

## Research on financial audit innovation based on big data

Yue Yu

Shanghai University of Finance and Economics, China

**Keywords:** Big data technology; Audit; Innovation

**Abstract:** With the continuous popularization of big data technology, major changes have taken place in all fields of society. For audit, the innovative application of big data related technologies shows obvious cutting edge. Compared with the traditional audit mode, the application of big data technology can achieve better results with higher coverage and accuracy, meet higher requirements with fewer resources and lower costs, and can warn and control the risks in enterprise operation early. Therefore, big data audit not only improves the reliability of audit, but also expands the scope of audit business to provide more sophisticated customer service. This paper briefly analyzes the innovative application of big data technology in the audit field. Firstly, it introduces the definition and characteristics of big data technology and audit work. Then it points out the basic concept and general function of big data audit along with the advantages, which compensate for the traditional mode, and its development opportunities in the current socio-economic environment. Finally, it discusses how to apply the big data technology in the audit field and shows the effective strategies and feasible measures of audit innovation.

### 1. Introduction

Accompanied by the rapid growth of information technology, generation and acquisition of data become easier so that the data is getting “bigger” in terms of scale and complexity. To deal with the obstacle of diversified and exponentially-growing data collection that traditional software tools cannot capture, manage and process within a certain period of time, big data technology comes into being with the characteristics of with the characteristics of massive, wide variety, high speed, and high accuracy.[1] Big data technology refers to a utility software that can extract huge, high growth and complicated information assets and analyze which could enhance productivity, profits and risk management. Big data includes four main technologies. By means of data storage, data can be kept in the cloud and can be accessed when necessary without considering space constraints. Through data mining, anomalies and correlations can be found within a massive portfolio of data, thus predicting probable outcomes and making meaningful decisions. In the way of data analysis, raw data can be turned into mature conclusions automatically by mechanical processes and algorithms. With data visualization, understandable ways to see trends, outliers and patterns in data such as charts, graphs, maps and other graphical representations are provided.

Unlike big data technology, which is an emerging product of the scientific era, financial audit has existed for many years as a guarantee of business order. Audit is an objective examination and evaluation of the financial statements of an organization by sample-based testing to make sure that the records are a fair and accurate representation of the transactions claim to represent, and the statements give a true and just view of the company’s financial position in all material respects. The typical procedure of audit includes tests of control, which is designed to evaluate the operating effectiveness of internal controls in preventing or detecting and correcting material misstatements at the assertion level, and the substantive procedures. The latter fall into two categories: analytical procedures and tests of detail. Analytical procedures tend to be appropriate for large volumes of predictable transactions like wages and salaries by a comparison to financial and non-financial data and the investigation or confirmation of identified fluctuations and relationships inconsistent with other information. Meanwhile, tests of detail on inventory, classes of transactions and disclosures can ensure the accuracy of account balances through inspecting documents and tangible assets and the

reperformance of internal control of the entity. As the economic climates are tougher and more uneven in modern times, auditing should rethink the way it is executed and accelerate the combination with advanced technology like big data to meet the higher requirements of data capture, analytics and privacy in the financial market, thus enhancing the quality and relevance of audit.[2]

## **2. The basic concept of big data audit**

To handle more complex data and business, big data audit appears which is to use the product of scientific and technological development - big data technology to process and analyze massive data to achieve the results of approving financial information and reflecting financial position.

In the era of big data, audit work should first effectively transform its objectives and gradually expand its tasks to improve the effect of audit work and give full play to the application value of big data technology.[1] In addition to finding various abnormal clues in controlling illegal acts and frauds, audit work also needs to consider the goal of identifying operational risks and promoting entity's development. It requires audit to optimize the operating system, play an active role in risk assessment, and better explore the development pattern, in order to display the significant value in decision-making analysis. Meanwhile, the transformation of the audit field also needs to pay attention to the specific audit content as audit work is faced with more complicated content. It involves not only simple data information such as previous amount and various expenses, but also more complex text information, audio information and video information.[2] If the content becomes more diverse, it will inevitably increase the difficulty of analysis, which puts forward higher requirements for the quantity and quality of data processed by big data technology. First of all, the audit needs to change the previous sample-based testing method and conduct a comprehensive audit analysis of entire populations of audit-relevant data to avoid any omission and deliver a higher quality of audit evidence and more relevant business insights.[3] Secondly, the accuracy requirements should be gradually reduced during the audit process as information value density is relatively low with the application of big data, so it needs to be optimized with the help of appropriate big data technology. Moreover, audit mindset also needs to change from the original causation to correlation, which requires the focus on exploring the relationship between different indicators to provide reference for decision-making and other work.[1]

## **3. Opportunities of big data auditing**

### **3.1 Enhancing the quality of audit work**

By convention, auditors had to sample the client's data to come up with useful insights. Although sampling has been effective for some time, it does not provide sufficient precision. Through the application of big data technology, a larger range or even entire population of data can be tested, so more accurate conclusions can be drawn. At the same time, previous auditing mainly relied on auditors to set rules to screen abnormal samples, and the audit effect is closely related to the experience level of auditors. However, machine learning technology can directly process and analyze huge amounts of audit data to form corresponding hypothesis results and risk prediction, so as to help auditors reveal the invisible data relationship.[4] This not only reduces the risk of human error, but also makes the anomaly of data be found earlier. After that, the subsequent audit process can be simplified or targeted to find effective audit evidence.

### **3.2 Improving the efficiency of audit work**

First, audit intelligence can be promoted by big data technology. Through using audit robot process automation instead of manual execution, repetitive transactional work such as external data mining, data collection, data comparison and analysis, system security inspection and audit working paper preparation can be completed in shorter time but with higher precision, which not only saves labor cost, but also makes audit resources be allocated more reasonably.[4] Secondly, a prominent feature of financial audit is sharing. Auditors can effectively realize the joint connection of relevant data that

can be synchronized beyond the limitation of time and space, and it can also be analyzed by multiple auditors simultaneously, thus greatly improving the work efficiency and the accuracy of audit. [5]

### **3.3 Strengthening the review of audit work**

Other than being costly, data collection and analysis can be quite time-consuming, especially in the absence of necessary analysis tools.[5] This is why companies choose to analyze data after each fiscal quarter or year – even though they know that frequent analysis will yield better results. This also leads to unnecessary losses caused by enterprises' failure to find problems in financial operation in time. However, by simplifying the data analysis process, big data reduces the audit lead time, greatly improves the timeliness of financial audit, and can review the financial operation status of enterprises anytime and anywhere at a lower cost. This continuous review not only effectively avoids the potential risks in operation,[5] but also paves the way for accurate control assessments as well as timely insights.

## **4. Application of big data technology in audit field**

At present, the concept of big data audit is constantly infiltrating, and some big data technologies have been proved to be effective in the audit field. The following describes the concepts and basic operation modes of five big data audit technologies.

### **4.1 Data mining analysis**

Data mining analysis is particularly important in the application of big data audit. It refers to the process of extracting new information that people are interested in, previously unknown and potentially valuable for decision-making from large databases. With the economic and informational development, many enterprises introduce artificial intelligence and information systems for management, such as enterprise resource planning system (ERP) and supply chain management system (SCM), forming a data warehouse containing a large amount of accounting and management information, which makes the use of data mining technology necessary.[6]

In the application process of data mining technology in audit, auditors should first understand the internal control and external environment of the audited enterprise, and determine a clear direction and goal for data mining.

Secondly, the data should be selected and sorted out. As the data may be random, incomplete and have complex data structure, it is necessary to preliminarily do the data pre-processing.

Then, a corresponding model needs to be established for data mining, design or select effective algorithms and revise repeatedly. For example, apply association rule analysis technology. Namely, by analyzing a given set of records, the correlation is derived to mine the relationship between data. It can be divided into simple correlation, temporal correlation, causal correlation and so on.[6] In view of the possible corresponding relationship between accounting subjects and data items, as well as the cross-checking relationship between the three accounting statements including balance, income and cash flow statement, auditors can use association rule analysis technology to mine the relationship between the data, thus finding data with abnormal connections and determining whether the sales revenue and cost of various products are reasonable or not. Another example is the application of cluster analysis technology which finds similar data from the data set and forms different clusters.[6] Minimize the distance of objects in the same cluster and reduce the similarity of different clusters, which can customize different types of data sets according to different needs. From this, auditors can easily identify dense and sparse areas and find the relationship between global distribution patterns and data attributes. The data in the enterprise's accounting statements will change accordingly with the change of operation, and it also has a certain regularity. If the data change is abnormal, it indicates that there may be false elements in the accounting statements. Like the substantive test of receivables, if group the data with similar characteristics, the receivable that is significantly different from other periods, with different amounts in the general ledger and sub ledger, generated around reporting date, with repeated bookkeeping and unconventional transactions can be found and listed as audit risks. Further example is the application of outlier analysis technology which can check deviation and

extreme value in the audit practice, and then disclose fraud and violation of laws.[6] As the data generated by these behaviors are generally a small amount of data hidden under huge normal ones. These abnormal or special information—outliers needs more attention from the audit work. Auditors can choose indicators that can reflect the financial situation, such as asset liability ratio, gross profit margin on sales, net profit margin on sales, current ratio, quick ratio, etc.[6] The ratio which is significantly different from other months and years may imply potential fraud. Also, when auditing the production cost of the enterprise's products, the isolated point analysis can be used to separate the cost data that is greatly different from the expected production technical standards.

#### **4.2 Real-Time risk warning**

In the past, the principle of follow-up audit was often adopted which only finds problems and gives feedback to some clues of violations.[1] Now it often reflects the characteristics of sustainability and forms a dynamic audit model. It can dynamically analyze the constantly enriched and updated data to continuously monitor and grasp the changing status of audit objectives, timely feedback abnormal problems and remind relevant personnel to take appropriate measures to avoid major accidents and economic losses through the characteristics of early admission, early warning, prediction and flexibility.

Generally, dynamic audit mainly includes four tasks: reviewing whether the income statement of the target unit is accurate, checking whether the target unit strictly implements the financial policies, verifying whether the target unit obtains the due economic benefits within the business scope, and whether the dynamic data of the target unit reflects fraud.[7]

Risk information collection is the primary segment of the early warning system of dynamic audit in the context of big data. By integrating the information systems of various regulatory objects, including banking audit information system, cross domain audit information system, insurance and securities audit information system, a new financial risk information supervision database is established. The information supervision database contains distributed secondary databases such as financial information databases of financial institutions, financial supervision information database, macroeconomic database, historical loss database and historical risk database.[7] The information collected from each upstream resource information system is extracted by big data collection technologies such as ETL, MySQL or oracle.[8] It is stored in the corresponding secondary database for later analysis. On this basis, the dynamic audit early warning system uses big data analysis technologies such as KDD, visual analysis and SQL query to further mine the relevant risk information to determine the risk index parameters, find the risk clues, generate the risk index and issue early warning reminders.[8] After that, it will be handed over to the professional risk analyst to define and verify the financial risk information with the determined category and hazard level and submit the final risk early warning results for release to prevent further evolution of financial risk to financial crisis.[9] For example, yellow warning can be given to financial risks close to critical value and low risk, which can be avoided through government intervention. For financial risks exceeding the safety threshold and with high risk, red warning can be carried out. In addition, in the risk identification and judgment stage, the data analysis results can also be transferred to big data adaptive learning systems and historical data analysis systems.[7]

#### **4.3 Multi-Domain data fusion**

To maximize the application value of big data technology in the audit field, attention need to be paid to the extensive collection and sorting of multi-domain data to avoid the impact of data island.[1] Audit often involves the cross analysis and application of various databases. It is necessary to ensure that it has cross database analysis ability and can use appropriate and reasonable analysis tools to better identify possible abnormal problems.

To achieve the informatization of audit and complete the audit work more effectively when processing complicated data, Multi-source Information Fusion Technology (MSIF) is needed which scientifically detect, correlate and combine the data required by the audit information resource system

based on the audit objectives and project management characteristics, and finally form new information characteristics for analysis and utilization.[10]

The specific workflow of MSIF participating in the audit is to classify industries according to different standards. Each industry can be used as an overall project as the starting point of analysis. Then, the embedded relationship between sub-databases is formed, which lays the foundation for the improvement of the general database. The information distribution of the general database penetrates each management link of the sub-database project, forming a multi-source sub-database project information group.[11] In terms of audit subject and audit timeline, each audit team and each period of audit have become the main sensor to obtain sub-projects. Other assurance business or non-assurance business and temporary business can only play the function of assisting in obtaining information. This process is equivalent to the collection of sub-project information in each sub-database. The information is continuously obtained through various types of sensing organs, processed and collected into the general database system. As each sensor has its own application range and accuracy, relying on a single one cannot meet the requirements of accurately understanding the unknown environment and research object.[11] Therefore, using a multi-sensor system and multi-source information fusion technology will improve the reliability and accuracy of understanding things and the environment. The existing data fusion platform can be used as a reference for the whole process of audit, such as the C3I information processing model adopted by the military, blackboard model and agent model.[10] Taking the agent model as an example, the multi-agent and hierarchical information fusion model in the multi-agent information fusion technology can be used as the processing framework of the whole process audit. Each information segment is regarded as an agent in the system, which provides data and knowledge from different sources for the fusion center, and the audit department can conduct audit inspection and analysis according to the fused data.

The application of MSIF in audit can be divided into three levels. The first level is during the audit process, for the same audit objectives like annual report audit, the audit team collects and processes the sub-project information when auditing each industry and sub-industry every year to obtain the information that has practical effect on the future audit. The second level requires the interconnection, correlation and information combination of the above audit sub-project information, and the information collected from all sub-projects shall be classified, combined and processed to form an audit information resource database related to the overall project. At the third level, the information after fusion forms the new information characteristics of the overall project. This development from single information value to integrated information value realizes the follow-up guidance value function of information. [11]

#### **4.4 Audit Software Development**

Now big data audit software has played a great role in the audit field. IDEA, which is an interactive data extraction and analysis software. It can collect various types of data files, record operation tracks and index extraction, and the data cannot be changed. To a great extent, it ensures data security and prevents fraud. ACL, which can process tens of millions of data and provide a flexible re-editing command program, thus enabling users to design the most suitable data analysis program according to their work experience or business needs. TeamMate, an audit project management software that provides services such as risk assessment, audit plan, review of working papers, etc. ZhongPu cloud set, which supports three-level penetration queries: table, account and certificate. It can also perform base setting, month on month trend, structure and other comparative analysis and graphical display of the occurrence of data in each period, and realize the trend analysis of loan balance and make two-dimensional cross display. More importantly, it can perform aging batch intelligent analysis on all transaction accounts, directly send letters or generate alternative tests. Also, set batch intelligent settings for related parties that will be automatically associated with audit drafts, statements and notes. ECPA is mainly used for the preparation of drafts in the audit process. It can automatically generate more than 200 manuscripts according to the electronic data of the enterprise in 10 seconds, spot check the vouchers from the voucher data, and establish a cross-index between various manuscripts. Auditors

can also program, control and customize the contents of templates and drafts through Evb macro language.

#### **4.5 Cloud audit**

When big data technologies apply in other field, it is often closely related to cloud computing, which should be relied on cloud storage and virtualization processing to promote the efficient utilization of massive data information and solve the problem of data information.[1] For example, the comprehensive analysis of cross-database data information can be optimized by using the cloud audit platform. The overall analysis and processing efficiency is high, which can effectively meet the development trend of increasing audit difficulty. The application of cloud audit mode can also realize the remote storage of data and analysis of information, which significantly improves the convenience of audit work, and makes the data sharing effect of relevant organizations stronger.[1]

At present, the software related to cloud audit includes Smartbi, which focuses on cross database query, high-performance storage, doubt generation, self-service analysis, etc. Its audit model can be shared and edited at the same time. The data navigation can open the data resource directory for analysis to business personnel, and carry out data quick query and desensitization preview under the control of authority. Through the virtual semantic layer, the obscurity of the underlying physical model is shielded, and the complex data relationships are transformed into models that can be recognized and used by business analysts. ACE Smart is an enterprise network architecture based on cloud technology. The cloud is updated in real time to support the two-way data interaction between the on-site person in charge and the preparer to realize interconnection. ShenYou, which enables auditors conduct off-site audits in advance through the query and analysis system, screen risk points in advance and determine the audit focus. In case of failure to go to the on-site audit, the auditee can provide the information required for the audit online through the system, which can not only accumulate the information of the auditee, but also leave traces in the information system and reduce the audit risk.

#### **5. Conclusion**

Ultimately, the introduction and application of big data technology has become an important development trend of innovation in the audit field. With the combination of big data technology, audit has obvious advantages and more prominent functions. Auditors will be able to use larger data sets and analytics to better understand the business, improve work efficiency and accuracy through automation and joint sharing, identify key risk areas with less cost and higher frequency, and deliver enhanced audit quality and coverage while providing more business value. Therefore, it is necessary to innovate and optimize the audit concept, organizational structure and specific technologies from multiple angles to create good conditions for the integration of big data technology in the audit field. This paper mainly discusses the transformation from traditional audit work mode to big data audit and its specific application. Through the introduction of data mining analysis, real-time risk warning, multi-domain data fusion, audit software and cloud audit, big data audit not only maximizes the audit value but also further deepens the significance of audit, from the guarantor of enterprise value to the helper of enterprise value creation. Arguably, with the development of global economic integration, multi-directional and multi-field data capture will make the audit work more complex and demanding, thus big data audit will be more standardized to provide better reference for decision-making.

#### **References**

- [1] Zhao, D. (2021). Research on the application of big data technology in audit field. *Modern Business*, 26, 163-165.
- [2] Ramlukan, R. (2015). How Big Data and Analytics Are Transforming the Audit. *Financial Executive*, 31(3 & 4), 14–19.

- [3] Tang, Jenna, and Khondckar E. Karim. "BIG DATA in Bussiness Analytics-Implications for the Audit Profession." THE CPA Journal, 2017, 34–39.
- [4] Zhang, Q.L., He, J.N. & Rui, P.S.. (2021). The road of internal audit innovation in the new era: from data audit to intelligent audit. Finance and Accounting Monthly (22), 78-83.
- [5] Chen, P., 2021. Research on the effectiveness of financial audit innovation under the background of big data. Accounting of Township Enterprises in China, 10, 127-128.
- [6] Cao, H.(2013). Application of data mining technology in enterprise audit. Financial Management, 04, 89-90.
- [7] Shi, X. F., and Gao, C. J. (2019). Construction and Test of Dynamic Audit Early Warning System under the Background of Big Data. Friends of Accounting, 06, 31–37.
- [8] Zhang, H. L. (2021). Construction of financial risk dynamic audit early warning system based on big data. Communication of Finance and Accounting, 13,123-127.
- [9] Zhang, X. K. (2021). Probe into the construction and implementation mechanism of dynamic audit early warning system in state-owned enterprises. Accounting of Township Enterprises in China, 10, 121-122.
- [10] Zhang, Y. (2014). Application of multi-source information fusion technology in the whole process audit of construction project. China Management Informatization, 12, 15-17.
- [11] Wang, J. F. (2010). Application of multi-source information fusion technology in audit. Finance and accounting, 12, 63-64.