INTOSAI
Working Group on Big Data

DEVELOPMENT OVERVIEW OF BIG DATA AUDITS PERFORMED BY SUPREME AUDIT INSTITUTIONS FROM 2016 TO 2021

September
2022

FOREWORD

In this information age, data is essential for ensuring good governance. That is why big data auditing has now become an emerging approach for supreme audit institutions (SAIs) to data collection, audit analysis, and risk detection.

We are pleased to present to the INTOSAI community this report on the Overview of Big Data Audits Performed by SAIs from 2016 to 2021. The report presents the findings of a research on big data audits conducted by members of the INTOSAI Working Group on Big Data, explores a framework for big data audit application, and taps into the experiences and practices of SAIs over the past five years, trying to understand the development trend of big data auditing. The report carries out a comparative study of SAIs' practices in technology application, data governance and fulfillment of objectives through surveys. The report also puts forward recommendations for SAIs to perform big data auditing. We hope this report will promote the use of big data in public sector auditing.

We would like to express our appreciation to the SAIs of Austria, Brazil, China, Denmark, Ecuador, Estonia, Finland, France, India, Indonesia, the Netherlands, Norway, Peru, the Philippines, Russia, Türkiye, the UK, the U.S., and the European Court of Auditors for sharing their practices in big data auditing, which are the basis for this report. We would like to thank the Secretariat of the INTOSAI Working Group on Big Data for compiling this report. We are also grateful for the valuable suggestions on the design of the questionnaire from the SAIs of Bhutan, Denmark, Ecuador, Eritrea, Estonia, Finland, France, India, Indonesia, Ireland, Mexico, the Netherlands, Peru, the Philippines, Portugal, Thailand, Türkiye, Ukraine, the UK, Vietnam, and AFROSAI-E.

We hope you find this report useful.

Hou Kai

Auditor General of the National Audit Office of China
Chair of INTOSAI Working Group on Big Data

ABOUT INTOSAI
Working Group on Big Data

INTOSAI Working Group on Big Data (WGBD) is a specialized working group approved by INTOSAI under Strategic Goal Three: Knowledge Sharing and Services. Its objectives are to identify the challenges and opportunities faced by SAIs in the era of big data, to summarize the knowledge and experience in the field of big data audit, and to strengthen bilateral and multilateral technical cooperation on big data.

Chair

China

Vice Chair

United States

Members

- Argentina
- Austria
- Bangladesh
- Belgium
- Bhutan
- Brazil
- Canada
- Chile
- Denmark
- Ecuador
- Estonia
- Finland
- Fiji
- France
- Georgia
- India
- Indonesia
- Kuwait
- Malaysia
- Mexico
- Netherlands
- New Zealand
- Norway
- Pakistan
- Peru
- Philippines
- Portugal
- Republic of Korea
- Russian Federation
- Samoa
- Senegal
- Slovakia
- Thailand
- Türkiye
- Ukraine
- United Kingdom
- Zambia

Observers

- AFROSAI–E
- Bulgaria
- European Court of Auditors
- Ireland
- Vietnam

CONTENTS

ABSTRACT

As the world shifts gear towards the information age, big data has played a pivotal role in promoting modernization of the national governance system and governance capacity, and the audit patterns of countries have also changed accordingly. Big data audit has increasingly risen as a major approach for Supreme Audit Institutions (SAIs) to obtain audit data, perform audit analysis, identify suspected violations and issue risk warning. This overview summarizes the research findings of big data audit theory of INTOSAI WGBD members; explores the establishment of a big data audit application framework based on the dimensions of technology level, governance level, and objective level; taps into the advanced practices of SAIs around the world over the past 5 years through close liaison between the WGBD and INTOSAI; explores and summarizes the development trend of big data audit; summarizes and compares the realities of SAIs in terms of technology application, data governance and fulfillment of objectives by means of questionnaire investigation, with a focus on the analysis and research of drawbacks. Through review and outlook, evaluation and feedback, comparison and reference of the big data audit work of various countries, this overview puts forward recommendations for SAIs in performing big data audit, aiming to play an active role of big data audit in the government auditing around the world.

THEORETICAL RESEARCH AND APPLICATION FRAMEWORK OF BIG DATA AUDIT

Big data features massive scale, rapid transmission, diverse sources, complex structure and inestimable value (Connelly et al., 2016)[1]. The advent of the information age and technological revolution has impelled big data to emerge as a technology resource for many countries to enhance public governance by the government. The core content of big data audit lies in effective use of the tremendous value contained in big data, analysis of data patterns, identification of abnormal data, extraction of useful information, and establishment of data visualization solutions (AICPA, 2014)[2]. SAIs collect and make full use of the massive data resources in accordance with the authority of law, and perform intelligent data mining and analysis of data that is cross-domain, cross-level, cross-sector and cross-system at all dimensions, which are effective measures for auditors to identify suspected violations, and make evaluation, conclusion and macro analysis on relevant issues (Hu Zejun, 2019)[3]. The establishment of a big data audit work framework will promote the shift of audit work priorities from improving the efficiency of public revenues and expenditures to enhancing information transparency, promoting public reputation, preventing and resolving major risks[4].

Compared with traditional audit, big data audit has registered remarkable innovations in three aspects: techniques and methods, application fields, and organization patterns.

---

1. Connelly R, Christopher J. Playford, V G, et al. The role of administrative data in the big data revolution in social science research, Social Science Research, 2016, 59: 1-12.

2. AICPA. Reimagining Auditing in a Wired World (White Paper). 2014. New York: American Institute of Certified Public Accountants.

3. Hu Zejun. 2019, National Auditing of the People's Republic of China. Beijing: China Times Press.

4. INTOSAI Strategic Plan 2017-2022

In terms of technological innovation, the audit analysis approach has evolved from sampling audit to full data analysis, analysis method from traditional inquiry statistics to big data technologies (Liu Xing et al., 2016; Alles et al., 2016)[5], and audit evidence collection from single data extraction and analysis techniques to the establishment of big data platforms and the integration of big data value chains (Niu Yanfang et al., 2017)[6].

In terms of the application domain, big data is not only applied to public finance performance auditing, but also radiates to the whole process from distribution, management, to use of public funds, state-owned assets and state-owned resources (Qian Hongdao et al., 2019)[7], realizing full coverage of big data audit (Ma Zhijuan, 2015)[8].

In terms of the organization pattern, the previous decentralized audit teams have been developed into an integrated supervision system featuring by its standard of process, professional capacity, network support and diversified oversight. The system makes full use of and consolidates data from different organizations and authorities, integrates big data technology with audit capacity (Zheng Wei et al., 2016)[9], formulates value goals and codes of conduct shared by auditors (Li Chengai et al., 2019)[10], and transforms the organization pattern in the context of big data (Xu Zongben et al., 2014)[11].

In view of the application of big data audit techniques, establishment of legal system and development of organization pattern, progresses are different among various SAIs (Giest, 2017)[12], and the high initial investment and technical threshold are among the main challenges for implementing big data audits (Omitogun et al., 2019)[13]. For the purpose of effective performance of audit work, identifying and eliminating risks and barriers, it is necessary to

---

5. Liu Xing, Niu Yanfang, Tang Zhihao. Reflections on Promoting Big Data Audit Development. Auditing Research, 2016(05): 3-7; Alles M, Gray G L. Incorporating big data in audits: Identifying inhibitors and a research agenda to address those inhibitors. International Journal of Accounting Information, 2016, 22, 44-59.

6. Niu Yanfang, Feng Zhanguo, Meng Xiangyu. Innovation and Practice of Audit Work from the Perspective of Big Data Value Chain. Auditing Research, 2017(5), 17-22.

7. Qian Hongdao, Xie Tianyu. The Amendment Trend and Its Logic of Audit Law under the Perspective of Full Coverage of Audit. Journal of Audit & Economics, 2019, 34(03): 22-31.

8. Ma Zhijuan, Liang Siyuan. Study on Full Coverage Audit Implementation Path of Government Environmental Accountability Audit in the Era of Big Data. Auditing Research, 2015, 000(005): 28-34.

9. Zheng Wei, Zhang Limin, Yang Li. Research on Data-oriented Audit Mode in a Big Data Environment. Auditing Research, 2016(04): 20-27.

10. Li Chengai, He Xiaobao. Exploration and Innovation of Big Data Audit Organization Modes. Auditing Research, 2019(05): 23-29.

11. Xu Zongben, Feng Zhiyan, Guo Xunhua, Zeng Dajun, Chen Guoqing. Frontier Topics in Management and Decision-making Driven by Big Data. Management World, 2014(11):158-163.

12. Giest S. Big data for policymaking: fad or fasttrack? Policy Sciences, 2017, 50(3):1-16.

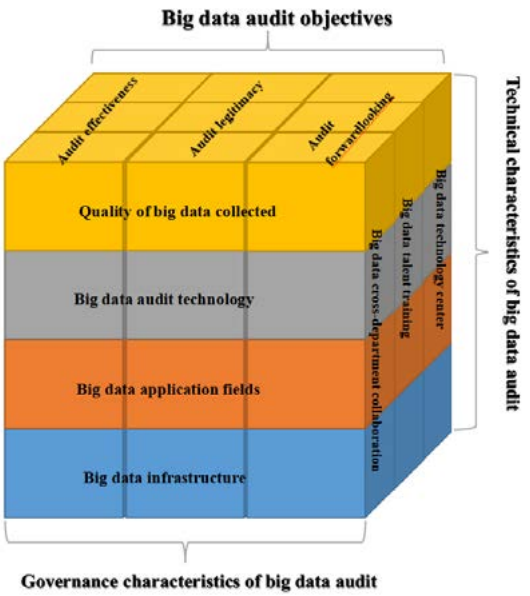
analyze and evaluate big data audits by taking the organization as a whole, and propose countermeasures by tapping into the deep-rooted causes.

Based on the research of scholars such as Appelbaum et al. (2017)[14] and Al-Sai et al. (2020)[15], big data is applied at three levels - governance environment, big data technologies and performance objectives. The effectiveness of audit objectives is jointly promoted by the factors of organization, technology and governance. In line with this theory, the framework of big data audit applications is divided into three levels - technology level, governance level and objective level (see Figure 1).

The indicators at technology level are designed to evaluate the big data technological innovation and implementation as well as management of audit institutions. They are measured from four dimensions: the data collection quality being more efficient, the techniques and methods (such as data modeling and algorithms) being more diverse, the audit field being more comprehensive, and the infrastructure (such as the big data audit platform module) being more capable.

The indicators at governance level are designed to evaluate the matching degree of organization, functional structure and big data audit techniques. They are measured from three dimensions: the establishment of big data technology centers, the training of big data talents, and full collaboration across departments and regions.

Figure 1 the framework of big data audit applications



The indicators at objective level are designed to evaluate whether the effectiveness of big data

13. Omitogun, A. and Al-Adeem, K. 2019, Auditors' Perceptions of and Competencies in Big Data and Data Analytics: An Empirical Investigation, International Journal of Computer Auditing, 1 (1): 92–114.

14. Appelbaum D, Kogan A, Vasarhelyi M A. Big Data and Analytics in the Modern Audit Engagement: Research Needs. Auditing A Journal of Practice & Theory, 2017, 36(4):1-27

15. Al-Sai Z, Abdullah R, Husin H. Critical Success Factors for Big Data: A Systematic Literature Review. IEEE Access, 2020:1-1.

meets the expected objectives and the post-implementation impact. They are measured from three dimensions: the efficiency of identifying suspected violations (effectiveness), the social influence promoted by enhancing information transparency (legitimacy), and the accuracy of risk prevention (forward-looking).

With the objective level as the orientation, the three levels interact and influence each other, and the technology level and the governance level are interdependent and mutually promoting. In practice, big data audit techniques are constantly innovating, organizational governance mechanisms continue to improve, and big data technologies and organization governance systems complement each other in harmony, together they facilitate the realization of audit objectives of SAIs.

BIG DATA AUDITING APPLICATIONS - BASED ON CASE PRACTICE

The INTOSAI Working Group on Big Data (WGBD) is a specialized group approved by the International Organization of Supreme Audit Institutions (INTOSAI) in 2016 aimed at achieving the strategic goal 3 of "Knowledge Sharing and Knowledge Services". With member and observers in over 40 countries and regions around the world, the working group is responsible for identifying the challenges and opportunities toward SAIs in the era of big data, summarizing knowledge and experience in the field of big data audit, and strengthening bilateral and multilateral technological cooperation. So far, the WGBD has held 5 annual meetings, on which representatives from various countries introduced the development and practices of big data audit in their respective countries. These meetings demonstrated that INTOSAI WGBD members have attached great importance to big data audit work by enhancing the application of diversified big data technologies in new scenarios, practicing big data audit in inter-department collaborations, reinforcing data sharing as well as enacting relevant regulations, in a way to achieve full coverage of big data audits. Specific measures are as follows:

1. Applications of big data audit techniques and methodologies

- **Innovation in Big data technology and audit methodology**

With the development and wide application of big data technology, it has evolved markedly beyond basic comparative analysis and trend analysis, from text recognition and data

association in the early stage towards cutting-edge technologies such as complex machine learning, AI (artificial intelligence) and block-chain. As can be seen from the cases shared by SAIs over the past 5 years (Figure 2), big data audit methods have evolved from the simple use of data mining and visualization technology to the use of diverse and relatively

complex model calculation methods. For example, the Government Accountability Office (GAO) makes extensive use of programming languages such as Python and R language, statistical analysis software such as SAS and Stata, visualization software such as Tableau, and GIS (geographic information system) software and block-chain such as ArcGIS and MapInfo in performing big data audit activities.

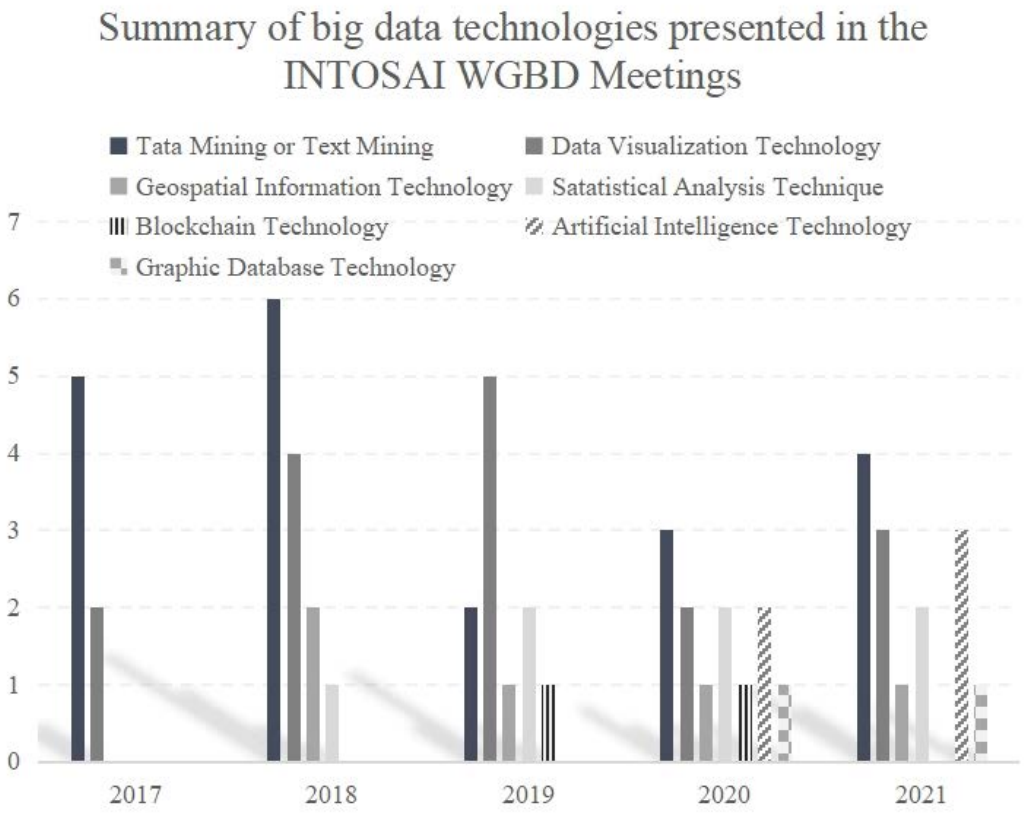


Figure 2 Big data technologies presented in the INTOSAI WGBD Meetings

- **Extension of big data audit coverage and breakthroughs in the application fields**

In traditional audit, data audit techniques are mostly applied in audit fields such as finance and taxation. It can be seen from the cases shared by SAIs that audit of public finance represents the

mainstream of big data audit work. Meanwhile, with different priorities of audit performance, SAIs have conducted audit on health care, criminal investigation, education, environmental protection, and disaster prediction. For example, SAI China shared resource-environment big data audit cases as early as 2017. From 2018 to 2020, SAI Denmark, France and GAO carried out fruitful explorations in the education field. In 2020, SAI Estonia introduced successful experience in emergency rescue and risk prediction. In 2021, the European Court of Auditors made innovations in multilingual news tracking and real-time audit of the implementation of the European Common Agricultural Policy.

2. The establishment and improvement of environment and hardware for big data audit

The establishment and improvement of policy environment for big data audit

Institutional arrangements are, in many countries, the prerequisite for promoting the sharing and opening of data resources, improving the value of data use, and building a big data audit pattern. In recent years, countries include India, China, the United States, Norway,Ecuador have promulgated policies on the mandate and collection standard of big data audit. To list a few examples, the United States issued the Digital Accountability and Transparency Act, which aims to promote the digitization of government work and establish unified government data standards; the Ecuadorian government signed the Transparency and Work Information Act and the National Public Data Registration Act, which authorize the auditors to visit the national information system portal and obtain information from the Ministry of Labor, property registration authorities and social security authorities, and guarantee for making full use of data in performing audit work.

The management and control of audit data quality and security

Audit data quality and data security are vital to ensure the normal operation of big data audit. Improving data quality and enhancing data security can help to expand audit scope, improve analysis quality and prevent internal control issues. Over the years, SAI Philippines, Mexico and the European Court of Auditors, have committed to optimizing the data operation environment and protecting data security. They introduced their principles for data security management and new data security management models in accordance with the big data procedures and the characteristics of their own projects. SAI Mexico established an evaluation index system to measure data quality from three dimensions of data source, data type and data characteristics, aiming to strengthen the overall quality management of big data.

The establishment and improvement of big data audit infrastructure

To carry out intensive and systematic big data audit on a larger scale, SAIs in some countries have built an audit platform that incorporates the entire process of data collection, preparation, storage, analysis and presentation, which enable the integrated management of auditors, audit procedures and technical means in one system. A total of 9 audit institutions have introduced the establishment of their big data audit platforms over the past 5 meetings. For example, SAI Indonesia and Norway introduced their audit analysis platforms and visualization platforms in 2017; SAI China, Russia and Brazil briefed the establishment of their audit platforms in 2018; SAI Estonia and SAI Turkey presented their data platforms named X-Road and VERA respectively in 2019; and SAI Philippines introduced its MIKA-EL (audit experience learning based on machine intelligence) intelligent platform which is capable of automatic identification of unusual transactions in 2020. Moreover, in 2021, the European Court of Auditors presented its infrastructure service platform, which incorporate the functions of data collection, data warehousing, and automated analysis for local authorized users of EU members.

3. The transformation and revolution of the organization patterns of big data audit

The establishment and training of big data audit talents

SAIs in some countries have launched various talent training programs targeted at improving data analysis and technology application capacity of auditors. Statistics show the main training modes adopted by SAIs include: centralized training (SAI UK), collaborative audit by technical expert group and audit group (GAO), external experts and professional data analysts joining audit assignments (SAI Finland), online technical training and regular technical discussion (SAI Austria), and the "experience mentoring" between new and senior auditors (SAI Denmark). In order to facilitate centralized management and better leverage the core technological advantages of talents, SAI China, India, Netherlands, U.K., as well as GAO and the European Court of Auditors, have established ad-hoc data centers or talent centers. For example, SAI China established the Department of Electronic Data Audit in 2014 to take charge of data management and analysis; GAO established an Innovation Lab in 2019 to tackle significant oversight challenges that exist across the government through explorations of data science and emerging technologies.

Integration and collaboration of big data audit resources

Many countries strive to unify collected data as a whole and make proper categorization, so

as to facilitate cross-department, cross-sector and cross-region cooperation. For example, the European Court of Auditors strives to maintain regular contacts with member states and institutions within the European Union for timely data update and sharing, and actively cooperates with academia and professional institutions in performing big data audit; and in the audit of Chesapeake Bay cleanup case, GAO paid more attention to state cooperation on big data[16]. SAI Estonia, cooperates closely with statistical department to leverage the latter' s rich experience in data processing for data verification.

- **The transformation of organization mode of big data audits**

Big data reinforces the coordination between audit team and relevant departments and organizations, which is reflected by the departmental collaborations of auditors and integration of audit data. The organization mode of audit project has also changed from the one with fixed members to network-supported and multiple-sourced manpower. Such transformation played an important role in facilitating the rapid adjustment of audit team in the current pandemic and reducing the inconveniences caused by remote working. For example, SAI Turkey relied more on the electrical documents and relevant data of their auditees to reduce the needs for mobility; and SAI Austria made quick adaption by providing more technical personnel to the audit team in a timely manner.

4. Objectives and approaches for improving the quality and efficiency of big data audit

The above-mentioned innovation of big data audit techniques and methods, the improvement of software and hardware infrastructure, and the transformation of organization governance patterns, are all serving the objectives of big data audit. For many SAIs, the innovative application of big data technologies and the transformation and development of organization governance patterns are led by their audit objectives respectively, at the same time, these objectives are also subject to the improvement of audit capacity and new conditions.

Big data has effectively improved the quality and efficiency of audit fulfillment through the following three approaches, first, it improved the ability and efficiency of revealing suspected

16. On account of the long-existing environmental issues such as eutrophication, low oxygen in waters, loss of ecological functions, and reduced fishery production in Chesapeake Bay, the US Federal Government and the state governments in the basin area established the Chesapeake Bay Program for cross-region governance in 1983. As water quality governance involves the interests of multiple stakeholders, a coordinating sector (Chesapeake Bay Program in this case) is required to take the lead to evaluate and improve the potential for cooperation between government authorities, with the active involvement of state representatives and a review of compliance performed on the basis of text analysis.

violations. For example, SAI Ukraine has developed a visualized fiscal expenditure data system that is open to the public and helped auditors to identify extreme values of medical purchase price in some regions, expanding the scope and effectiveness of audit. Second, it enhanced transparency and attracted more public attention to the audit work. Growing public awareness of audit institutions have further advanced their social influence and credibility. For example, SAI Austria released more audit findings in through interactive interfaces to facilitate public understanding; the digital version of GAGAS (Generally Accepted Government Audit Standards) designed by GAO with interactive functions and user behavior analysis has attracted more social

CURRENT DEVELOPMENT AND CHALLENGES OF BIG DATA AUDIT-AN ANALYSIS ON QUESTIONNAIRE RESULTS.

attention. Third, it improved the accuracy of risk warning issued by audits. For example, GAO and SAI Estonia applied big data technology in natural disaster prediction and gained accurate estimated results; and experts from the non-governmental organization-DataKind, suggested to establish a fire prediction model through analyzing location data and the relevant public information and distribute fire alarms timely and effectively based on the analyzed results.

Although SAIs have engaged in extensive practices in the field of big data audit, the development of countries (regions) have not yet struck a balance. For the purpose of a comprehensive and in-depth summary of big data audits performed by SAIs, especially understanding the existing challenges and difficulties, the INTOSAI WGBD Secretariat prepared a questionnaire with 44 questions centering on three levels – technology level, governance level and objective level, in a way to summarize the experiences of the big data audits performed by SAIs.

As of late September 2021, SAIs of 22 countries and regions (15 of which are developing countries) responded the questionnaire, including 8 countries in Asia (Vietnam, Thailand, Philippines, Estonia, Turkey, India, Bhutan, Indonesia), 8 countries in Europe (France, the United Kingdom, Finland, Norway, Ireland, Ukraine, Portugal, Denmark), 3 countries in America (Mexico, Peru, Ecuador), and 3 countries/organizations in Africa (AFROSAI-E, Eritrea and Sudan). Most of the SAIs are equipped with due independence mandated by law (17/22), administered under at least two levels of organizational structure from central to local audit institutions (12/22),

and most of SAIs are focusing on financial audit, and supplemented by performance audit and compliance audit. (13/22).

The feedback shows the respondent SAIs have engaged in years of extensive practice in the field of big data audit. More than 50% of the SAIs are able to collect data from multiple sources in accordance with the law, with more flexibility in applying big data analysis technology, and most of them have applied more than 3 kinds of big data technologies. SAIs have different priorities in big data audit, more than 50% of which focus on enterprise audit, financial subsidy audit, and government budget audit. Countries such as Sudan, Vietnam, and Ireland also applied big data audit techniques in resource, environment, finance, education, health insurance and social security audits.

While progressing by leaps and bounds, countries also have witnessed difficulties and problems such as inadequate big data technologies, lack of professional talents, weak cross-department collaboration capacity and outdated management.

1. Low accuracy of collected data and extensive data delay

Due to the wide variety and diverse sources of data, variation of data structure, and limited technological means, 41% of the SAIs noted that the quality of collected data is unsatisfactory, 45% regard the accuracy of the collected data as average or poor; 36% rely on audit assignments for data collection rather than a regular collection mechanism; 45% are concerned about the timeliness of data collection.

Reasons for the low accuracy of data and data delay lie in various data sources, types and format. The sources of data include financial data (28%), business data (24%), management information data (19%), network data and others (29%). The data format includes spreadsheet (30%), text (29%), image (16%), spatial geography data and others (25%). Half of the SAIs responded that the audit data standardization in their work are far from satisfactory due to such diversity of data sources and formats.

2. Insufficient techniques and less attention on data platform establishment

The results show that in the phase of data collection, analysis, and presentation, structural

database and spreadsheet softwares such as SQL Server, Oracle, Access, Python, and Excel are most frequently used (see Figure 3), accounting for around 60% of the total.

There are relatively few cases of geospatial technology and professional visualization technology, and even fewer application cases of advanced technological means such as AI, deep learning, and neural networks. More than 60% of SAIs have encountered barriers in exploring the frontier technologies and techniques of big data.

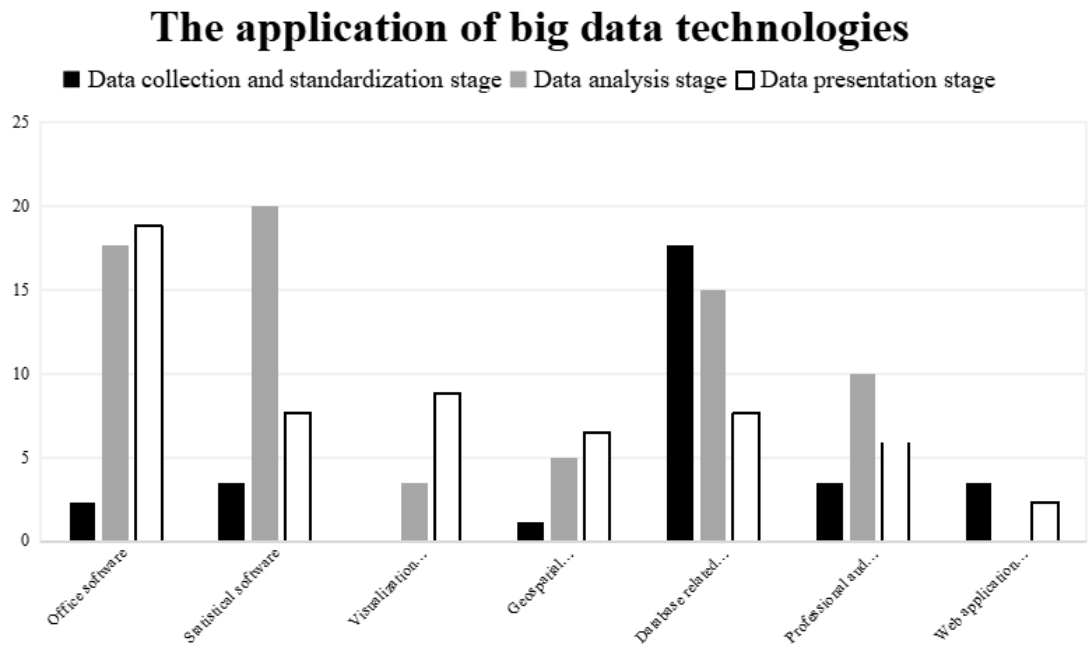


Figure 3- Statistics on the application of big data technologies

There is relevance among the insufficient support of big data audit techniques, lack of big data talents, and less developed big data platforms. Half of SAIs expressed their difficulties in addressing lack of big data technology talents, low level of big data application and less targeted activities. Although SAIs in many countries have indicated that they will train more experts on big data audit, the results of the questionnaire show 12 (55%) out of the 22 SAIs paid "average or below" attention on their big data talents, and 6 SAIs have not launched any relevant training programs.

Among the 16 SAIs that have launched such programs, their training mainly include fixed vocational classes (10/16), with less interactions (such as joint training with professional institutions; visits and exchanges). Only 36% of the respondents have established big data audit platforms.

3. Imperfect data sharing model and institutional management

The development of big data audit of SAIs is inseparable from the overall planning of responsible authorities, and the inter-department data sharing mechanism. However, at the present stage, although half of SAIs have established big data technology centers (approximately 45%), but lack a systematic organization pattern and professional talents to some extent. Among the respondents that have established big data technology centers, only 40% have set up the position of big data CTO (chief technology officer), more than 50% believe that the main purpose of the establishment of such centers is to solve issues related to professional technology and to achieve centralized management and control of data rather than the centralized management and communication of technical talents. In 50% of the data centers of SAIs, technical personnel accounts for less than 10% of the total staff, apparently inadequate for running such centers.

Meanwhile, the results show that 45% of SAIs are lack of data sharing mechanisms and available regulations in the application of big data; 41% of SAIs have limited/insufficient cooperation on big data with other countries or regions; 32% of SAIs confronted weak or problematic implementation of inter-department data exchange. This indicates that, despite of the ongoing cooperation on big data audit among many SAIs, there' s still a long way to go for many SAIs to establish a well-performed inter-organization data sharing and cooperation mechanism.

4. Under-performed big data audit in enhancing credibility and risk prediction

Although big data technology has remarkably expanded the audit field, the questionnaire respondents have not performed well in issuing accurate risk warning and enhancing social influence by applying big data technologies. Although 86% of the respondents believe big data audits are conducive to performing forward-looking audits and discovering potential risks in time, there are 63% of the respondents noted that big data technologies are not quite accurate in analyzing and predicting economic development, and about 86% of the respondents indicated that the accuracy of big data technologies is less satisfactory when it comes to predicting the risks of diseases and natural disasters.

Based on the all-around audits and relevant data sources, big data audit can perform analysis of the causes and trends of general violations of laws and regulations by audit objects and issues of common concern among the public, make in-depth exploration of the value beneath, better serve macro decision-making and institutional improvement, and enhance the social influence. The questionnaire indicates that most audit institutions believe that big data audit did improve

audit efficiency, discover audit trails and hidden risks, analyze the situation and expand the audit field, yet the role of enhancing social influence is not fully played in practical application. Among the respondents, only 18% believe that making full use of big data audit will help to promote the social influence of audit institutions, and the advantages brought by big data technologies cannot be fully utilized and shared by most of them

In summary of the feedback of the 22 respondent SAIs, we make a list from three levels – technology level, governance level, and objective level (see Table 1).

Table 1 Summary of survey results based on big data audit framework

Level	Indicators	Evaluation of Results
Technology Level	Quality of Collected Big Data	The data collection quality of the respondents are high on the whole, with slight insufficiency in terms of timeliness, collection mechanism and data accuracy.
	Big data audit techniques	Most SAIs can make comprehensive application of multiple data techniques and methods, yet the use of cutting-edge technologies needs to be strengthened.
	Big data application fields	Although big data audit techniques have been applied to multiple audit fields, the main audit objects mostly still focus on public finance.
	Big data infrastructure establishment	The establishment of big data platforms remains relatively low.
Governance level	Big data technology center	Many SAIs have established big data technology centers, the technical personnel serving in these centers are slightly insufficient.
	Training of big data talents	The training mode of big data technology talents appears single, and the importance placed on big data talents is insufficient.
	Cross-department collaboration	Although many SAIs have carried out cross-department and cross-region audit cooperation, there still exist some problems such as insufficient data sharing and barriers in cross-region and cross-department cooperation.
Objective level	Effectiveness	Big data technology can significantly enhance the audit efficiency and effectiveness of member SAIs.
	Legitimacy	The application of big data technologies has not performed well in improving the audit credibility and social influence of member SAIs.
	Forward-looking	The application of big data technologies has not performed well in realizing accurate economic forecasts and risk prediction.

BIG DATA AUDIT APPLICATION FRAMEWORK AND PRACTICAL COUNTERMEASURES

At technology level, a horizontal comparison of questionnaire results indicates that, as for most respondents, the establishment of the infrastructure platform and the innovative application of big data technologies did improve the quality of data collection and expand the application fields. For example, SAIs in 8 countries including Sudan, Turkey, Ecuador, Peru, the United Kingdom, Indonesia, Denmark and Portugal have established big data audit platforms, which contributed to the high quality of collected data and greatly expand the scope of audit. Six SAIs including SAI France, Turkey, U.K., India, Indonesia, and Portugal applied multiple big data technologies at all phases of audit, with higher quality of data collection and more fields compared with other countries. The data and cases indicate that the establishment of the infrastructure platform and the improvement of big data technical capacity can greatly enhance the quality of data collection, expand the audit field, and lay a solid foundation for building a centralized, unified, and comprehensive big data audit and supervision system.

At governance level, the respondents with established technology centers tend to pay more attention to talent training, with higher possibility of conducting cross-department and cross-region cooperation. For example, SAI Vietnam, France, Sudan, Turkey, U.K., India, and Norway attached great importance to the training of big data talents with diversify talent training modes. Meanwhile, the above 7 countries have performed well in domestic big data audit cooperation and data sharing. This indicate that the establishment of technology centers will be helpful in improving big data audit at the governance level and support the management of big data audit talents.

At objective level, the higher evaluation of indicators at the technology and governance levels, the more remarkable improvement in audit efficiency. For example, SAIs in 4 countries including Sudan, Turkey, India, and Indonesia score high on the governance and technology levels.

Statistics indicate that they also highly recognize the efficiency and effectiveness of big data audits.

Therefore, it can be inferred that the technology level, governance level, and objective level of big data audit are corralated in their development (see Figure 4). In practice, it is reflected in three aspects as follows.

First, with the support of big data technologies and the improvement of big data audit capacity, SAIs achieve massive data comparison, analyze, excavate and reveal doubtful issues and questions that were difficult to discover in the past, and improve audit efficiency and effectiveness; second, the application of big data audit techniques and the legal guarantee related to big data audits greatly eliminate blind spots in audit supervision, thereby maintaining the authority and credibility of audit supervision; third, the full coverage of big data and the cooperation of various authorities improve the ability to prevent and resolve major risks related to economy, finance and disaster, and ensure the sound operation of domestic economy and society through timely information feedback and risk warning.

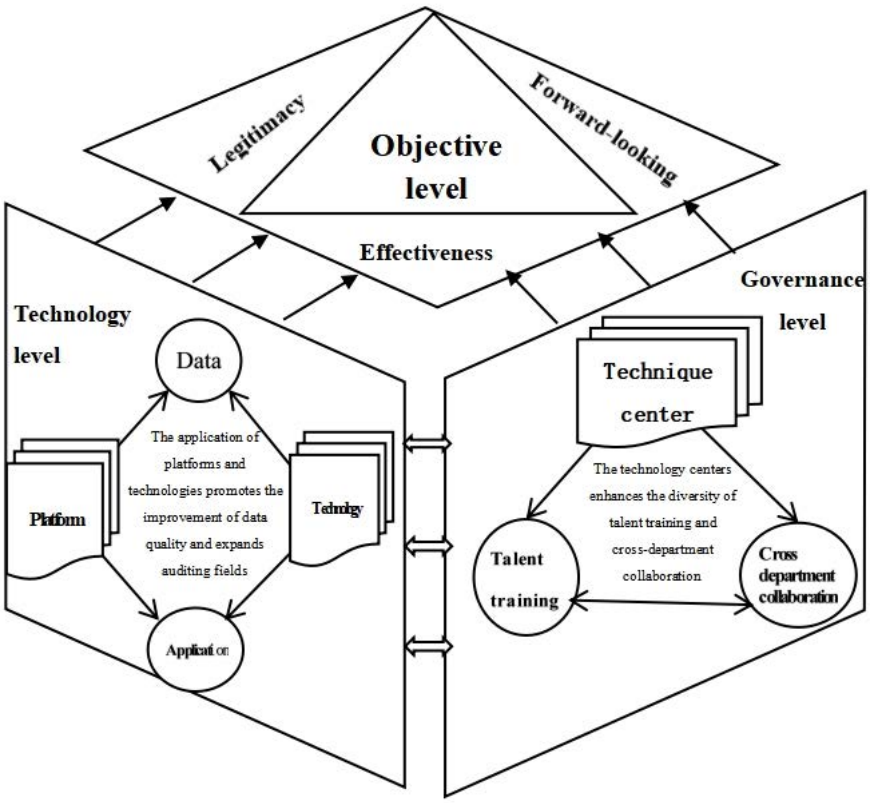


Figure 4 Schematic diagram of the mutual promotion relationship among the technology level, governance level and objective level of big data audit work

In combination with the theoretical research findings, questionnaire results and real practices, it can be found that the development of big data audit in various countries are uneven. However, the overall trend tends to be much clear. Big data audit is gradually evolving from a single data mining and analysis method to a multiple-field and cross-department audit activity that relies on infrastructure such as big data audit platforms as the hardware support, policies, systems and governance regulations as environment guarantee, combines techniques and methods with the practical audit requirements, and covers data collection, storage, transmission, analysis and presentation. Compared with traditional audits, it performs better in improving the audit efficiency, enhancing the credibility of audit findings and the ability of risk prevention.

Based on the feedback of respondents, the existing deficiencies of big data audit in various countries, and the frontier professional experience, the following three suggestions are made for the application and development of big data audits:

1. To accelerate the establishment of integrated data platforms, regulate data collection standards and explore cutting-edge technologies.

An intelligent big data platform shall integrate the function of data collection, preprocessing, storage and management, analysis and mining, presentation and application, and comprehensively integrates the key technologies of big data, which will greatly promote the full coverage of data standardization, and the capacity of audit organizations in mining and evaluating audit questions. It is recommended to build an integrated big data audit platform in the first place. The platform shall combine the functional requirements, design strategies and application environment of audit data collection, customize solutions of standard data collection technology in the platform, and integrate data mining, analysis and visualization to achieve functions such as audit evidence collection, correlation of clues, detection of doubtful issues and risk prediction.

2. To establish big data technology centers, formulate training plans, and conduct cross-department and cross-region cooperations.

By aligning the structure of governance, functions and business with the big data technologies, it will regulate audit activities and improve work efficiency, and reduce professional risks. It is recommended to select professionals with experience and technical capacity to establish

a big data audit techniques center or data analysis working group, and carry out exploration and follow-up learning of new techniques and methods by means of mutual promotion of on-site extension and data analysis, as well as combination of academic exchange and centralized training. The center (or group) shall be responsible for coordinating cross-tier, cross-region, cross-sector, and cross-department exchanges and audit collaboration, in particular, the exchanges among SAIs may refer to international seminars, training or sessions. The purpose of which is to understand, discuss, and learn the advanced technologies and experiences from various countries and localize these experiences to meet the real needs and achieve audit objectives.

3. To incorporate big data audit into national governance to improve audit credibility and the accuracy of risk warning.

The future objectives of big data audit shall include the following three aspects. First, to promote the efficiency of audit work, improve the quality of audit recommendations, and make full use of audit data in national governance; second, to enhance transparency as well as publicity of audit findings and rectifications; third, to improve the capacity of economic monitoring and risk warning and the ability to prevent and resolve major risks. It is recommended for supreme audit institutions to further reinforce their technical capacity, optimize forecasting models, expand the scope of audit collaboration, improve governance and coordination, and give full play to the effectiveness, credibility and forward-looking insights of big data audits

Appendix

Questionnaire Results Analysis
Development Prospect of Big-data Audit of Supreme Audit Institutions

Level	Primary indicators	Secondary indicators	Questions
Technology level	Data collection quality	Data collection barriers	7. Whether laws, regulations and policies related to big-data audit have been issued in your country to ensure the access to audit data of the SAI? _____ A. Yes B. No If "Yes", please list the regulations, policies and the department issued them (additional page can be attached) 8. (Single-choice) In terms of lawfully collecting the data from auditees, the SAI in your country_____ A. Almost cannot collect the data B. Can only collect a small part of the data C. Not sure about that D. Can collect most of the data E. Can collect the data without a hitch
		Data collection timeliness	9 (Single-choice) Whether do you agree that the SAI in your country can collect the required data in time? _____ A. Totally agree B. Agree C. Not sure D. Disagree E. Totally disagree 10 (Single-choice) In your opinion, whether the SAI can collect the required data regularly? _____ A. Can collect the required data at any time B. Can collect the required data regularly every month C. Can collect the required data regularly in one year D. Cannot collect the data regularly E. Others_____
		Quality and accuracy of collected big data	11 (Single-choice) Whether do you agree that the data collected by SAI in your country is accurate? _____ Totally agree B. Agree C. Not sure D. Disagree E. Totally disagree 16 (Single-choice) In terms of data standardization, the data collected by SAI in your country includes the following types: _____ A. Mainly unstructured data, a few semi-structured and structured data B. Mainly unstructured and semi-structured data, a few structured data C. Mainly semi-structured data, a few unstructured data and structured data D. Mainly semi-structured and structured data, a few unstructured data E. Mainly structured data, a few semi-structured and unstructured data
		Data Sources	14 (Multiple-choices) The following types of data collected by SAI in your country include_____ A. Transanction data B. Financial data

Level	Primary indicators	Secondary indicators	Questions
Technology level	Data collection quality	Data Sources	C. Internet data: such as data from Internet and social media related to the auditee D. Management data: such as personnel, official documents, regulations, etc E. Data of industry or monitoring data F. Others_____ 15 (Multiple-choices) The following formats of data collected by SAI in your country include _____ A. Table data B. Text data C. Graphic data D. Audio data E. Video data F. Geo-spatial data
	Big data audit techniques	Data collection technologies	17 (Multi-choice) What kind of big data technologies are used in data collection and preparation by SAI in your country? _____ A. Direct database collection B. WEB data collection C. Sensing device collection D. System log collection E. Others_____ 18 What kind of big data tools are used in data collection and preparation by SAI in your country? _____ A. SQL Server B. Oracle C. Python D. Scribe E. Flume F. Crawler4j or Scrapy G. IoT H. Others_____
		Data storage and transmission technologies	19 (Multiple-choice) What kind of big data technologies are used in transmission of audit data by SAI in your country? _____ A. Mobile devices B. Internet C. Cloud and other synchronous transmission functions D.5G E. IoT technology F. Others_____ 20 (Multiple-choice) What kind of big data technologies are used in storage of audit data by SAI in your country? _____ A. General server B. Data center system or database C. Distributed storage D. Cloud storage platform E. Others_____
		Data analysis technologies	21 (Multiple-choice) What kind of big data technologies are used in data analysis by SAI in your country? _____ A. Statistical analysis

Level	Primary indicators	Secondary indicators	Questions
Technology level	Big data audit techniques	Data analysis technologies	B. Natural language C. Data mining D. Special audit tools E. Business intelligence F. Neural network G. Artificial intelligence, machine learning or deep learning H. Others_____ 22 (Multiple-choice) What kind of big data tools are used in data analysis by SAI in your country? _____ A. Excel B. SAS C. SQL or R language D. Python E. Oracle F. BP (Back propagation) G. HADOOP H. ARCGIS or GIS 1. Others_____
		Data presentation technologies	23 (Multiple-choice) What kind of big data technologies are used in data visualization by SAI in your country? _____ A. Office software B. Statistical software technology C. Graph database technology D. Professional visualization technology E. Geographic information technology F. Natural language processing G. Others_____ 24 (Multiple-choice) What kind of big data tools are used in data visualization by SAI in your country? _____ A. Excel B. SPSS or SAS C. Neo4j D. Tableau E. PowerBI F. ArcGIS or MapInfo G. SQL or Python H. Dundas I. QlikView J. Others_____
		Data security management	25 (Multiple-choice) What kind of big data technologies are used in data security management by SAI in your country? _____ A. Improved data destruction methods B. Blockchain technology C. Firewall D. Transparent encryption and decryption technology E. Others_____
	Big data application fields	Big data application fields	38 (Multiple-choice) The SAI in your country has used big data technology in the following fields_____ A. Enterprise auditing

Level	Primary indicators	Secondary indicators	Questions
	Big data application fields	Big data application fields	B. Audit of financial institutions C. Resource-environment auditing D. Education auditing E. Medical and social insurance auditing F. Financial aid auditing G. Government budget auditing H. Policy review auditing 1. Traffic management G. Criminal investigation K. Customs auditing
	Big data infrastructure establishment	Big data platform establishment	26 (Fill in the blanks) Whether SAI in your country is building or has built an integrated data platform for auditing_____ The name of the platform is_____ A. Yes B. No
Govern ance level	Big data technology center	Establishment of the big data technology center	29 (Fill in the blanks) Whether SAI in your country has established a special technical department or center of big data? _____If Yes, year of establishment_____ A. Yes B. No
		Functions of the big data technology center	29-1 (Multiple-choice) The role of the technical department or center _____ A. Solve professional technical problems B. Centralized training of talents C. Technical communication among various talents D. Centralized management and control of data E. Others_____
		Scale of the big data technology center	29-2 (Fill in the blanks) The total number of staff in the technical department or center is _____
		CTO (chief technology officer) position	28 (Fill in the blanks) Whether SAI in your country has appointed a position similar to technical director of big data to take lead and promote the development and application of big data in SAI?_____ Yes B. No
		Talent training modes	31 (Multiple-choice) The SAI in your country trains talents for big-data audit by the following ways_____ A. Centralized professional training B. Cooperate with professional institutions, such as universities and other scientific research institutions C. Take a temporary post in other organizations D. Visit other countries for exchange E. Others_____

Level	Primary indicators	Secondary indicators	Questions
	Cross-department collaboration	Cross-department collaboration pattern	35 (Single-choice) Whether do you agree that the SAI in your country fully cooperates with other countries or regions based on big data? _____ Totally agree B. Agree C. Not sure D. Disagree E. Totally disagree
		Data sharing mechanism	36 (Single-choice) How do you think the data exchange and sharing among departments of SAI in your country A. Very good. The data can be exchanged and shared among departments without obstacles. B. Good. The data can be exchanged and shared among most departments. C. Normal. The data can be exchanged and shared among some departments. D. Poor. There is little data exchange among departments. E. The information system is not available, so it is impossible for data exchange. F. Not sure
Objective level	Audit efficiency	Improving audit efficiency with big data	39 (Multiple-choice) What role do you think big-data analysis can play in the work of the SAI_____ A. It is conducive to improve the efficiency of audit work B. It is conducive to discover the clue to problems more accurate C. It is conducive to carry out forward-looking audit and discover risks D. It is conducive to analyze the overall situation of specific fields E. It is conducive to expand the audit coverage, and more funds, projects and auditees can be included. F. It is conducive to enhance the social influence of government audit G. There is no obvious effect 40 Whether do you agree that the big data technology is an important means to improve the audit efficiency of SAI Totally agree B. Agree C. Not sure D. Disagree E. Totally disagree
	Audit service quality	Advance social influence	42 Whether do you agree that SAI in your country has improved the social influence of audit by performing big-data audit? _____ Totally agree B. Agree C. Not sure D. Disagree E. Totally disagree
	Economic forecast and risk warning	Macroe-conomic forecast	Whether do you agree that SAI in your country has performed more accurate analysis and forecast of economic development by using big data and related technologies? _____ A. Totally agree B. Agree C. Not sure D. Disagree E. Totally disagree
	Economic forecast and risk warning	Risk Warning	44 Whether do you agree that SAI in your country has performed more accurate risk forecast (such as diseases and natural disasters) by using big data and related technologies?_____ A. Totally agree B. Agree C. Not sure D. Disagree E. Totally disagree

Secretariat of INTOSAI Working Group on Big Data

Add: National Audit Office of the People's Republic of China
No.17 Jin Zhong Du Nan Jie, Fengtai District, Beijing 100073, China

Tel: + 86 (10) 50992782

E-mail: intosai-wgbd@audit.gov.cn