

Rule-free Records Classification for Scanning Service Provider

A renowned scanning service provider experienced remarkable improvements by implementing PLANET AI's Intelligent Document Analysis (IDA) software suite. The company could move beyond its previous, rule-based document classification system. The adoption of IDA's machine learning approach facilitated quick deployment, swift adaptation to regulatory changes, and a significant expansion of business operations due to decreased resource requirements.

THE CLIENT

Our client has been offering scanning services as part of business process outsourcing (BPO) to healthcare providers, the public sector and enterprise customers for over 50 years. The client specializes in digitizing, indexing, and archiving documents and processes a volume of one billion pages annually.

THE CHALLENGE

The client's daily processes involve scanning and reading incoming boxes of documents. Before implementing the IDA software suite, they used a market-leading Intelligent Document Processing (IDP) product for **rule-based document classification**.

However, our client encountered significant difficulties in automating their document scanning and archiving processes. They struggled with an **automation rate of only 50% for document classification**. As a result, a substantial amount of manual work was

OUTCOMES ACHIEVED

91%
out of the box automation rate¹⁾

80%
reduction in time spent on
manual corrections

¹⁾The automation rate refers to the percentage of document processing tasks that can be completed without human intervention.

required to correct and validate documents that fell below the necessary confidence thresholds.

Additionally, the client's existing rule-based document classification approach was resource-intensive and expensive to maintain, especially when it came to modifying existing document classes or creating new ones. The **introduction of regulatory changes** further emphasized the need for a more precise categorization system that their previous solution couldn't effectively deliver.

THE SOLUTION

During the evaluation phase, IDA showcased a remarkable performance with an **automation rate of 91% for document classification out of the box**. This achievement can be attributed to IDA's state-of-the-art machine learning capabilities, providing outstanding accuracy and confidence scores by analyzing both textual and visual document features.



Unlike the previous solution, IDA follows a rule-free, few-shot learning approach. This allowed for a **swift initial setup in less than three days instead of the lengthy 240-day setup time with the previous rule-based solution**. Regulatory changes necessitated a significant expansion in document classes, growing from approximately 150 to over 300. IDA's ability to quickly adapt to new document classes enabled the client to meet the requirements promptly.



The implementation of IDA had a remarkable impact, leading to **an 80% reduction in time spent on manual corrections**, thanks to enhanced straight-through processing. Additionally, it successfully met the client's goal of keeping the **rate of false positives below 1%**. This is significant because a false positive – in general terms, like incorrectly classifying a legitimate email as spam – can lead to unnecessary manual review. In this context, maintaining a low rate of false positives ensures that most documents are processed correctly on the first attempt, thus minimizing human intervention.



Overall, the substantial improvement in automation empowers our client to significantly **scale their business**, being able to complete more indexing projects in the same time.

OUTLOOK

Scanning and processing large batches of documents can become incredibly challenging when dealing with PDFs that contain 100 or more consecutive pages. The process of separating the documents currently requires a significant amount of manual effort, which usually involves inserting separator sheets between each document.

However, in the future, the client plans to address this issue by leveraging IDA's **document splitting feature**. This will enable them to train a neural network to automatically detect the start and end of multi-page documents or split documents rule-based after a fixed number of pages.



Transform your document management with PLANET AI

Schedule your exclusive IDA demo today and take the first step towards effortless, intelligent document analysis today.

[Book your demo](#)

ABOUT PLANET AI

PLANET AI is a research-driven company dedicated to developing software products with human-inspired cognitive capabilities for information processing and understanding. By utilizing proprietary deep learning technology, PLANET AI empowers organizations to unlock information trapped in documents, seize digitization opportunities, and eliminate manual effort for data capture. The Intelligent Document Analysis software suite offers comprehensive capabilities for customers with the common desire for short time-to-value automation and high-quality data capture, extraction, and understanding.

PLANET AI serves a variety of customers that include Fortune 500 companies, scanning service providers, as well as software vendors in business process automation and content management. Since its beginnings in 1992, PLANET AI has established itself as a global technology leader in cognitive computing. In 2023, German IT provider Bechtle acquired a majority share of PLANET AI.