IDA – Intelligent Document Analysis

Highest Quality Data Capture, Extraction and Understanding

PLANET AI's Intelligent Document Analysis software suite (IDA) offers comprehensive capabilities for customers with the common desire for short time-to-value automation and high-quality data capture, extraction, and understanding.

WHAT MAKES IDA UNIQUE?

Highest OCR and ICR accuracy

- For handwriting and the most challenging scenarios, such as poor-quality scans, historical scripts etc.
- > 99 % accuracy for machine-printed text, > 95 % accuracy for handwriting

Highest accuracy for classification

- For the most challenging document scenarios by analyzing textual and visual features
- > 80 % cost savings due to less manual efforts

Short time to value

- IDA is based on machine learning, not on pre-defined rules
- > 91 % out-of-the-box automation rate

Deployment on-premises

- To ensure compliance with privacy or security regulations
- Optimized for CPU and GPU usage

IDA FEATURES

Thanks to the modular approach, all features can be licensed individually based on your use case.

IDA Recognition is the core feature for data capture in PLANET AI's IDA suite and delivers OCR and ICR results with exceptional accuracy despite scenarios like distorted, poor-quality scans with machine-print and difficult-to-read handwriting.

- ▸ Leveraging the patented PerceptionMatrix
- ▸ Available as a client-server application and as a software development kit (SDK)
- ▸ Versatile JSON output format
- ▸ Add-ons for barcode recognition, historical scripts, and more

IDA Classification offers exceptional accuracy for challenging document scenarios by analyzing both textual and visual features, even when minimal variation is apparent. Furthermore, it provides a rule-free, few-shot learning approach that dramatically accelerates the setup and maintenance of workflows, in contrast to rule-based or manual approaches.
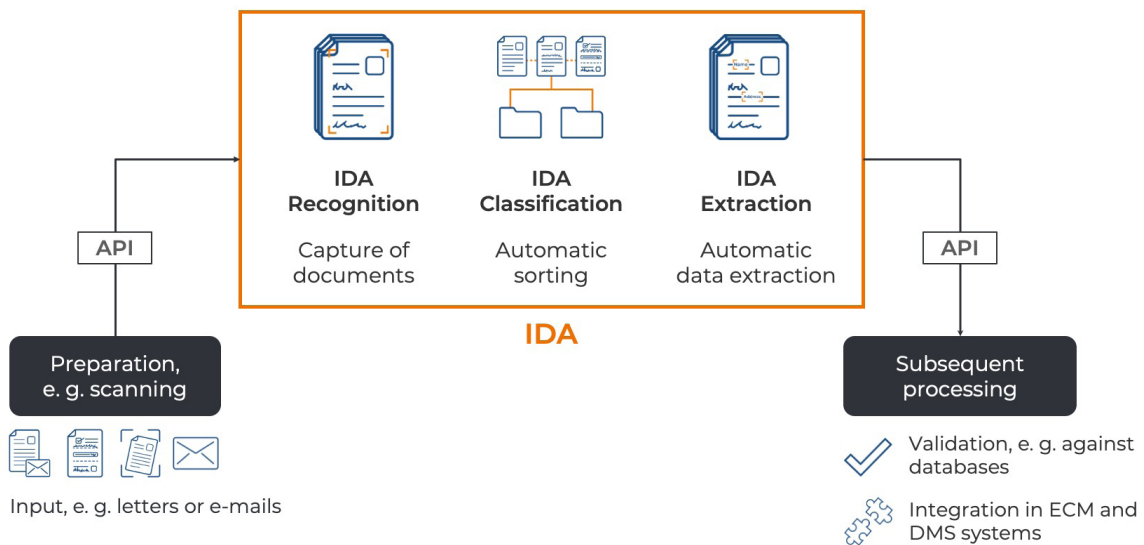
▸ No-code training with a browser-based interface

▸ Intelligent document splitting

▸ Leveraging unmatched OCR quality

IDA Extraction provides advanced information retrieval for both structured and unstructured documents. Utilizing machine learning capabilities like smart zonal data extraction and large language models (LLM), it enables accurate capture of data points within documents. This significantly reduces the need for manual labor in validating results or setting up rule-based workflows.

▸ Refined zonal data extraction for structured and semistructured documents

▸ LLM-based extraction for unstructured documents

▸ No-code solution accessible via browser-based interface

This figure demonstrates how IDA can be integrated into a document workflow:

ADD-ON

When handling large volumes of documents with handwriting, historical scripts, or poor-quality machine-print, even with advanced solutions like IDA, attaining 100% accurate OCR results requires significant manual correction efforts. In such scenarios, **IDA Search** offers complete keyword retrievability, leveraging PLANET AI's patented technology. The PerceptionMatrix ensures that all possible transcriptions of a given text are preserved and accessible through search requests.

▸ Search based on using regular expressions, operators, and filters.

▸ Internationally awarded keyword-spotting

USE CASES

Document Indexing

▶ Critical component of document management systems
▶ Goal: structure and categorize documents to facilitate information retrieval

IDA enhances the indexing process across various document types, significantly reducing expenses for manual labor.

Forms Processing

▶ Essential for a range of subsequent enterprise operations, including document archiving, customer service and accounting
▶ Goal: extract information from pre-defined fields

IDA excels at simplifying data capture from both semi-structured and structured documents.

OEM Integration for OCR

▶ Common practice utilized by software providers to embed optical character recognition (OCR) technology
▶ Goal: seamlessly incorporate third-party software solutions into a company's proprietary products

IDA provides a turnkey product to implement its core data capture feature (IDA Recognition).

Records Classification

▶ Instrumental in governing an organization's documentation
▶ Goal: systematic approach for creating and locating files (file plans)

IDA optimizes records classification with a rule-free approach based on minimal training data.

PLANET AI is a research-driven company that develops AI applications with human-inspired cognitive capabilities, enabling customers to unlock and process information from unstructured documents, automate data capture, and achieve the most accurate results.