minimal trust permissions and providing an uncoupled graphical interface as a proxy for the client to submit document modifications. Audit trail functionality is provided through redundant storage of all document versions with itemized changes visualized through change tracking software libraries. Future work should focus on incorporating flexible approaches to indexing incoming data elements to allow intelligent search and retrieval protocols independent of the original database schema.

References

[1] Sprague, R. (1995) Electronic Document Management: Challenges and Opportunities for Information Systems Managers. *MIS Quarterly*, **19**, 29-49. http://dx.doi.org/10.2307/249710

[2] Topi, H., Valacich, J.S. and Hoffer, J.A. (2005) The Effects of Task Complexity and Time Availability Limitations on Human Performance in Database Query Tasks. *International Journal of Human-Computer Studies*, **62**, 349-379. http://dx.doi.org/10.1016/j.ijhcs.2004.10.003

[3] Fielding, R., Irvine, U. and Gettys, J. (1999) Hypertext Transfer Protocol—HTTP/1.1.

[4] O'Tuathail, E. and Rose, M. (2006) Using the Simple Object Access Protocol (SOAP) in Blocks Extensible Exchange Protocol (BEEP).

[5] Rescorla, E. and Schiffman, A. (1999) Secure Hypertext Transfer Protocol (S-HTTP).

[6] Wang, C., Wang, Q., Ren, K. and Lou, W. (2010) Privacy-Preserving Public Auditing for Data Storage Security in Cloud Computing. 2010 *Proceedings IEEE INFOCOM*, San Diego, 14-19 March 2010, 1-9.

[7] Fraser, N. (2012) Diff, Match and Patch Library.

[8] Kher, A. (2004) The XML Diff and Patch GUI Tool.

[9] Dix, A. and Miles, V.C. (1992) Version Control for Asynchronous Group Work. Department of Computer Science, University of York, York.

[10] Collins-Sussman, B., Fitzpatrick, B. and Pilato, M. (2004) Version Control with Subversion. O'Reilly, Sebastopol.

[11] (2004) Audit Documentation. Auditing Standard No. 3.

[12] Bronson, S.N., Hogan, C.E., Johnson, M.F. and Ramesh, K. (2011) The Unintended Consequences of PCAOB Auditing Standard Nos. 2 and 3 on the Reliability of Preliminary Earnings Releases. *Journal of Accounting and Economics*, **51**, 95-114. http://dx.doi.org/10.1016/j.jacceco.2010.06.002

Scientific Research Publishing (SCIRP) is one of the largest Open Access journal publishers. It is currently publishing more than 200 open access, online, peer-reviewed journals covering a wide range of academic disciplines. SCIRP serves the worldwide academic communities and contributes to the progress and application of science with its publication.

Other selected journals from SCIRP are listed as below. Submit your manuscript to us via either submit@scirp.org or Online Submission Portal.