J-CAATs: A CLOUD DATA ANALYTIC PLATFORM FOR AUDITORS

Sherry Huang
Jacksoft Commerce Automation Ltd., Taiwan, sherry@jacksoft.com.tw

Shi-Ming Huang
National Chung Cheng University, Taiwan, smhuang@gmail.com

Abstract

In this article we present our experience designing and implementing a new platform of computer-assisted audit techniques (CAATs), called J-CAATs, in the cloud environment. In developing this new tool, we sought to combine what we felt were useful features from two existing general audit software, i.e. ACL and IDEA. As its main characteristic, J-CAATs allow an auditor to use mobile-phone, Pad and Computer devices in the formulation of one and the same data analysis in the cloud. We illustrate this claim by discussing a prototype implementation of the J-CAATs based on an internet web approach. We demonstrate the look-and-feel of this audit tool by means of examples, and show the verification and validation of the system which compares with ACL and IDEA. We feel that the new tool provides advantages in the areas of portability, mobility, computational efficiency, and team collaboration.

Keywords: CAATs, Cloud Computing, Computer Auditing, Data Analytic, Business Analysis

1. INTRODUCTION

The increasing complexity of business environment in cloud and the high risks associated with violations of business rules have created the need for on cloud auditing. Many studies have shown that to improve the effectiveness and efficiency of the auditing works, auditors need to improve team problem-solving ability and collaborationskill. Therefore, an auditor needs to have a mobility and intelligent working environment by using information technology. The current research has begun to explore the construction of an on line group data analysis systems for auditors, but these studies are mostly limited to conceptual or experimental nature of the feasibility study, the lack of a systematic and complete

©copyright by Sherry Huang, Shi-Ming Huang

collaboration and data analytic technology research. The collaborative audit and data analytic mechanism in cloud is different from traditional human-based or computerized audit. It includes more complex human and system inter-cooperation, anauditing software as a service is become an important issue to enable auditors can work portability, mobility, computational efficiency, and team collaboration.

This article discusses some issues involved in the design and implementation of a computer tool for auditors in cloud environment. Our experience with these issues occurred while developing such a tool. The work has been heavily influenced by Audit Command Language (ACL) and Interactive Data Extraction and Analysis (IDEA).We felt that there were strong points in each of the software and that it would be interesting to see if the strengths could be combined in the cloud environment. The resulting interface is very similar in appearance to ACL, but the underlying implementation and semantics are derived from Relational Algebra and Relational Calculus. In fact, we implemented the system by first writing an interpreter for anaudit command subset in a relational database and then progressively mutating it to resemble the language.

At this point we had enough of a framework in place to begin building a new platform ofCAATs in the cloud. This process is ongoing, but we feel that we are well on the way to building a complete and useful piece of software.

2. RELATED WORK

Computer-assisted audit techniques (CAATs) are a growing field within the computer auditing and forensic profession. It is dedicated as more advanced software for data analysis and fraud detection in the recent year. Functions of CAATs include importing computerized data; thereafter other functions can be applied: the data can be e.g. browsed, sorted, summarized, stratified, classified, joined, duplicated, fuzzy duplicated, benford analyzed, etc. The use of CAATs allowed the auditor to searchthrough millions of transactions for audit-specified criteria in hours. Further,CAATs permitted the auditor to adjust the criteria after reviewing the initialresults. All audit phases can be supported with modern audit software (David Coderre, 2009).

Two of the most popular software packages are ACL and IDEA (Jim Kaplan, 2012). Both commercialized in late 1980s, such as ACL in 1987 and Caseware IDEA in 1988. This is due to the fast growth of micro-computers. In the modern Internet age, no auditor need workin isolation. Traditional manual audits (Audit 1.0) have existed for centuries fulfilling many needs. Although the IT audit by using CAATs (Audit 2.0)emergedfrom the 1980s, and most all businesses are currently computer based, still very few auditors are ITenabled(Protiviti,2015). This delay of IT adoption can be partially attributed to the

©copyright by Sherry Huang, Shi-Ming Huang

conservatism and rigidity of the profession, as wellas the calcifying effect of increasingly obsolete regulation (Liu and Vasarhelyi 2014), but also to the lack of quality tools thatwould allow traditional auditors (i.e., those without IT and analytics training) to automate the functions that they currentlyperform manually (Brown-Liburd, Issa, and Lombardi 2015).

A new Audit 4.0 concept piggybacks on technology promoted by Industry 4.0 to collect financial and operational information, as well as other audit-related data from an organization and its associated parties (Dai and Vasarhelyi, 2016). It analyzes, models, and visualizes data in order to discover patterns, identify anomalies, and extract other useful information for the purpose of providing effective, efficient, and real-time assurance. It will significantly change the auditing profession by automating current procedures, enlarging their scope, shortening timing, and eventually improve the overall assurance quality.

To strengthen the effectiveness of internal audit functions, 360 degree Internal Auditing concept is addressed by *(*Maxwell, Andrew, Eichhorn, Michael, Seddighi, Hamid and Smith, Peter, 2012). A new cloud concept of CAATs is needed to enable the auditors can work any time, any device, any platform, and anywhere.

3. SYSTEM ARCHITECTURE

The J-CAATs is a meta-language framework for an auditor. We added ACL-like and IDEA-Like features in several stages (see Figure 1). First, we altered the meta-language parser so that the syntax would resemble that of ACL or IDEA. This created a major advantage to enable auditors who do not need to learn a new language and can emphasized on data analytics. The Share-and-Leave scheme has applied to enable an auditor can easily share their audit project and develop the collaboration network, i.e. Auditor Social Network.

The system is developed by using the most popular internet protocol, i.e. HTTP and HTTPs, and open source codes, i.e. HTML 5, PHP and JAVA script. This communication layer enable our system can be used in any time, any device, any platform, and anywhere.

The cloud analytics engine is based on a relational database management system. It can analysis a big data with high performance. We design the basic data analytic and data mining functions by using SQL store procedures to become an external library. In addition to that PHP family supports many advanced data mining functions, such as DBScan (dbscan.php) ,Naive Bayes (naivebayes.php), Anomaly Detection (anomaly_detection.php) etc.(Burtini, G. (2011)). They are all open source codes. It makes that the system become powerful and extendable in their analytic functions.

©copyright by Sherry Huang, Shi-Ming Huang

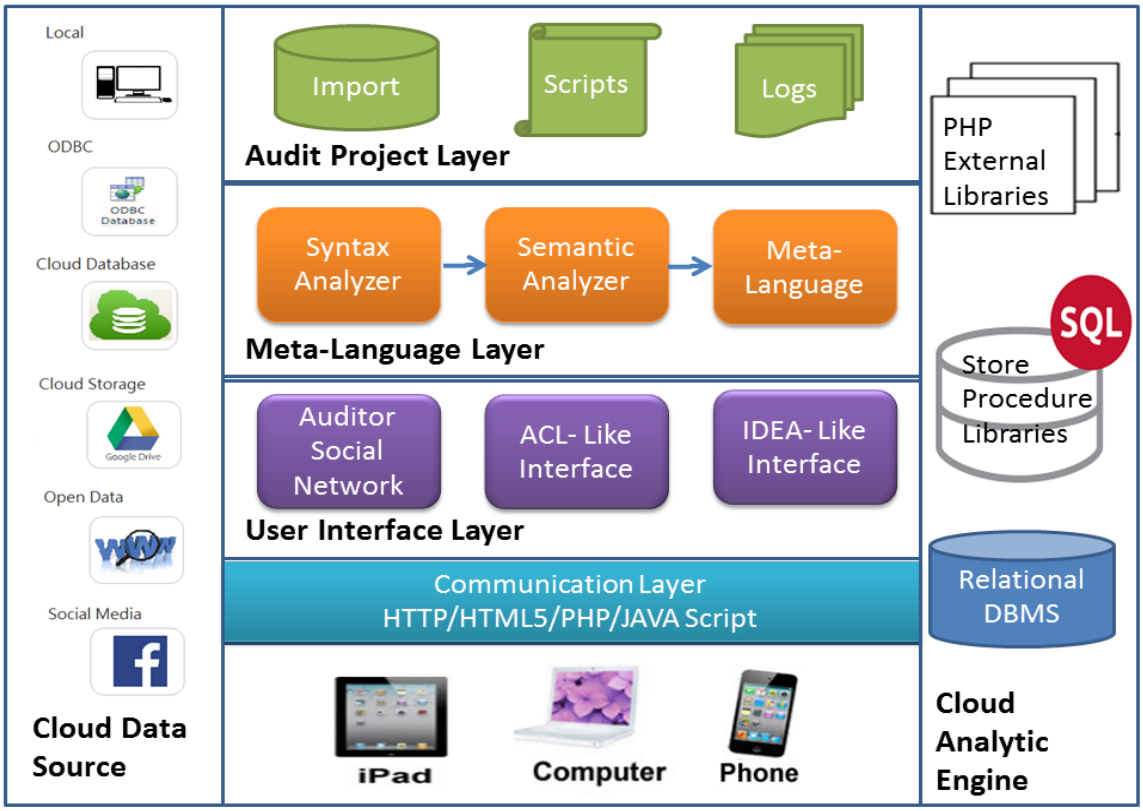


Figure 1. System Architecture of J-CAATs

4. SYSTEM DEVELOPMENT AND PROTOTYPE

To prove the concept of the idea, a prototype system is developed in WAMP environment. The acronym WAMP refers to a set of free (open source) applications, combined with Microsoft Windows, Apache web server, OpenSSL forSSL support, MS-SQL database and PHP programming language.

The prototype interface, Figure 2, is much similar to the most CAATs, such as ACL and IDEA. The major functions are on the top of menu. One side is the project navigator and the other side is the demonstration of data for read only access. User can choose to change different language in any time.

©copyright by Sherry Huang, Shi-Ming Huang

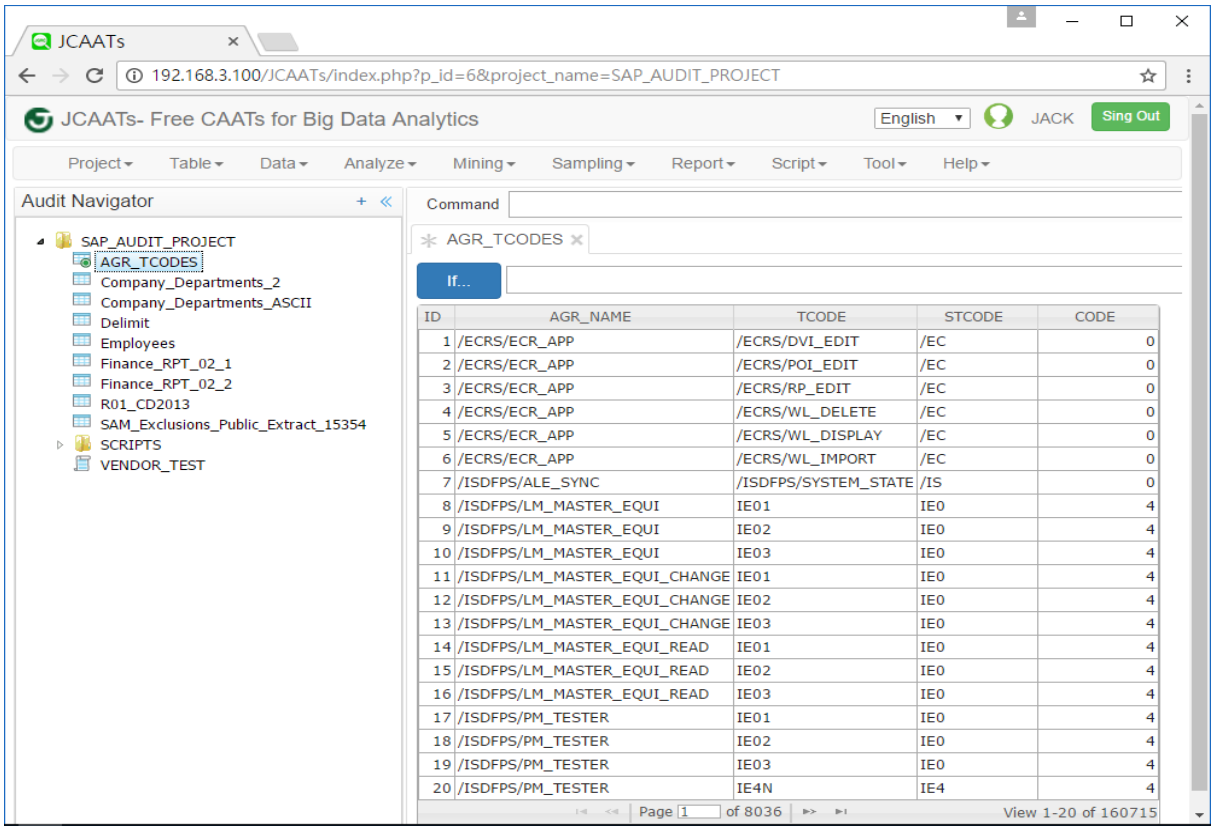


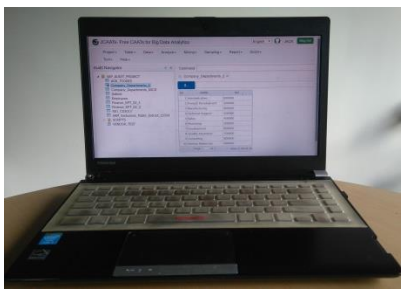
Figure 2. J-CAATs Prototype Interface

5. VERIFICATION AND VALIDATION

Three different layer testing have been applied to enable we can go deep to understand the correctness and performance of the prototype system.

5.1 Interface

To test our system can work for different platforms and devices; we use the most popular OS platforms, i.e. Window 10 for a Toshiba Note-Book, iOS for iPad Air and Android for Sony mobile phone. Figure 3 shows JCAATs interface on the different devices. Table1 shows the comparison of the OS device specific with ACL and IDEA.



Note Book



iPad Air

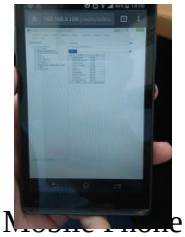


Figure 3 J-CAATs on the different devices

©copyright by Sherry Huang, Shi-Ming Huang

Table 1. Comparison of OS Specific

CAATs / Device OS Specific	J-CAATS 1.0	ACL 12	IDEA 10
Microsoft Window	Yes	Yes	Yes
iOS	Yes	No	No
Android	Yes	No	No

5.2 Data Source

There is an import wizard interface for an auditor to import data into J-CAATs. J-CAATs can process Unicode and None-Unicode data. To now a date, we have only completely tested the local source files, but not all cloud data sources or external data sources. Table 2 shows the comparison of the data preparation.

Table 2. Comparison of Data Preparation

CAATs / Local File	J-CAATS 1.0	ACL 12	IDEA 10
Access (Microsoft Office)	Yes	Yes	Yes
CSV/Delimited	Yes	Yes	Yes
DBF	No	Yes	Yes
Excel (Microsoft Office)	Yes	Yes	Yes
Fil (ACL data file)	Yes	Yes	No
Fixed length text file	Yes	Yes	Yes
ODS (Open Office spreadsheet documents)	Yes	No	No
Print-Report file	No	Yes	Yes
PDF	No	Yes	Yes
XML	Yes	Yes	Yes
XBRL	No	Yes	Yes

J-CAATs has enabled to import and manage more than 1,048,576 transaction records' file which is the max transaction for Microsoft Excel 2010 version. Our largest test file, so far, is a CISCO fire wall log for 2,037,013 records.

5.3 Analysis Feature

There are 15 major features for CAATs as the following (Wikipedia, 2017). Table 3 shows the comparison of the analysis features. J-CAATs will develop all the features in the next version.

©copyright by Sherry Huang, Shi-Ming Huang

CAATs / Analysis Features	J-CAATS 1.0	ACL 12	IDEA 10
1. **Age analysis**: making age analysis (stratification by date).	Yes	Yes	Yes
2. **Benford's law**: S finding abnormal distribution of specific digits accordingly to Benford's law.	Yes	Yes	Yes
3. **Calculated field**: adding extra calculated fields into the table/file.	Yes	Yes	Yes
4. **Drill-down (Table)**: drill-down features by zooming in (filtering) on selected rows in the table.	No	Yes	Yes
5. **Drill-down (Pivot)**: drill-down features through pivot table.	No	Yes	Yes
6. **Matching (Duplicate)**: finding matching items for a specific field in a table/file.	Yes	Yes	No
7. **Fuzzy Matching (Duplicate)**: finding matching items for a specific field using fuzzy comparison.	No	Yes	Yes
8. **Sample (Random)**: selecting a random sample of rows from the table/file (population).	No	No	No
9. **Sample (Monetary unit)**: selecting a monetary unit sample of rows from the table/field (population).	No	Yes	Yes
10. **Sequence check (Gap)**: finding (identify) gabs (in sequences) for a specific field.	Yes	Yes	Yes
11. **Sort field**: sorting (indexing) by a specific field.	No	Yes	Yes
12. **Sort multiple fields**: sorting by multiple fields.	No	Yes	Yes
13. **Statistics**: calculation and presentation of various statistics on a specific field.	Yes	Yes	Yes
14. **Stratification**: stratification on number (amount) values in specified intervals.	Yes	Yes	Yes
15. **Total row**: displaying a total row for the table/file, e.g. accumulated numerical value.	Yes	Yes	Yes

6. CONCLUSION

The development of the key portions of system took two years of part-time effort. This comparatively rapid development has been possible because we were able to handle different device interface, external libraries and data sources. In any project of this size-and in particular in the area of computing-there are lots of extensions that can be made. Some that are particularly relevant are source-level debugging and interpreting. Again, we hope to adapt

©copyright by Sherry Huang, Shi-Ming Huang

procedures already common in PHP implementations to our situation. Most PHP implementations have very good, intelligent and with some work it is feasible that these can be adapted to J-CAATs.

This article concentrates on the tools, but clearly there are other issues that are important for computer auditing. A major one isa well-formed computer auditing language. The language-that is, syntax and semantics-plus the tools are combined into what can be called the run-time environment. This environment is the typical user's view of the audit project. Users want a language that is not only flexible but also provides them with a variety of tools for performing collaboration auditing in the cloud environment.

Finally, the lab experiment and industry field study will be applied to validate the system performance and verify the usefulness of the project result. Without the real case studies, we are still far behind to make our dream to success.

References

[1] William, Messier Jr, Glover, Steven and Prawitt, Douglas, 2016, Auditing and Assurance Services: A Systematic Approach, McGraw-Hill Education, ISBN 9780077732509

[2] Jim Kaplan, 2012 Survey Report on DataAnalysis Audit Software,Auditnet, https://www.auditnet.org

[3] David Coderre, 2009, Internal Audit: Efficiency through Automation, John Wiley and Sons, ISBN 978-0-470-39242-3

[4] Natasha DeKroonand Brian Karp, 2013, An Auditor’s Guide to Data Analytics, The Institute of Internal Auditors, https://www.theiia.org/

[5] Maxwell, Andrew, Eichhorn, Michael, Seddighi, Hamid and Smith, Peter (2012) 360 degree Internal Auditing: A complimentary approach to strengthen the effectiveness of Internal Audit functions? Audit and Risk. ISSN 2048-8408

[6] Dai and Vasarhelyi, 2016, Imagineering Audit 4.0, Journal of Emerging Technologies in Accounting Volume 13, Number 1, 2016, pp. 1–15

[7] Protiviti. 2015. 2015 Internal Audit Capabilities and Needs Survey. Available at: http://www.protiviti.com/en-US/Documents/Surveys/2015-Internal-Audit-Capabilities-and-Needs-Survey-Protiviti.pdf

[8] Liu, Q., and M. A. Vasarhelyi. 2014. Big questions in AIS research: Measurement, information processing, data analysis, and reporting.Journal of Information Systems 28 (1): 1–17.

[9] Brown-Liburd, H., H. Issa, and D. Lombardi. 2015. Behavioral implications of Big Data’s impact on audit judgment and decision makingand future research directions. Accounting Horizons 29 (2): 451–468.

[10] Burtini, G. (2011). Learning Library for

©copyright by Sherry Huang, Shi-Ming Huang

PHP.http://github.com/gburtini/Learning-Library-for-PHP
[11] Wikipedia, 2017, Computer-aided audit tools, https://en.wikipedia.org/wiki/Computer-aided_audit_tools

©copyright by Sherry Huang, Shi-Ming Huang