INDIUM

Indium's Gen AI-powered solution helped a real estate leader achieve 700x faster data extraction

SUCCESS STORY

Manual extraction – A time-consuming bottleneck

Our client, a digital middle-office solutions provider in the real estate and private equity space, builds innovative solutions for virtual data management, data integrity, and security. Their challenge? Extracting relevant data from various documents – a complex and time-consuming process that hampered their efficiency.

Here's a closer look at the specific challenges they faced:

Manual data extraction: Extracting data from documents like asset valuation reports, real estate assessments, cooperative venture contracts, and lease collection records was a tedious, manual process.

Document complexity: These documents were unstructured, large, and contained tabular data with varying complexities, making manual extraction even more challenging.

Time consumption: Manually extracting data significantly slowed down their workflows.

The need for automation!

The client needed a solution that could automate data extraction, improve efficiency, and reduce the time spent on this crucial task.

Efficiency: Reducing the manual workload associated with data extraction.

Accuracy: Extracting data with high accuracy to ensure reliable information.

Document variety: Handling a wide range of document types with varying complexities, including unstructured data and tabular data.

An ingenious solution:
Automating extraction with Indium's Gen AI expertise

Understanding our client's need for an intelligent and efficient solution, we at Indium crafted a multi-pronged approach powered by Gen AI. Here's how we tackled the challenge:

▶ **Building a robust infrastructure:** We leveraged the power of the AWS cloud environment, utilizing AWS Bedrock services and building robust backend APIs. This foundation provided a scalable and secure platform for our custom-developed solutions.

▶ **Document parsing with Generative AI:** Powered by AWS Bedrock, a core element of our solution was a custom-built document parser. This intelligent tool analyzed the structure and content of each document type, identifying and extracting relevant data fields with high accuracy.

▶ **Tailored pipelines for different documents:** We didn't employ a one-size-fits-all approach. Instead, we designed specialized data extraction pipelines for each document type (asset valuation reports, real estate assessments, etc.). This ensured optimal performance and accuracy for each document format.

▶ **Advanced AI models and tools:** We incorporated a range of powerful AI models and tools to achieve the best possible results. OpenSearch provided a flexible and scalable search platform, while FAISS facilitated efficient retrieval of similar documents. Additionally, we leveraged the capabilities of foundation models like Titan and Cohere and advanced retrieval-augmented generation (RAG) to enhance the extraction process.

▶ **Integration with scanned documents:** Recognizing the presence of scanned documents in the mix, we integrated AWS Textract. This powerful tool accurately extracts data from scanned documents, ensuring seamless processing of all document formats.

▶ **Data quality at the forefront: Maintaining data accuracy was paramount.** We implemented rigorous Data Quality (DQ) checks across the extracted data, employing filtering mechanisms to guarantee clean and reliable outputs. This ensured that the client received trustworthy data ready for further use.

By combining these elements, we created a comprehensive solution that addressed the client's specific needs for efficient and accurate data extraction from complex and varied documents.

Quantifiable success: How Gen AI delivered measurable results

The impact of our AI-powered solution was clear and measurable. Here's how GenAI helped our client achieve significant improvements:

▶ **Soaring accuracy, 87% across all documents:** One of the primary concerns for our client was data accuracy. Our solution delivered exceptional results, achieving an impressive **87% accuracy rate across all document types.** This ensured that the extracted data was reliable and ready for further analysis and use without extensive manual verification.

▶ **Dramatic reduction in manual effort from days to hours:**
The time-consuming nature of manual data extraction was a major bottleneck for the client. Our solution streamlined the process, resulting in a staggering **700x reduction in manual effort.** This translates to tasks that previously took days to complete, which are now achievable in a matter of hours. Imagine the reallocated resources and the potential for increased productivity!

▶ **Significant cost savings, a 4x advantage:** The efficiency gains brought about by automation also yielded substantial cost savings for the client. By eliminating the need for manual data extraction, the client achieved a **4x reduction in costs.** These savings can be re-invested in further growth initiatives or used to deliver greater value to their clients.