

## China's Audit in the Era of Big Data

Hello everyone, it gives me great pleasure to attend this meeting in Mexico City. I am honored to share with you the experiences and practices of China National Audit Office (CNAO) in big data audit. As we have seen in our daily lives, modern information technology such as big data, cloud computing, and AI profoundly changes the way of thinking, production, life and learning of human beings, and gives a strong impetus to the national governance capability to advance towards a higher peak. At the same time, national audit enjoys rapid development in various countries, responsibilities of which are become increasingly apparent, i.e., strengthening of accountability, promotion of good governance, tracking and supervision for performance of sustainable development obligations in various countries. In the era of big data, launching of big data audit echoes with the development needs of the times. In recent years, National Audit Office is always committed to taking effective measures to intensively collect all kinds of electronic data, actively gropes for new ideas and new ways for electronic data audit, and vigorously makes progress in audit informatization construction with “big data” as the core. Today, I will introduce the major practices of CNAO in big data audit in recent years from seven aspects. At the same time, we will also discuss the four difficulties and challenges encountered in big data audits.

### I. Major Practices of CNAO for Big Data Audit

#### (I) Improving policy environment to promote big data audit.

In 2015, the Chinese Government announced *Action Plan for Promoting Big Data Development*, which put forward the implementation of National Big Data Strategy so as to promote China's big data application and development. In 2016, China released *The Thirteenth Five-Year Plan for National Economy and Social Development*, which proposed to regard big data as fundamental strategic resources, comprehensively take action for promoting big data development, accelerate the sharing, development and application of data resources, and contribute to industrial transformation and upgrading and social governance innovation.

In 2014, *Opinions on Strengthening Audit Work* was printed and issued by National Audit Office, which specified that “Competent authorities, financial institutions, state-owned enterprises and public institutions should depend on the requirements for audit work, and provide audit institutions with electronic data information and the necessary technical documentation associated with the fulfillment of functions of such entities and sectors”. At the national level, *Opinions on Strengthening Audit Work* offers a favorable environment for promoting the audit informatization construction. In 2015, *Framework Opinions on Several Important Issues in Improvements of Audit System* and related supporting documents were further printed and issued by National Audit Office, which expressly requested audit

institutions at all levels to “establish work mode for big data audit, improve capability, quality and efficiency of audit work, and expand breadth and depth of audit supervision”.

## **(II) Improving operating bodies, and rendering favorable organizational support.**

At present there are two departments related to the data auditing, which are the Department of Electronic Data Audit and the IT Center. Department of Electronic Data Audit mainly charges for work as follows: centralized management of audit business of electronic data, organizing audit business electronic data acquisition, acceptance and finishing work, and creativity to carry out cross industry, cross sectoral, cross regional data analysis, comprehensive analysis and use of electronic data; at the same time, business guidance of the computer technical center, and guide the local audit institutions electronic data audit business. IT Center mainly charges for work as follows: organization and coordination of IT application of the audit system; providing IT services for the CNAO; management of the audit network and hardware; IT training and assessment examination.

In addition, various functional divisions of National Audit Office also set up analysis teams of different operational fields to specifically carry out in-depth data analysis dominated by operational matters in their own fields. In addition, for the sake of national-class large projects, experienced professionals were transferred throughout the country to set up a cross-sectoral data analysis team. With the relevant functional departments and various types of analysis team, National Audit Office not only achieves the centralized management of electronic data, but also reinforces the application of big data mining and analysis technology, which guarantees the acceleration of big data audit work.

According to the data audit situation, local audit institutions set online audit department, data analysis center or IT center etc, which mainly charges for local governments data audit work.

## **(III) Continuously promoting audit informatization construction under the goal of audit efficiency improvement.**

Since 1998, National Audit Office has vigorously promoted informatization construction. “Golden Audit Project (Phase I)” kicked off in 2002, which was completed and accepted in November 2004. “Golden Audit Project (Phase II)” kicked off in 2007, which was completed and accepted in 2012. From 2015 onwards, “Golden Audit Project (Phase III)” with “Big Data” as the core has been under construction, which Digital audit support capability is established, i.e., “Overall analysis, systematic research, troubleshooting, decentralized verification and precise positioning”. All efforts have been made to give full play to revelation, resistance and prevention functions of audit.

Nowadays, in the network construction, CNAO has build a wide area network(WAN) with more than 40 dispatched agencies and 37 provincial audit institutions. In the software

deployment, the Office Automation (OA) system has become the main software of audit management in an IT environment by each audit institution at all levels nationwide, supporting the informatization of audit management effectively; the Auditors' Office (AO) system, has issued to the national audit institutions more than 100 thousand units, mainly used in the audited units of financial data processing and analysis, audit methods to sum up and share and so on.

#### **(IV) Upgrading data collection, and building big data resource library.**

National Audit Office constantly improves various systems, clarifies work requirements for access to electronic data of various regions and relevant departments, and standardizes data validation, statistical analysis and other systems and processes in various fields. Nationwide auditing authorities at all levels constantly enrich electronic data, collect and sort up the industry-wide and region-wide electronic data, form basic data resource library, and initially build analysis environment for hardware, operating systems, database software and data analysis tools.

#### **(V) Strengthening data analysis, and innovating technical means.**

The audit data analysis methods include the following three levels: the first one is based on the audit expert experiences and routine audit analysis technology auditing, performance for the users of the records in the database to access and query, through the SQL language to interactive to describe the query request, or according to the needs of the query by development tools customized query software, called "query analysis". The second one is based on the audit analysis models and multi-dimensional data analysis technology, users put forward their own assumptions, then the various tools through repeated, recursive query, to verify or deny the hypothesis, from the user's point of view, this is finding facts through drilling data. So this level belongs to "verification analysis". The third one is based on the data mining technology, users find specific schema from 'big data' to predict trends and behaviors of the data analysis model, it can tap the potential data model, find information that are neglected by the users and help auditors to make forward-looking, knowledge-based decisions. Big data analysis can be called as "finding analysis".

Big data audit has two notable features: One feature is comprehensive use of diversified data from complex sources, and the other feature is innovative use of intelligent data analysis technologies and methods, such as cloud computing, intelligent mining, complex networks, natural language understanding and other emerging big data analysis technologies. Over the past few years, a series of data association analysis have been launched in the course of the audit for macroeconomic operation and implementation of major national policies, thereby demonstrating the macroeconomic performance and revealing the major risks and hidden dangers. In various professional audit fields, National Audit Office makes full use of

multi-industry data from fiscal administration, tax, industrial and commercial registration, enterprise and the like, carries out cross-industry, cross-level and cross-system comprehensive data analysis, profoundly look for clues, and precisely detects problem. For example, in the implementation of the national precision policy audit, fiscal poverty alleviation data are used to track the true destinations and effects of fiscal poverty alleviation funds. In the implementation of environmental audit, geographic information data and relevant data on forests, wetlands, grasslands, farmland coordinates and others are used to track the changes in the ecological environment.

#### **(VI) Facilitating the establishment of international standards for accounting software data interface, and resolving data diversity problems.**

Since 2003, National Audit Office has made unremitting efforts to roll out national standards for software data interface. In 2004, Standardization Administration of the People's Republic of China approved and issued *Information Technology: Data Interface of Accounting Software* (GB/T 19581-2004), which greatly simplifies data collection. This standard is also widely supported by Chinese and foreign software enterprises. In March 2015, the International Organization for Standardization approved the establishment of "Project Committee for Audit Data Collection (ISO/PC 295)". Secretariat of ISO/PC 295 is based in National Audit Office of the People's Republic of China. ISO/PC 295 consists of 17 Member States and 23 Observer States up to now, which is quickening the pace to develop international standards for audit data collection.

#### **(VII) Actively fulfilling various tasks of INTOSAI Big Data Working Group, and joining hands with national auditing authorities to address the challenges from big data.**

In recent years, INTOSAI has reached a consensus that Supreme Audit Institutions play a key role in promotion of harmonious and inclusive social sustainability and tracking and supervision for performance of sustainable development obligations in various countries. Supreme Audit Institutions must innovate the audit methodology, and further expand the scope and spectrum of audit. Many Supreme Audit Institutions have gained some experience in improving the efficiency and quality of audit by means of big data technology. However, big data audit has also brought about significant challenges to the existing audit organization management, audit talent cultivation and audit data collection and collation.

The 68th meeting of the INTOSAI Governing Board approved the motion for the establishment of a new Working Group on Big Data (WGBD) with SAI China as its chair and SAI U.S. as vice chair. The XXII INCOSAI endorsed the motion and listed the new working group under Goal 3 - Knowledge Sharing Committee. As chair of the Working Group, CNAO successfully held the first meeting of WGBD from 17 to 19 April 2017 in Nanjing, China. At the meeting, 49 delegates from 18 SAIs discussed the development of government auditing

under the big data environment. In the working session, delegates from 13 countries talked about challenges and opportunities faced by the SAIs in the big data era, shared their experiences in big data audit, and discussed prospects for the future. The meeting has determined 4 strategic goals of INTOSAI WGBD as follows:

- Identify opportunities and challenges SAIs are facing in a big data world to make recommendations to SAIs.
- Summarize the know-how, experiences and good practices concerning big data audit to help SAIs boost relevant skills.
- Develop guidelines and research reports to share knowledge and support capacity building activities in big data audit.
- Strengthen bilateral, regional and INTOSAI-wide cooperation among SAIs in big data audit.

To this end, INTOSAI Big Data Working Group is launching a series of research projects with focus on “Basic Theory of Big Data Audit”, “Technology Innovation of Big Data Audit”, “Management Reform of Big Data Audit” and “Application Practice of Big Data Audit”. All interested colleagues, who are present today, are highly welcomed to contact INTOSAI Big Data Working Group, join in this collective and jointly address challenges from big data.

## II. Challenges Encountered by Big Data Audit

In addition to introducing our audit practice, I would also like to discuss with you the challenges we face in the development of big data audit. These challenges are specifically exposed in terms of data collection and processing, planning and construction of data center, data analysis technology, risk management, etc.

### (I) Data collection and processing.

On one hand, collection scope for audit-oriented big data covers performance-related financial data, operational data and management data of audited entities. No doubt, this is the main data source for audit-oriented big data. On the other hand, unstructured documents, Internet data and even the future Internet of Things show ever-increasing value for audit data analysis, which should also be included in the collection scope. It is generally believed that high data volume is preferred compared with low data volume. The biggest advantage of big data lies in enormous room for data selection.

However, it is precisely because of this that we have to be bothered by such a dilemma when enjoying the richness of data: Is storage capacity strong? Is sorting capacity strong? Is

analytical capability strong (timeliness and sufficiency)? How to identify the authenticity of the data? How to select the data association? How to classify and screen the data so as to prevent from being overwhelmed by “data noise”?

In the context of big data, approach to effective identification, sorting, classification, extraction, absorption, distribution and storage of the data is one of the important challenges for big data audit work.

## **(II) Planning and construction of a data center.**

Like any mass storage system, an audit data center must be constructed to have high scalability, high performance, fault tolerance, scalability, low operating cost and other features, which should also satisfy the requirements of audit-oriented big data for complexity, uncertainty, dynamics and other technical characteristics.

From the perspective of data scope, an audit center covers operational data and management data of audited entities, as well as the Internet information. From the perspective of function, multiple operational functions of data processing, storage, query, comprehensive analysis and service are involved. Among them, data sorting and analysis are core functions of an audit data center. From the perspective of workflow, original data are acquired from the audited entities and the Internet. The original data in different storage patterns are stored and standardized by means of secure and controllable database software. Finally, high-performance parallel database technology is applied to provide standardized and normalized data analysis services.

These audit demands make planning for audit data center much more complex than big data architecture design for a single industry. Approach for an audit data center to apply big data storage management technology to achieve overall design and gradual implementation is an important basis for big data audit.

## **(III) Innovation of technical methods for audit analysis.**

It is generally believed that big data analysis is characterized by “Cross Fusion + Intelligent Mining”, which are two major difficulties in the current innovation of technical methods for audit analysis.

Audit data analysis of National Audit Office has smoothly proceeded for more than 10 years. Over these years, National Audit Office goes through data separation analysis stage (data analysis in professional audit fields). In recent years, Chinese auditors are carrying out a variety of cross-domain, cross-level, cross-system, all-dimensional and cut-through data analysis.

From the perspective of data usage, we enter big data analysis phase. Another feature of big

data analysis lies in widespread application of various depth mining and intelligent learning algorithms. At present, National Audit Office is strengthening the use of cloud computing, data mining and other technologies in order to raise application level of “big data” mining and analysis technologies.

#### **(IV) Big data analysis risk and quality control.**

A variety of risks exist in audit-oriented big data in terms of authenticity, integrity and interpretativeness. With specific progress in big data audit work, clear understanding and necessary prudence should be maintained. For example, authenticity is the value cornerstone of all data, which is also a congenital defect of big data. For example, from the perspective of integrity, big data are currently still in the form of “Isolated Island”. Strictly speaking, no authority can have access to sufficient data in terms of breadth and depth. Moreover, from the perspective of causality and relevance, big data analysis is made with the help of machine, but machine is merely able to demonstrate the relationship between the data. Causality must be thought and judged by people. It is likely to lead to errors or even dangerous conclusions in case of big data analysis with emphasis on relevance rather than causality. Approach to controlling risks of big data audit and improving quality of the audit is a quite critical challenge which even influences the future development direction of big data audit.

Despite so many hardships and difficulties, the worldwide auditors should cherish a common dream, i.e., radiating audit-oriented big data wisdom. Such big data wisdom should substantially cover four aspects, namely, “Perception”, “Self Learning”, “Integration” and “Collaboration”. “Perception” and “Self Learning” are concepts of artificial intelligence. Through inter-data association and multi-dimensional and multi-angle intelligent analysis, big data analysis mines hidden behavior mode, which constantly perceives the risks through self-learning. In terms of “Integration” and “Collaboration”, extensive-range data integration is a must to find out the relationship between the data. Big data is an interdisciplinary concept, which requires work specialization. It is necessary to achieve work specialization between various operations, and achieve the data sharing and synergy.

At the end of my speech, dear colleagues, big data audit is not a technical change, but rather a type of strategic asset in the development of audit undertaking. Our audit undertaking is bestowed upon with unlimited opportunities by the era of big data. All auditors of this generation shoulder lofty mission and responsibility to face up to challenges and bravely make breakthroughs. Let us continuously increase mutual exchanges, strengthen mutual cooperation, and make our concerted efforts for promoting big data audit to continuously write brilliant chapters all over the world!

Thank you!