

## CASE STUDY

Technology

Providing intelligent data extraction solutions to analytics relation specialist



### About the client

The client is a global consulting firm providing data solutions through AR-driven strategic analytical insights.



**Headquarters**  
Kansas City,  
MO



**Launched**  
2012



**Employees**  
200

### Objective and challenges

The analytical insights provider wanted to optimize and automate its entire data extraction & analyzing process. To be the go-to expert as a leading data solution provider, the consulting firm reached out to Netscribes to develop an AR-based platform to assimilate accurate and informative data from various online media to obtain analytical solutions.

### Challenges and obstacles

- It was a manual process to provide holistic data insight at scale.
- Extracting and analyzing data from online media like blogs, podcasts, and videos was not feasible as it involved tangible human efforts and time

These limitations made the data extraction and analysis an exhausting task for this data solution firm's employees.

#### Phase I: Text analysis



- To develop an intelligent module, multiple components like Data Collectors, Data Extractors, Data Parsers, and NER Models were incorporated
- To extract data from PDF, HTML, and Excel-based reports, the module was implemented
- To control the quality of the mined dataset, they were sent to the QC team to be reviewed
- To increase machine accuracy and decrease manual involvement, the ML module was re-trained based on the QC review

## Phase 2: Image analysis



For optimizing the intelligent module we worked on improving the image processing step. Here's how:

- To revamp the image quality, Stable Diffusion techniques were applied
- To classify images, NS proprietary CNN ML models were implemented
- To annotate images based on Object Detection, proprietary R-CNN ML models were utilized

## Phase 3: Video analysis



- To extract data from videos, they were split into image frames and audio
- To revamp image frames, they were refined through the image analysis module
- To convert the audio into text, they were screened through standard transcription techniques. Next, the extracted text was processed through the text analysis module

The output of the intelligent pipeline was an analysis of the reports extracted from multiple sources, like databases, social media, and other publicly-available information. The optimized module added more attributes that offered more insights.

## Results delivered

Through our joint effort, the firm witnessed:

Improved  
TATs by  
**50%**

Increased  
efficiency by  
**30%**

Human involvement  
reduced to  
**20%**

Data accuracy  
optimized by  
**70%**

## Benefits

Armed with Netscribes' intelligent data extraction, the client overcame its setbacks and limitations with a solution that enabled them to gauge the competitor landscape and analyze data from more data sources sustainably. It was automated with reliable data analysis and the least manual intervention which in turn maximized the brand's productivity.

**Optimize existing operations to gain market advantage with Netscribes  
AI and automation solutions.**

[Contact us](#)



Proprietary and Confidential, Copyright © Netscribes Data & Insights. All Rights Reserved.  
The content of this document is confidential and meant for the review of the recipient only.

[www.netscribes.com](http://www.netscribes.com)

