



DECODING

THE RELATIONSHIP BETWEEN OCR, RPA, AND DOCUMENT UNDERSTANDING

INTRODUCTION

Software bots used for **Robotic Process Automation (RPA)** have automated tedious and time-consuming tasks to a great extent.

Replicating operations that require frequent human interaction with computers, RPA has significantly reduced manual workloads and turnaround time pressures. As such, it frees up the work force to focus on critical aspects of the business such as productivity, growth etc.

Gartner predicts that thru' 2022, 90% of all large global organizations will be using RPA for several operational tasks. Further, with the ability to automate rule-based, repetitive operations, the adoption of this technology is projected to increase to \$1.89 billion by the end of 2021.

Among the many routines and mundane tasks, the manual process of scanning physical documents and using the scanned images to enter data stored therein, is a prime use case for RPA.

Integrating RPA and Optical Character Recognition (OCR) with Machine Learning (ML) models, further drives the gains in productivity, cost savings and reducing waste.

World over, businesses are taking advantage of this technology making RPA bots with OCR & ML the de facto standard for streamlining document handling, processing, and storage and retrieval processes. These cases are varied and covers all aspects of automating the workflow of invoices, scripts, expense statements, bills, shipping/supply chain documents etc. and other types of forms used by the enterprise.

With this background, let us review how RPA leverages OCR technology, and how the integration of RPA, OCR, and Machine Learning models facilitates document understanding and workflows in the business domains

Understanding Key Concepts

Robotic Process Automation (RPA)

Software robots, commonly called BOTs, are a set of coded commands that instruct mechanical or electronic devices on various tasks it is required to perform. Developing and deploying BOTs to automate repetitive and/or rule-based tasks is called Robotic Process Automation.

Robotic Process Automation was engineered to deliver software solutions that could emulate human actions and interactions.



However, with the gradual onset and adoption of technologies like Artificial Intelligence, these robots have now acquired enhanced self-learning capabilities. Machines can now read documents the same way a human mind does!

This positions RPA as one of the most reliable and cost-effective solutions to automate routine repetitive tasks across the enterprise.

Optical Character Recognition (OCR)

Optical Character Recognition technology captures text contained in physical documents and images and converts it into digital formats. These digital assets can then be used to improve information processing and other computer operations.

Coupled with Machine Learning, OCR can deliver error-free results. Furthermore, when integrated with Artificial Intelligence, its ability to interpret the converted data can generate actionable insights.



OCR is the foundation on which Document Workflow Automation and Document Intelligence Platforms are built.

Document Understanding

Systems powered by AI to extract and interpret data from documents automatically is called Document Understanding. It includes the extraction, interpretation, and management of data from any file format - be it PDF, images, handwritten documents, physical printouts, etc.



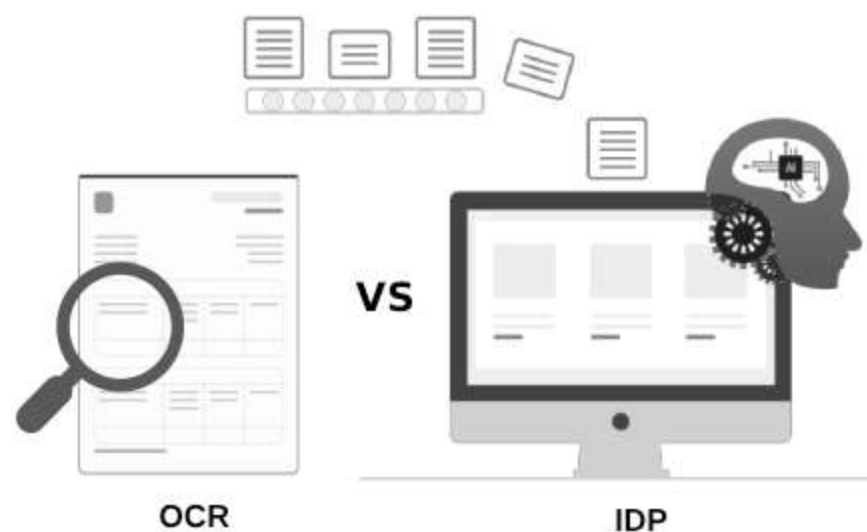
Document understanding is the orchestration of trained machine learning models, software bots and data processing APIs. The process works by enhancing the capabilities of traditional RPA bots by using AI-based OCR technology to enable data extraction from documents without being restricted to templates, formats, etc.

Why is

Traditional OCR

Misfit for

Modern-day Use Cases?



Traditional OCR software is highly dependent on templates. As such, if the template does not match a standard format, the application will produce problematic output to deal with.

A classic example is that of a form. Traditional OCR software may recognize the content of the form, but it fails to accurately convert entries made over blank spaces.

Another issue with traditional OCR is that it doesn't have the ability to validate the correctness of the output and rectify any mistakes on its own.

Lastly, OCR applications tend to get baffled by complex documents. Tables, font changes, headings, etc., aren't accurately converted which upends the format in which data is required.

The Case for ML and RPA

With the inclusion of Machine Learning and Robotic Process Automation, the challenges with traditional OCR technology can be managed effectively. With BOTs, the job becomes more precise and is less time-consuming. Software Robots perform tasks that require human intervention and use requisite applications on our behalf to complete a particular process.



Some extraordinary examples of how RPA has facilitated document understanding includes:



Deployment of an Intelligent Document Processing (IDP) Application Program Interface (API). When a document is downloaded from the web, an IDP API automatically submits the same for processing.



Automating the selection of appropriate APIs to use to extract data and add it directly to a CRM, ERP, or other technology platforms in the organization's ecosystem.

Although RPA delivers significant benefits, its singular setback is that BOTs do not possess the ability to think as humans do. This gap of cognitive ability is bridged by Machine Learning models, that in conjunction with RPA renders the document understanding process both meaningful, and productive. The on-the-fly adaptive nature of RPA & ML combination drives successful outcomes.

The seamless integration of OCR, RPA & ML technologies makes possible the process of document understanding and relevant data extraction a reality.

A very common example of this model is its use in data entry and data tallying operations where OCR expedites data extraction, RPA automates routine and manual tasks and trained Machine Learning models perform intelligent actions such as deleting duplicate entries and errors.

The Workflow of Document Understanding



Retrieving Documents

The first step in the workflow is to access a document that may reside in local storage or in the Cloud. Software bots fetch the relevant documents required for extraction and processing.



Type Of Documents

A mechanism to classify retrieved documents into its type like invoices, expense receipts, trade documents, shipping documents etc. is triggered.



Classification

With RPA and IDP one can easily automate the process of document classification and also decide very precisely as to which approach of data extraction would be best suited for the process.

Further, once the analysis of the extracted data gets completed, the metrics obtained will be studied and stored to train the models for future reference.



Data Extraction

The next step is to convert and digitize the text in the documents using an OCR platform. The extraction of data may not be full-proof and free of errors, depending on the ability of the software itself.

It calls for a mechanism that can help in correctly analyzing the nuances of every document, assisting in standardizing the documents in the right way. It is here that Machine Learning plays a significant role.



Document AI and Use of Bots

Document AI performs an analysis of the output that it has generated once the extraction gets completed. It becomes important to rectify the errors and develop insights from the data to ease the job of the end-users.



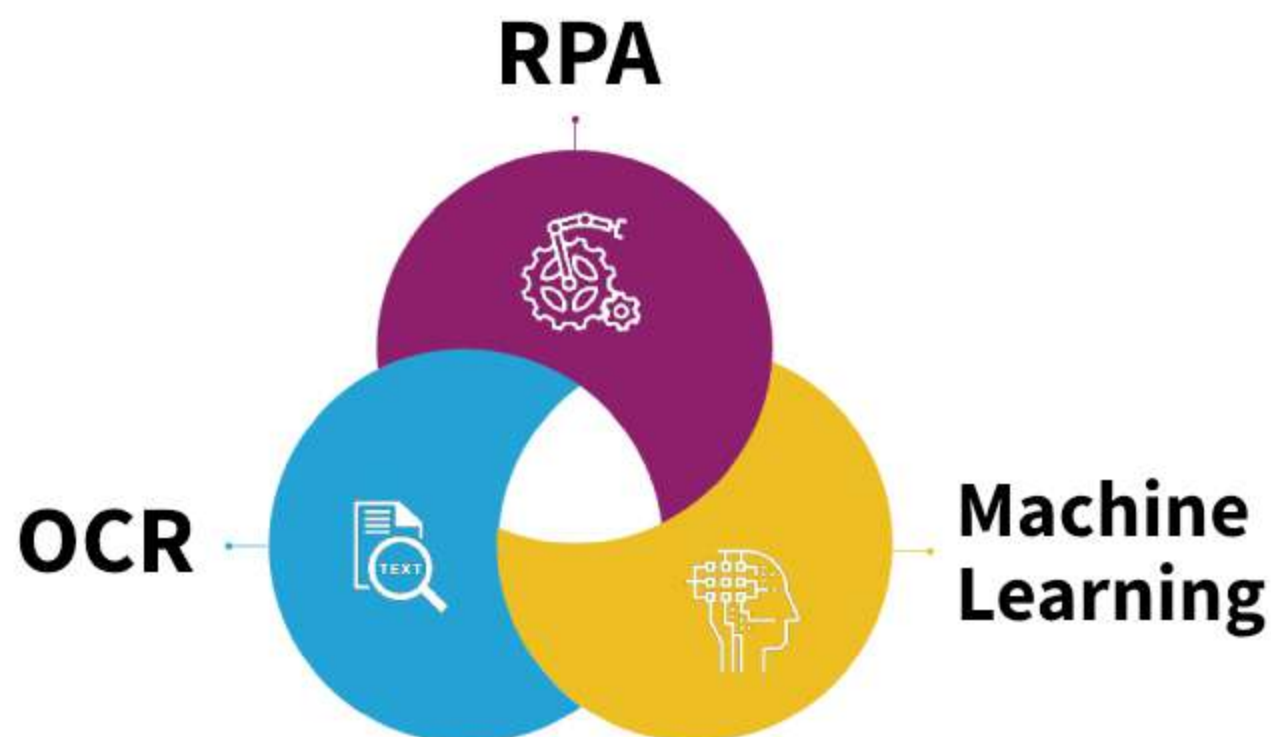
Final Validation

Sometimes even bots may not provide high accuracy owing to the complexity of the data involved. Hence, there is a final mechanism where the bots raise a request for validation such that a final human recheck is done to ensure that the extraction and interpretation is as per standards.

The steps for final validation are:

- I. The steps involved in this process start from a rule's engine in the IDP platform.
- II. Reconciliation of the data with existing records is the next procedure, which can be automated by bots.
- III. Any shortcomings in these procedures call for adding documents to the exception queue to allow final human validation.
- IV. All documents that pass the reconciliation process are then sent for approval.

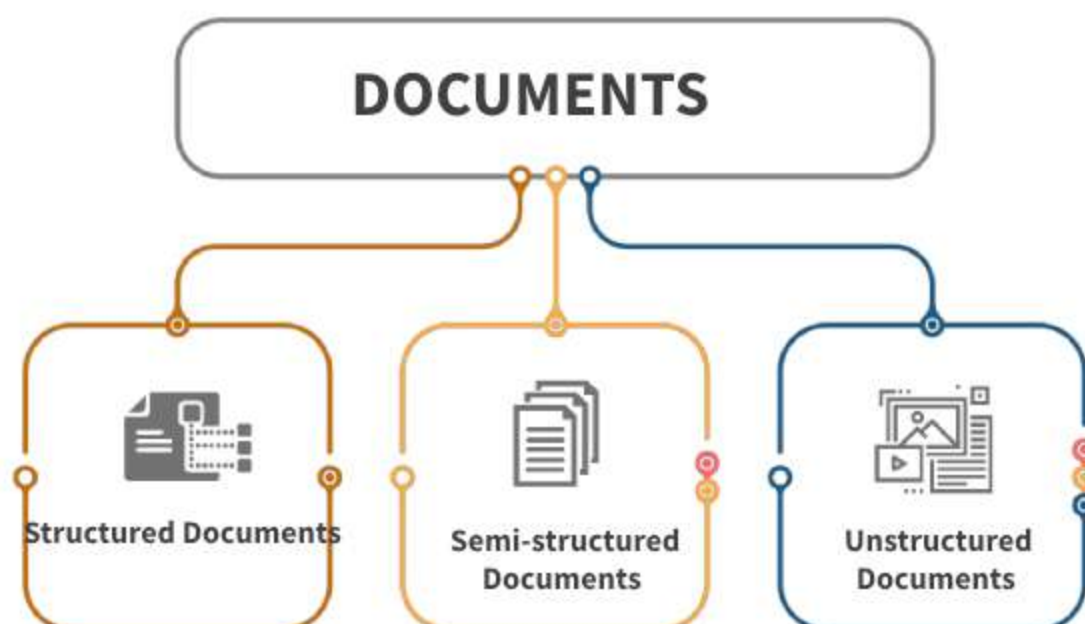
Implementing Document Understanding



Organizations generate data that is stored in a variety of templates, formats, styles, etc. As such the workflow of an integrated OCR, RPA & ML solution is not standardized and has to be customized to handle the variables in data and document types.

The document type dictates the method best suited to extract data from it. Hence, different methods are used that enable software bots to automate the data extraction process irrespective of the type of the document.

Generally, documents are divided into three types- Structured, Semi-structured, and Unstructured.



Structured Documents

Structured documents are created as fixed and standardized templates.

For instance, a government form like the IRS-1040, will have the exact same fields and line items for any person using it to file his tax returns. Hence, extracting data out of many such forms can be easily automated by using a rule-based approach that remains uniform for the entire process.

However, minor changes or deviations from the standard template skew the output. Therefore, rigid template-based OCR is becoming obsolete.

Semi-structured documents

Semi-structured documents contain the same information through the form, but where the information appears in the form may differ slightly. Fixed forms are structured documents wherein the data to be extracted is in the exact same position on the page. Similarly, semi-structured documents where data does not appear in fixed positions are called free-form documents.

Consequently, a rule-based approach will not reap accurate results for free-form documents since consistent uniformity is absent in the documents. It is here that Artificial Intelligence comes into play to process and extract information out of complex documents.

Machine Learning models are in a continual learning mode to decipher these variations in data positioning. After the initial learning phase, the accuracy of data extraction improves owing to the element of dynamism.

Hybrid models work best in cases where a balance between rule-based and model-based approaches gives the desired outputs.

Unstructured documents

Lastly, there are unstructured documents where no standard template is being followed.

Here, using a rule-based approach will not work because there are no repetitive characters in the documents. Hence, RPA does not allow users to extract meaningful data out of such documents.

Thus, software bots need to use AI-based OCR to extract and convert it into a structured form. Training ML models to recognize different patterns and translate written text into actionable data is essential for processing data in unstructured documents.

The use of Named Entity Recognition, Natural Language Processing (NLP), and Word Embedding makes it easier to process such information.

Conventional Ways of Document Interpretation



Manual Document Processing

The manual method of document processing relies entirely on human intervention. Users peruse the entire document and then manually enter the required data into internal systems.

Data thus entered is then routed through the approval process and when okayed, the final action is taken. Quite evidently, data entry processes increase the effort, time and cost burden of the entire operation.



Template-based OCR for Document Processing

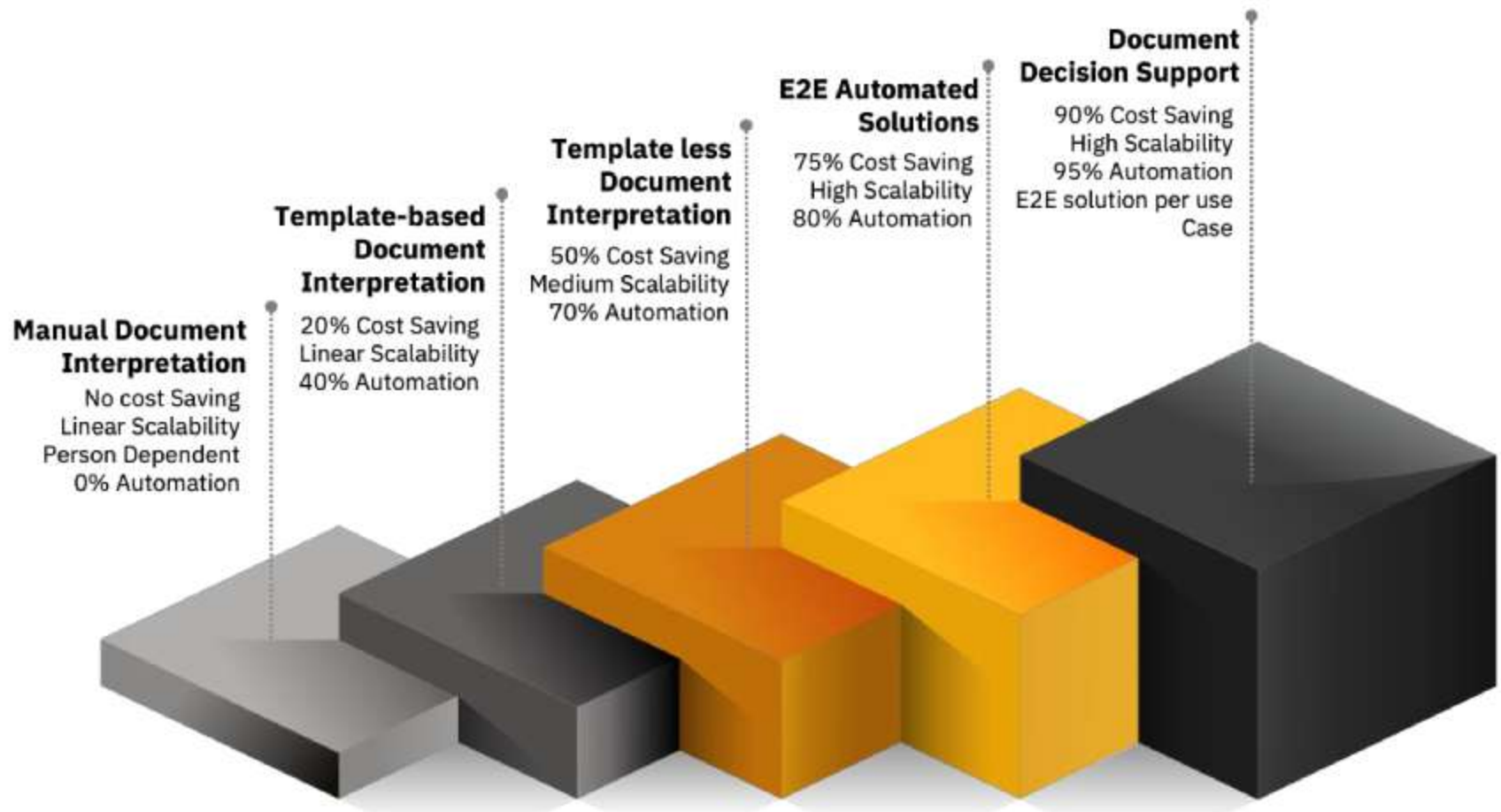
Template-based OCR technology is a step towards automation from conventional manual processing of documents.

Here, the basic premise is that with the use of template-based tools, you can create a particular layout for similar documents.

Further, the integration of these template-based tools with traditional OCR allows for easy data extraction. This method is faster and more accurate than manual entries.

However, the issue is that unstructured documents cannot be processed using template-based OCR technology.

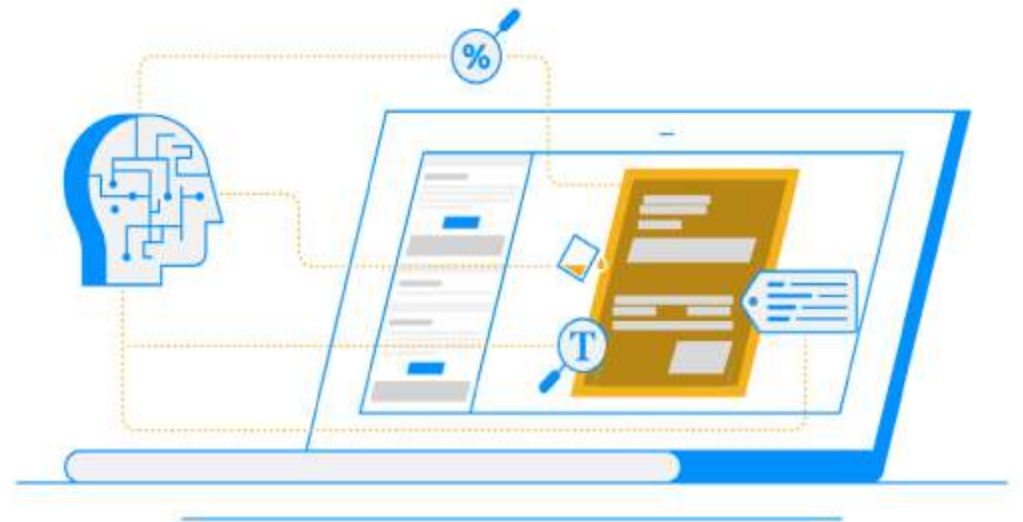
The Real KlearStack Game Changer



Template-less Document Interpretation

Template-less document processing is the next generation technology organizations can leverage to process unstructured documents.

Template-less data interpretation depends on Artificial Intelligence. With AI and machine learning models, the capability of the system to learn-as-it-processes scales significantly.



With such self-learning capabilities, the model is trained to accommodate variances between documents, thereby producing usable outputs.

Industry surveys report that template-less data extraction solutions can deliver up to 50% in cost reduction and bring 70% automation to document processing operations. These solutions also provide scope for non-linear scalability, making it a solid choice for data-related tasks.

End-to-End Automated Solutions

Scalability requires solutions that automate every step of document processing from capturing, classification, splitting, and validation, thru' enrichment to enable document hyper-automation.



This eliminates human intervention from the document process freeing up workers to focus more on critical aspects of the business.

Additionally, KlearStack's E2E solution incorporates Electronic Document Interchange (EDI) to expedite data transfer and sharing with applications in the eco-system.

Document Decision Support System

The topmost form of document understanding is the Document Decision Support System. These are computer-based document management systems that use communication technologies and models to support decision-making activities.

Guaranteeing phenomenal levels of automation and cost savings, Document Decision Support Systems streamline the workflow for organizations of all sizes. Provisioned with E2E solutions, the extent of business scalability realized is superior to any other document understanding solution.



The Document Decision Support System provides access to all critical information assets, generates comparative data figures and insights, and analyses and informs implications of choosing different data-related decision alternatives.

KlearStack @ Work, Works

KlearStack is a G-local multiple awards winning solution used by Global Brands and Fortune-100 companies across many verticals. Some of the use cases include:

BANKING



Many renowned banks and financial service providers have leveraged the AI-based OCR services of KlearStack to expedite real-time data extraction and automate several critical processes related to KYC, accounts payable, and loan sanctioning.

TRADE FINANCE



Trade finance involves the sharing and distribution of millions of paper-based documents. KlearStack's AI-based OCR solution automates the document processing for transactions, document reviews, compliance checks, application forms, etc.

INVENTORY MANAGEMENT



KlearStack automates the vital task of managing inventory data with its IDP and OCR solutions. The RPA services automate error-free inter-system reconciliations.

INVOICE PROCESSING



No templates, no formats. Process and extract data from invoices and manage payments more effectively.

The KlearStack Advantage

- Template-less, end-to-end automated data extraction solutions.
- Guaranteed maximum degree of automation achievable for document understanding.
- Customized interpretation technology for business insights.
- Improved Work Efficiencies and Productivity.
- Reduced Cost, Time, Effort, and Waste.
- Reasonable TCO and Low Payback Period.



www.klearstack.com