Optimizing Records Classification for File Plans with IDA

File plans play a crucial part within **records management systems**, ensuring control over an organization's written materials. A file plan generally provides a **structured framework** for both the creation of new files and the retrieval of existing ones. Complexity can vary significantly, depending on the industry and the use case. File plans include guidelines for classification, folder structure, naming conventions, and retention periods, among other aspects. **Records classification** is a pivotal step when implementing a file plan, encompassing the categorization and labeling of large document volumes.

Explore how the IDA software suite can optimize your records classification.

Reduce manual efforts
Outstanding accuracy for the most difficult classification scenarios

Accelerate time to value and minimize maintenance
Low-effort training for changing document layouts

Ensure compliance
On-premises or private cloud deployment

COMMON CHALLENGES

File plans should not be seen as rigid frameworks; they **require frequent updates and optimizations**. Reasons for updating a file plan may include changes in legal requirements related to document layouts or retention periods. Legacy software solutions often aim to automate the implementation of file plans based on rules. Yet, **maintaining this rule-based records classification** is both complex and costly.

Operational efficiency may be compromised due to the **substantial manual effort** required to classify documents, including validation and correction. Documents often vary only minimally, which can test the limits of rule-based approaches as well as AI-based solutions that rely solely on textual features. Ineffective records management poses the risk of escalating storage costs, unauthorized access to sensitive information, and, consequently, non-compliance to regulations and laws.

THE SOLUTION

PLANET AI's Intelligent Document Analysis offers a **rule-free approach to document classification based on machine learning**. IDA requires minimal training data and thus low efforts in adapting to new or changed document classes. It classifies documents with exceptional accuracy by **analyzing both textual and visual features**, even when there appears to be minimal variation.

IDA delivers **outstanding OCR and ICR results**, particularly in challenging scenarios like handwriting. In downstream processes such as document classification, high-quality input data is essential for making accurate decisions and minimizing manual corrections for increased automation. IDA can be deployed **on-premises or in a (private) cloud** as a Java application or containerization using Docker.

HOW IT WORKS

IDA workflow for records classification:
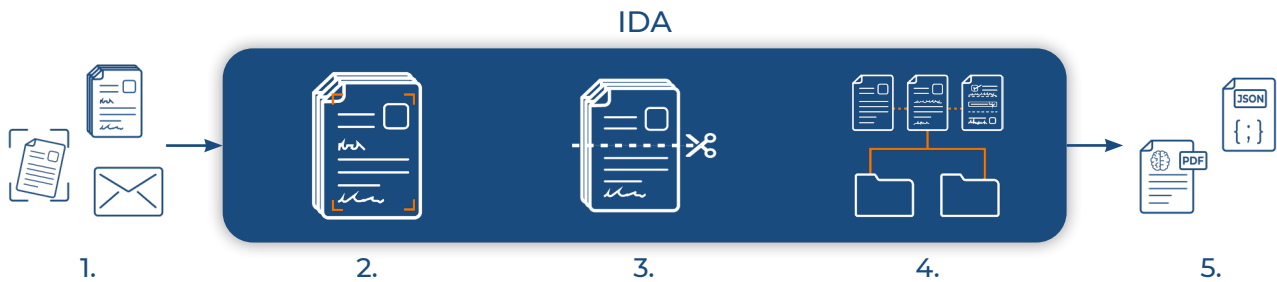
1. Input: Physical and electronic documents via scanner, mailbox, email etc.

2. Recognition: OCR and ICR capability based on patented PerceptionMatrix

3. Document splitting: Automatic separation of large consecutive documents

4. Document (or page) classification: Rule-free document categorization with AI models, trainable with minimal training data

5. Output: PDF or PDF/A (all conformance levels) with text layer containing recognition results and optionally highlighted metadata and/or JSON with positional information, confidence score etc.

Subsequent tasks may involve metadata extraction, including data extraction to capture data fields from records.

CUSTOMER SUCCESS STORY

File plans and records classification find application in various industries and verticals where efficient records management is essential, such as government agencies, the public sector as well as the healthcare industry.

Our renowned client has been offering **business process outsourcing services** to healthcare providers for over 50 years. Regulatory changes necessitated a **significant expansion in document categories**, growing the file plan from approximately 150 to over 300 classes. IDA's rapid and accurate adaptation to the more granular document classes allowed the client to meet the requirements quickly. Moreover, implementing the few-shot learning approach led to an **80% reduction in manual efforts** due to increased straight-through processing.