



+91 (44) 2230 2751

info@systechusa.com

184-187 Anna Salai, Little Mount,  
Saidapet, Chennai, Tamil Nadu 600015

# Driving AI Enabled Data Ingestion

## Solving Complex Challenges in Data Strategy and Analytics for a Pharmaceutical Giant

### About Client

The client is a multinational pharmaceutical firm based in India, with a global footprint spanning over 100 countries. It ranks among the largest specialty generic pharmaceutical companies in the world.

### Business Need

The client aims to ensure timely drug availability, optimize inventory, and swiftly adapt to market fluctuations. They also seek to bolster data-driven decision-making, streamline the supply chain for cost-efficiency, and swiftly respond to market trends and competition.

### Business Objectives

The core goal was to furnish the client with a robust Microsoft Azure-based data and analytics platform. This platform aims to tackle key business challenges by establishing a centralized data repository and a scalable data ingestion system. The latter leverages AI models for document processing to extract data from a diverse range of files shared by distributors.

### Challenges

The client encountered obstacles in aggregating data from primary and secondary sales, performing OCR on PDFs, and recognizing entities in Excel files. The absence of a unified data model and centralized data hub amplified these challenges. Manual collection of fragmented data from multiple distributors and the inconsistent data formats further complicated matters, hindering a comprehensive view.

### Strategic Approach

The strategy involved a two-pronged approach to data extraction and processing. Structured data from ERP systems and third-party aggregators underwent data integration, feeding into a centralized enterprise data hub. For unstructured data, OCR techniques were employed on PDF files from distributors, followed by entity recognition. Excel files were parsed to extract cell-by-cell information, which was then processed for entity recognition.



+91 (44) 2230 2751

info@systechusa.com

184-187 Anna Salai, Little Mount,  
Saidapet, Chennai, Tamil Nadu 600015

## Solution Overview

The resulting architecture was a unified enterprise data hub, collating data from third-party aggregators, in-house ERP, and distributor invoices. This enabled a multi-faceted analytics framework, inclusive of a sales dashboard providing insights into primary and secondary sales and stock movement. The platform's flexibility also accommodates the deployment of additional dashboards, self-service analytics, and predictive models.

## Results

Leveraging Microsoft Azure's capabilities, the client overcame their key challenges, leading to efficient stock management through data insights. Automated data pipelines fed into the centralized data hub, and a framework was established for the seamless onboarding of new distributors. The analytics dashboard generated actionable insights across sales and stock metrics, allowing the client to scale their data-driven decision-making effectively.

