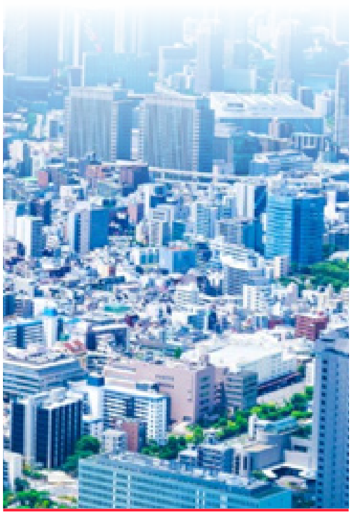
Tokyo Shoko Research Reduces Data Entry Time By 80% with ABBYY AI OCR

With 133 years of history, Tokyo Shoko Research has its own unique approach to digitalization. The company implemented ABBYY Document AI technology to greatly improve the efficiency of the process of converting information obtained by surveyors nationwide into a database. This significantly reduced the time required for the task—by up to 80%.



Tokyo Shoko Research (TSR) has been conducting corporate credit investigations for 133 years, since its founding in 1892. TSR services include the TSR Report (Domestic Company Research Report), which is created based on information collected by field researchers who visit companies directly nationwide, as well as sales and marketing support utilizing Japan's largest corporate information database. Additionally, the company has a business partnership with Dun & Bradstreet (D&B), which holds the world's largest corporate information database, and provides access to over 600 million corporate records globally.

ABBYY's AI OCR is being utilized in the company's business reform efforts, enabling improvements in productivity. We had the opportunity to speak with Mr. Mitsuo Kawahara, President and CEO; Mr. Gen Tanaka, Executive Officer, Management Division, Corporate Planning Department and Systems Division; and Mr. Tetsu Tokiwa, Deputy General Manager of the Systems Division, about the implementation of AI OCR and its effects.

Challenge

Leveraging data freshness and accuracy through AI OCR

"TSR, now in its 133rd year, continues to uphold traditional values even in an era where information is widely accessible online. While TSR manages information on 1.6 million companies annually, only a small percentage of that data can be sourced from the web. The majority is gathered directly by TSR's field investigators stationed nationwide—one of TSR's key strengths—who manually transcribe real, on-site information," said Mr. Kawahara, President & CEO.

"For us, data freshness is vital. To deliver truly valuable data, daily updates are indispensable. Until recently, entering data into our systems from financial statements—obtained as copies, PDFs, or photos—took significant time. Our partnership with ABBYY substantially changed that; automatically digitizing collected information has drastically reduced time and improved data quality," he said.

Mr. Kawahara
President & CEO

Mr. Tokiwa
Deputy General Manager of the Services Division

Mr. Tanaka
Executive Officer

"Even in today's environment, where various information can be accessed online, we continue to deliver unique value by utilizing information gathered firsthand from investigation sites. On the other hand, keeping pace with technological advancements is also essential for TSR to enhance its corporate value. While OCR itself is not a new technology, high hardware and maintenance costs and low recognition accuracy had previously hindered its full-scale use for documents like financial statements and real estate registers. ABBYY's AI OCR, however, offers both the benefits of cloud services and outstanding recognition accuracy. These advantages are the main reasons we adopted the solution," said Mr. Tokiwa, Deputy General Manager of the Services Division.

At TSR, documents obtained from investigation targets and external information providers are used as inputs for corporate information research and registration. Until now, the process of converting these documents (paper, PDFs, images) into text and storing them as structured data required manual labor, which posed challenges to operational efficiency. While some OCR tools and outsourcing were previously used, the high cost of these products meant that the return on investment (ROI) was not sufficient, making it unfeasible for nationwide deployment.

ABBYY was proposed by IBM, which is responsible for the development and maintenance of TSR's database system. In light of the growing interest in practical and commercial applications of generative AI, the solution was selected for its proven global track record as part of a pilot to assess the potential of AI OCR. This represented an advanced use case on global scale—automating data extraction from unstructured documents with AI OCR to address challenges in corporate information management.

Solution

Cross-functional team tackles project complexity and varying document structure

This project targeted both printed financial statements and real estate registration documents in PDF format. Both document types featured multiple layout variations and lacked standardized formatting rules, which meant that simple document definitions were insufficient. As a result, customization and programming were required. While the project progressed as planned, the level of complexity turned out to be higher than initially expected.

Following approximately three months of proof of concept (PoC), TSR transitioned to a full-scale project with an 11-month timeline. Since this project involved core database registration operations, they established a cross-functional structure that included field investigators, specialized departments, and headquarters. FlexiCapture's flexible document definition capabilities successfully addressed all of these challenges.

AI OCR and core database system automatically linked

For real estate registration documents (PDF), after automatic input using AI OCR, a minimal visual check is performed, and the data is automatically linked to the central database system. The previous manual workflow has been completely revamped. For financial statements, data is linked via photos taken with mobile phones, enabling the manual workflow performed by surveyors nationwide to be semi-automated, significantly simplifying manual input tasks.

Value

Data entry time reduced by up to 80%

	Before implementation	After implementation
Time required for registering new financial statements:	Average **13.5 minutes per case**	Average **2.5 to 3.5 minutes per case** (depending on PDF/photo format)
Time required for registering new real estate registration records:	Average **15 minutes per case**	Average **5 minutes per case**

In addition, the scope of outsourced work was reduced, impacting costs.

Post-implementation maintenance

In this implementation, AI OCR was integrated with an existing business application. As a result, seamless operability with the existing system was ensured. However, due to the variations in document layouts, the document definitions and the structure of the associated programs became more complex, leaving some challenges in terms of maintainability and operational management. For tasks where the layouts are standardized and the work is consistent, simple training is sufficient to ensure smooth adoption. In addition, both ABBYY and IBM representatives who participated in this project responded swiftly and proactively when technical difficulties arose, and the outcomes delivered were favorable. Overall, TSR says that it is highly satisfied with their support.

Future outlook

TSR aims to leverage ABBYY's superior data extraction performance while integrating it with other technologies for structured data to enhance flexibility, responsiveness, and operational efficiency.

> "Financial statements are often provided in PDF or paper format, so many companies can benefit from the automation of data entry demonstrated in this case study. We hope that companies that have not been able to achieve this with traditional OCR will use this case study as inspiration," **said Mr. Tanaka, Executive Officer.**

From left, TSR Deputy General Manager Tokiwa; President Kawahara; General Manager Tanaka; ABBYY Country Manager, Japan, Tomohito Shimizu