



A Guide on **Document Understanding**



Contents

03	Introduction
04	What Is Document Understanding
05	Get Documents Processed Intelligently
06	Document Understanding Framework
08	How to Automate Document Processing
08	Structured Documents
09	Semi-Structured Documents
09	Unstructured Documents
10	Handwriting, Signatures, Checkboxes, and Other Special Cases
10	Guide on How to Choose the Best Solution
11	Machine Learning Models in Your Business
12	Conclusion
13	References

Introduction

If there's one universal challenge that every industry deals with, it's the buildup of paperwork. According to Forbes, as much as 2.5 quintillion bytes of data are created every day.¹ That's a lot of data in need of processing! From human resources (HR) forms to invoices and contracts, businesses across the world spend countless hours reading documents, pulling information from them, and making sure the correct action is taken based on that information. It's a daily routine, forcing employees to focus on manual, time-consuming work instead of the higher-value tasks that drive business objectives and company growth.

Fortunately, there's a solution. You can minimize the manual, routine work of data processing and get your employees back to higher-value tasks.

Advances in Artificial Intelligence (AI) have led to the development of software robots that can read documents of any difficulty, pull desired information, and pass it to where it needs to go--with a high level of accuracy and reliability. This enables automation of various document-intensive business chores which usually require highly-manual processing. Robotic Process Automation (RPA) can cope with these routine business tasks with essential time and cost savings. In turn, document understanding, also referred to as intelligent document processing, is one of the fastest-growing segments of the RPA industry. In particular, Gartner predicts that RPA

can save finance departments 25,000 hours of avoidable paperwork annually.²

The automation of document processing has many business applications. Some of the most common use cases include invoice processing, insurance claims, new-patient onboarding, proof of delivery, and order forms. It can also be used for more advanced forms of data processing such as multilingual document processing and mobile capture.³

Intelligent document processing brings several key benefits:

Accelerate operational efficiency: Spend less time and costs on high-volume document processing and more on the real business investments.

Improve customer experience: Accelerate the speed of your responses to customers in order to win their respect with exceptional customer experience.

Mitigate risks: Eliminate the risk of human error to save you and your customers from related losses.

Improve employee experience: Save employees from operational mundanity and let them focus on higher-value tasks and objectives.

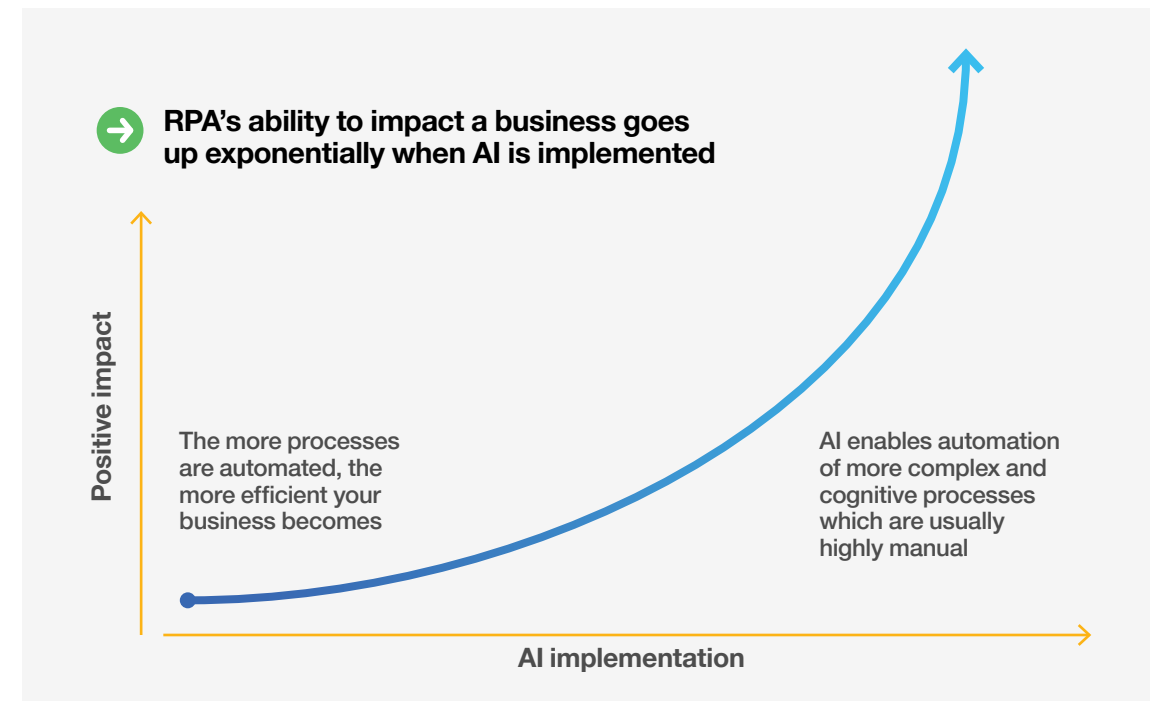
According to a recent publication by Everest Group Research, owing to automations with intelligent document processing, businesses have seen a 9-17% reduction in non-people costs

such as facilities and technology. The study reports that this implies USD 200,000 – 400,000 saved per year.⁴ This reduction in costs helps fuel a positive return on investment (ROI).

In this white paper, you will learn how document understanding:

- **Reads and processes** structured documents, such as forms or passports.
- **Understands and draws information** from semi-structured documents, such as invoices or receipts.

- **Can help you process** unstructured documents, such as contracts and emails.
- **Copes with special elements**, such as handwriting, signatures, checkboxes, and tables.
- **Automates** the processing of documents, saving time and resources for your teams as well as helping to decrease costs and risks of human error.
- **Integrates** with our rich partner ecosystem, allowing you to access cutting-edge technology incorporated into the UiPath platform.



What Is Document Understanding?

Document understanding is the term used to describe reading, interpreting, and acting on document data automatically. What's most important, all this is performed by software robots. Intelligent document processing technologies incorporating machine learning capabilities, allow the robots' artificial intelligence to understand documents as digital assistants. This way, we can say that document understanding emerges at the intersection of document processing, AI, and RPA.

UiPath has designed Document Understanding – an AI-enhanced solution developed to help organizations read, understand, and process different information within documents automatically – with the help of robots. We've spoken about how automation of this process can bring a wealth of benefits to businesses, but let's delve a little deeper.

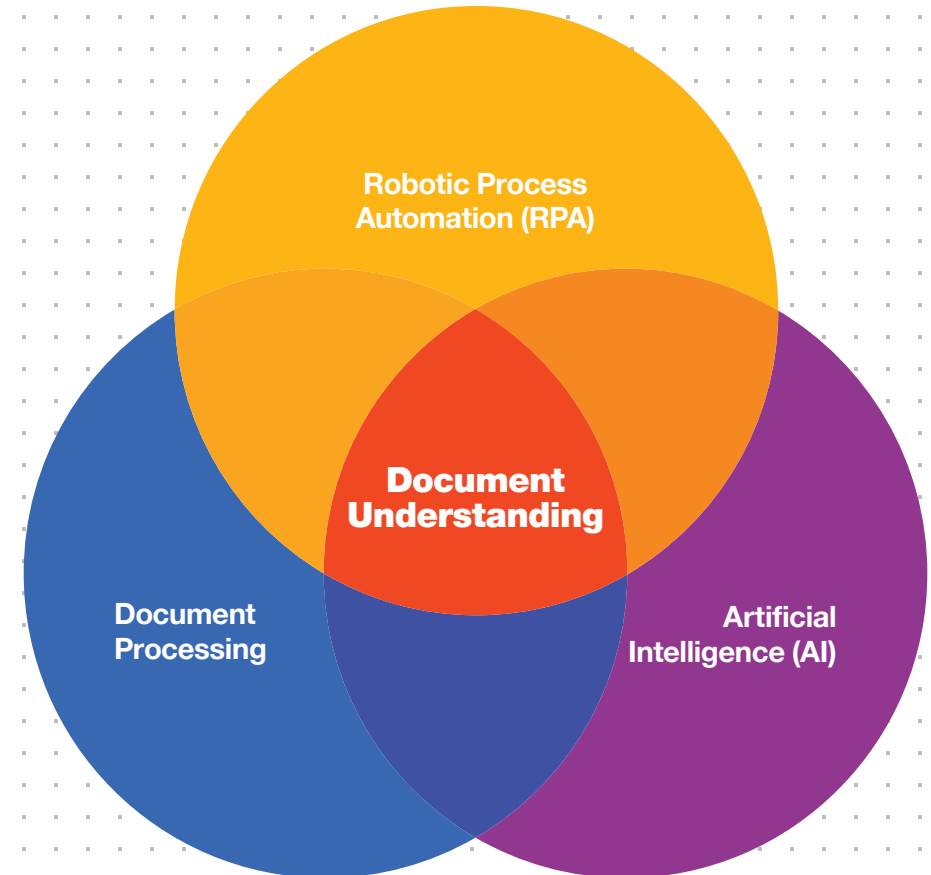
Think about how documents are typically processed. Both humans and robots need to read them, understand the content, and then act on it. With some documents, like a passport or a simple form, it's easy to spot the data and process it. If you need to find out if a person has American citizenship, you can quickly find it in their passport. But with other documents, like legal forms or bank statements, it can be difficult

and time-consuming to validate the information. Just ask anyone who's had to review tax or legal forms how much time it takes! Now imagine how much time people need to process high-volume documents or those with no fixed structure, like contracts. It can be a difficult and imperfect manual process.

Let's consider how important information kept in our documents impacts our business. Invoices, receipts, contracts, and bills hold critical, detailed information. This information impacts our profitability and efficiency. This is a challenge any business may encounter – and it needs a solution – a system of document automation that enables consistent interpretation of this data and allows businesses to divert human resources to valuable tasks.

UiPath Document Understanding addresses complex document processing and eliminates manual extraction headaches by having robots read, understand, and act on documents using their automated intelligence skills. These robots can work at speeds and with a high accuracy eliminating the risk of errors. They automatically determine the location of document data, even when the type or format of the document changes. The robots will read documents with the accuracy your business requires.

➔ **Document understanding emerges at the intersection of document processing, AI, and RPA.**



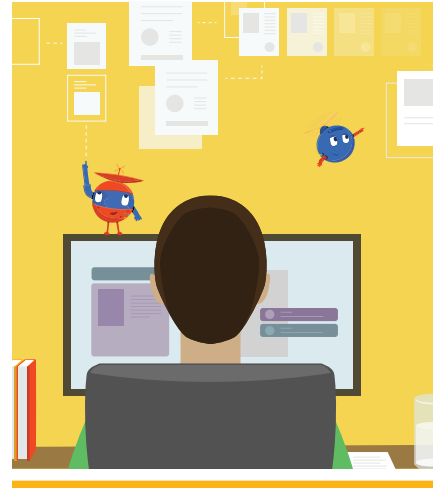
Get documents processed intelligently

Teach robots to understand documents using intelligent drag-and-drop skills for data extraction and interpretation. UiPath provides flexible technology for fast and accurate document processing.



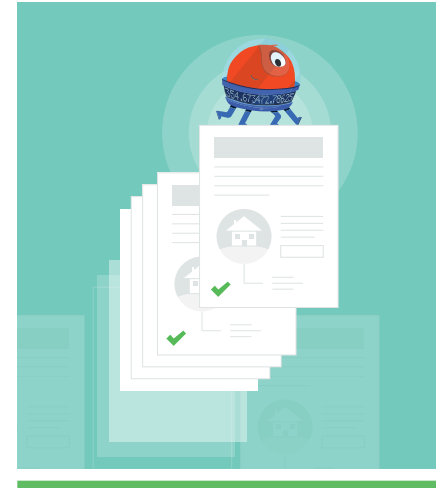
Flexibility

A combination of different approaches help process various types and formats of documents. Even if the format of the document changes, AI can still determine the location of data.



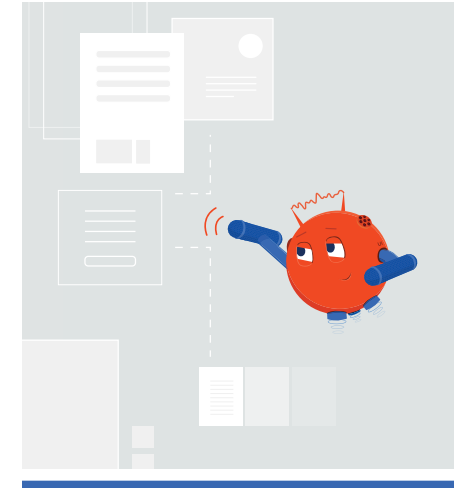
Intelligence

The robots can understand and act upon documents, even if those are rotated or skewed. They can ask for employee validation if needed and will automatically learn from the human-validated data.



Accuracy

The more you work with these models, the more trained and accurate they become. This eliminates the risk of human error, thus, ensures compliance and reduces time employees spend on rework.

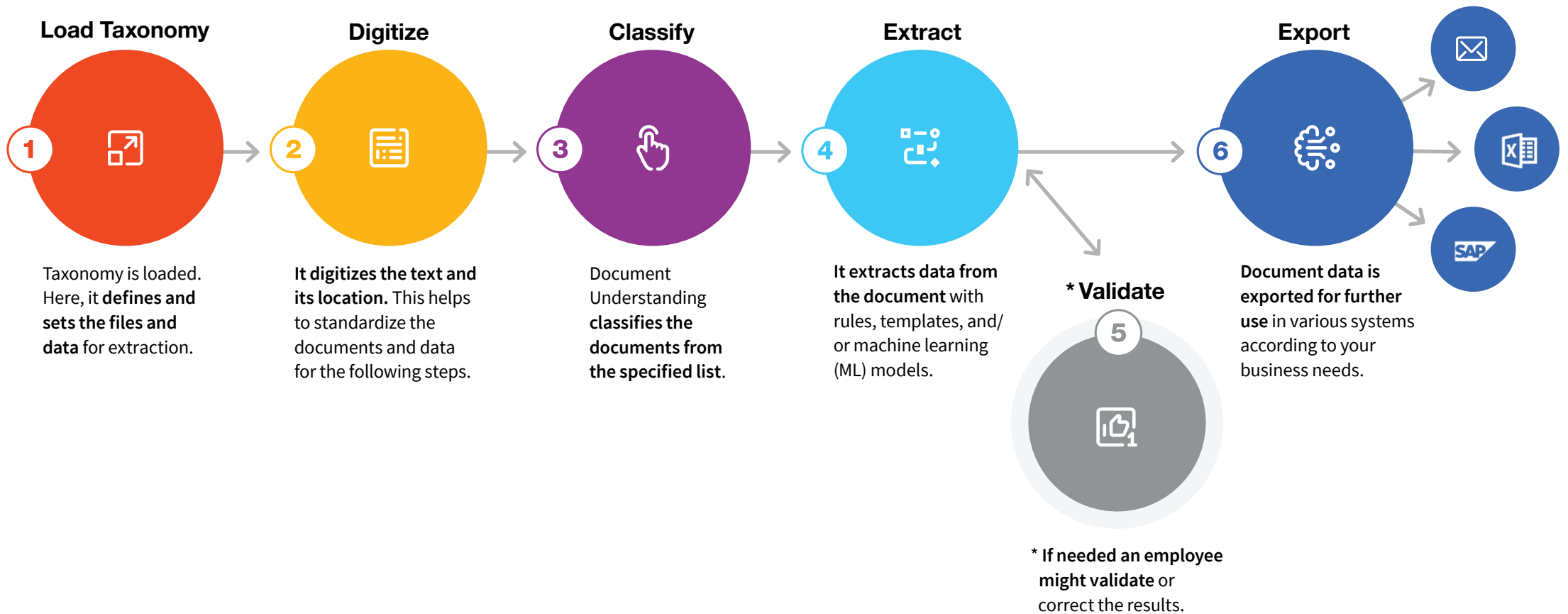


Fast

You can seamlessly automate high-volume document processing for a wide range of documents to save costs and time with end-to-end automation.

Document Understanding Framework

UiPath Document Understanding automations operate according to the framework shown here:



Let's take a closer look at extracting document data, a crucial component of the Document Understanding framework.

When extracting document data, you can take three different approaches. UiPath Document Understanding can use either a rule-based or a model-based approach, or those can be combined into a hybrid approach. Different approaches are used based on the types of documents being processed. For example, documents with fixed positions of data, or as we'll later call them, "structured documents," will need a rule-based approach. Documents with varying templates and less structured data, or "semi-unstructured" documents, will require an ML model. Hybrid solutions are crafted for documents that are mostly unstructured or require a special approach.

A rule-based, or template-based, approach takes actions based on specific, user-specified rules. Think of when you block someone's email address to avoid getting emails from that one individual. When you click that "block sender" button, you are creating a rule. Or imagine a query in a simple search functionality – like in a Word document or in a PDF. Each time a query is entered, these search functions execute the search based on the user-specified rules.

Machine Learning (ML) models, on the other hand, learn how to respond to dynamic situations. This means they are taught how to find and extract data when no static rules or templates can be applied. This is done by exposing them to a dataset and telling them how to act in different

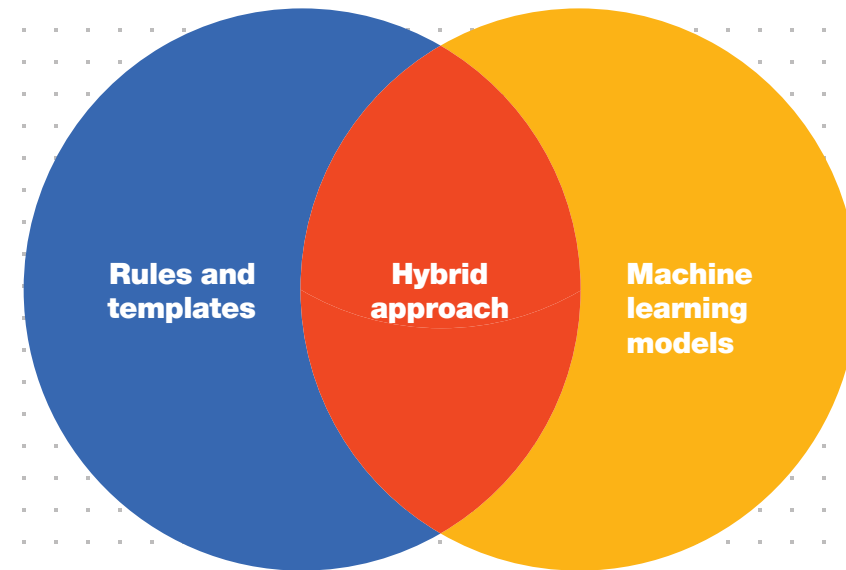
situations. Once taught, the ML models can apply what they've learned to unforeseen events. Let's go back to our email example. Yes, blocking an annoying email address is good. But do you know what's even better? A spam blocker – an example of ML being used to filter out hundreds if not thousands of original emails after learning how to recognize spam email from non-spam email. Now, let's imagine our search engine scenario again. While rudimentary search engines scan for relevant data based on rules, advanced search engines like the ones in Google or Bing use ML. This way they can identify trends in queries, synonyms, related topics, and help find content that is relevant to the user, although it may not be the same as the user query.

In fact, ML models need time to learn. Fortunately, you can make use of out-of-the-box models that have been pre-trained to process common and not-so-common documents. They can be also retrained based on your custom data – the more you work with the model, the more accurate results you'll get. Additionally, UiPath can help you build custom models for your specific business needs.

In some cases, you need to combine rules, templates, and ML models to ensure more effective data extraction. This is called **a hybrid approach or multi-approach**. A hybrid solution draws upon specific, pre-determined rules to extract structured data, while also using ML models to recognize and process less structured parts of a document. Correspondingly, this combination grants higher accuracy. Different extractors are set up separately for each



When extracting document data, you can take three different approaches



document or even field. As seen below, the hybrid approach allows you to combine any extractors within the Document Understanding framework. A minimum confidence threshold is set for each document or field to define which extractor is optimal for a specific part.

Some structured documents may need much more than just rule-based methodologies as some data cannot be extracted with the help of rules or templates. Likewise, solely model-based methods do not work for all unstructured and semi-structured documents. We highly recommend using a multi-approach data extraction when you are dealing with

complicated documents and want to achieve the highest levels of accuracy during the data extraction process.

But how do you decide which approach to use – rules, templates, ML models, or a combination of those? As we've already mentioned, all documents can be divided into three categories – structured, semi-structured, and unstructured. It's crucial to understand the difference between those since the processing of each of these categories requires different approaches – and we can guide you through document understanding to automate document processing intelligently.

How to Automate Document Processing

Using both rule-based and model-based approaches, we're able to handle data from different document structures. A document's structure type refers to how much of a document is set and fixed from sample to sample. Whether your documents are structured, semi-structured, or unstructured, UiPath Document Understanding can teach robots to read, analyze, and process your documents.

Structured Documents

In structured documents, the **structure and layout of the document are fixed**, information fields are labelled, and positions are consistent. Think of the forms you fill out at a doctor's office. Every new patient's form is identical, with the same "NAME" and "ADDRESS" boxes in the same places. There may be cases such as checkboxes, where the important data is not labeled through text, but in symbols or marks. Only the answers vary. This is a structured document.

For these types of documents, the recommended approach is to use a template or rules to extract the information. This provides the highest accuracy and is easiest to implement.

A rule-based, or template-based, approach uses a set of rules to determine what data to read and collect. For example, a rule can be established letting the robot know that the first box of every patient's form will be the patient's name, and numbers in the top right corner will be the date. In other words, you can set up specific rules to create a template. When using

UiPath Document Understanding, you can either use our preexisting document templates or create your own rules to read any structured data.

Let's think again about the new patient forms we've been discussing. A doctor's office can use Document Understanding to automate the process of taking information from new patient forms and sending it to each patient's insurance company. Knowing that patient forms are structured, we can establish rules that tell the robot where to find the information insurance companies need and how to send it to the insurance companies. This frees the nursing staff to care for patients instead of dealing with insurance documentation. In particular, this saves money by reducing the number of hours spent on repetitive paperwork, boosts revenue by increasing how quickly patients move through the office and how many can be seen in a day, and mitigates the chance of human errors.

Rule-based approaches are easy to understand, easy to implement, and quick to deliver your ROI.

While rule-based approaches are perfect for structured data, they struggle with less-structured documents. Because a rule-based solution relies on fixed parts, even minor changes in form structure can cause rules to break down. For instance, remember the insurance company in our doctor's example? The insurance company might be getting patient information from hundreds of different doctor offices in hundreds of different formats. A rule-based tool would require a hundred different rules to understand it all, a level of preparation that is not feasible.

That is why UiPath Document Understanding offers other solutions for processing less structured documents. Let's begin with semi-structured ones.

Types of Documents

Structured

Repetitive forms and documents with a structured template that can contain handwritten text, signatures, checkboxes (forms, passports, licenses, time sheets)

Semi-structured

Documents containing fixed and variable parts like tables (invoices, receipts, purchase orders, medical bills, bank statements, utility bills)

Unstructured

Documents with no fixed format (various contracts, agreements, emails, disease descriptions, drug prescriptions, news, voice scripts)

Semi-Structured Documents

Semi-structured documents have all the same information, just arranged in different locations on different documents. Think of any company that lets employees expense their lunches on business trips. If the company is big enough, it could receive tens of thousands of receipts every year from all sorts of countries and restaurants. The employees in HR could process the expenses manually, but it is tedious low-value work. You would just be asking for someone to misplace a decimal and reimburse \$1,000 instead of \$100. Instead of relying on humans solely, you can train a robot to read the amount that needs to be reimbursed and either make the reimbursement or let management know if someone is spending beyond their limits. You can also program the machine to only act automatically if it meets a certain threshold of confidence that it is correctly understanding the document. If the confidence level does not meet that threshold, the document is sent to HR for review.

Handing the task over to a robot alleviates HR hours spent reading lunch receipts, frees their time to spend on

value-driving activities, and reduces the chances of error. Semi-structured documents require a bit more customization than user templates or rules-based approaches, but that's why **a pre-trained or customized ML model** is usually recommended. Alternatively, in some cases you can opt for a **hybrid approach, that is to combine both rules and ML models for extraction.**

ML models though, are not without some drawbacks. They are non-deterministic, meaning all documents will be returned with an accuracy or confidence score. Because the model is learning patterns, rather than operating off concrete rules, it may make initial mistakes right after corrections. The solution to these drawbacks, however, is simple – the more samples the ML model processes, the more patterns it learns to ensure accuracy.

Unstructured Documents

Unstructured documents are complex with no fixed structure at all. The main difference between a semi-structured and unstructured document is the absence of keys for extracted values. Keys are consistent fields to look for in a document, such as a vendor name, date, price, amount, total, or currency in an invoice. Keys may change position from document to document in a semi-structured document, but in an unstructured document, there are no keys to look for at all.

An example of an unstructured document would be customer service request emails. In this example, each email is going to be unique. For ML models to accurately process these emails, the robots need to learn how to translate written text into actionable data, like a phone number. In a structured or semi-structured document, phone numbers are identified by the field "phone number." But for an unstructured document, the model will need to learn that specific 7- or 10-digit number patterns should be extracted as phone numbers. This form of data extraction requires robust ML models that are tuned specifically

for this task, possibly combined with the hybrid approach if some parts of the document are more structured. In more complicated cases, this is where our partnership network comes into play.

For unstructured documents, UiPath Document Understanding can be paired with our rich ecosystem of partners to meet your needs in unstructured data extraction. Partners like Indico have models capable of processing unstructured documents that are deployable through the cloud or on-premise. Similarly, you can leverage our **document understanding ecosystem**⁵ to cover any other business needs in document processing.

Quick Reference Tool

Rule-based approach

Takes actions based on predefined rules or templates. They can be set up in days, but as needs grow more complex, required maintenance time can become cumbersome.

Machine learning models

Use AI and a trained dataset to apply learnings to new dynamic situations. They take longer to set up, but are more scalable over time and as complexity increases.

Handwriting, Signatures, Checkboxes, and Other Special Cases

Some documents turn the difficulty up a notch by featuring handwriting, signatures, checkboxes, tables, or non-Latin alphabet characters.

We're ready for that possibility. UiPath Document Understanding offers solutions that can be used to extract document data from these special cases.

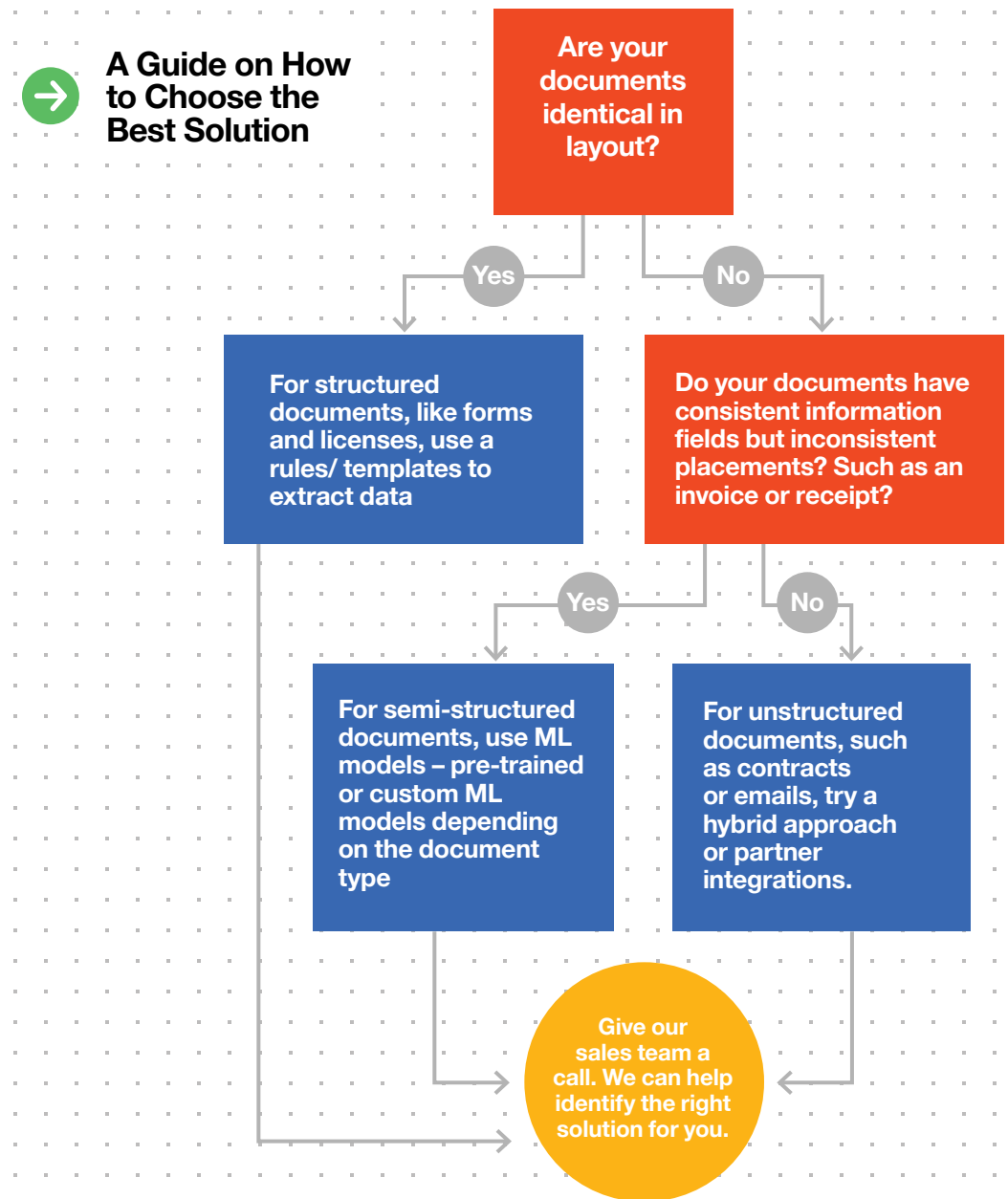
For example, human handwriting is unique to each individual. This presents obvious challenges for standard Optical Character Recognition (OCR) engines picking up typed text. In these cases, UiPath Document Understanding applies special built-in technology for handwriting, enabling it to be extracted letter by letter. While the process can be more difficult and time-consuming, it's necessary to extract the data with a high degree of confidence and accuracy.

For documents that contain special cases such as signatures and checkboxes, Document Understanding

can detect when a document is signed, or when a checkbox is checked. Similarly, it can extract data from tables keeping the right correlation between the text fields and corresponding values.

Another special case is identifying characters from a non-Latin script or alphabet like Japanese or Chinese. Because ML models need exposure to many samples before finding patterns, foreign languages can present challenges. UiPath Document Understanding works with most common languages, and support for both Japanese and Chinese scripts is expected in our next releases. Users who need to extract characters from Middle Eastern, Asian, or Indian scripts can leverage our diverse network of partners for this functionality.

Because there are so many variables that go into choosing the right model, we have outlined a simple visual guide on choosing the best solution for your document processing needs.



Note: This flowchart offers a high-level overview of the existing technology. The decision on which technology to use should be based on your specific use case and business needs.

Machine Learning Models in Your Business

What makes Document Understanding so powerful as a data solution is its ability to bring the versatility of AI to RPA. Because your ML models will learn as they encounter more document samples, the speed and accuracy of your document processing automations will compound over time, as we can see in the Chevron example.⁵ Owing to automatic document understanding, **Chevron achieved:**

- **Faster processing** times for long, detailed drill reports.
- **More resources** for analysts to focus on higher-value tasks.
- **More available information** for business leaders and decision makers.



After using document processing automations to enhance their workflows surrounding drill reports and data, Chevron experienced faster processing and a freeing up of employee resources.

For a world-wide beverage manufacturer, UiPath Document Understanding enabled robots to **quickly read beverage and snack product labels to ensure that details were correct before they were printed** on products. This small task helps this company accurately print labels on over 20 billion products a year.

Bringing the power of ML models and AI-enhanced document processing to your business can bring both measurable results and fast ROI. According to *Everest Group Research's Smart RPA Enterprise Playbook*,⁶ businesses that implement AI-enhanced automations see a multitude of benefits.

The first is often a dramatic reduction in total cost. As processes are automated, facility and technology costs can decrease as much as 33-38%.⁷ Employee expenses can also be reduced by as much as 27-40% after implementation.⁸ Altogether, implementations were found to deliver a steady net savings of 25-40% in addition to a 3-year risk-



The process's accuracy mitigates the company's risk of incorrectly printing nutritional labels – a lawsuit waiting to happen that could damage the company's reputation.

adjusted return of investment (ROI) of 250% and an expected payback period of 12 months.⁹

Cost is not the only benefit. Businesses also reported an increase in process accuracy, improvements in creating new business models, and improved customer experience. It's clear that an investment in ML modeling and AI-enhanced document processing provides short term and long term benefits to your business.



Your machine learning models will learn as they encounter more document samples, and the speed and accuracy of your document processing automations will compound over time.

Conclusion

Every day businesses spend a lot of time reading, understanding, and processing document data. We've demonstrated how automating this process can save businesses both money and time, not to mention mitigate risk and uncertainty. Automating your document processing needs with UiPath Document Understanding creates a solution carefully crafted to fit your business needs.

We understand that your documents may come in many different formats, such as structured, unstructured, or semi-structured. We know that they may have special parts such as handwriting or checkboxes. We also realize that you may want a solution that provides for the full automation lifecycle or that can be used out-of-the-box. UiPath Document Understanding has been built to address all these needs.

To handle various formats of documents, Document Understanding offers rule-based, ML-based, and hybrid approaches to suit each business case. We can help identify which pre-built model fits your document processing challenge to find the ready-to-deploy solution you need. We can retrain or build a custom model from scratch for original challenges and opportunities. And for unstructured documents or less common complex cases, we can combine the best of rule-based templates and ML models, drawing on our rich partner ecosystem to help. See for yourself – [sign up for the Enterprise Cloud Platform trial to test out Document Understanding capabilities.](#)

[Start Cloud Platform Trial](#)



To handle various formats of documents, Document Understanding offers rule-based, ML-based, and hybrid approaches to suit each business case.

References

¹<https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/#50c418f060ba>

²<https://www.gartner.com/en/newsroom/press-releases/2019-10-02-gartner-says-robotic-process-automation-can-save-fina>

³<https://www.everestgrp.com/2019-06-intelligent-document-processing-market-is-exploding-expected-to-grow-70-80-over-next-2-years-press-release-50323.html>

⁴<https://www.uipath.com/company/rpa-analyst-reports/everest-smart-rpa-playbook> page 145

⁵<https://www.uipath.com/blog/leveraging-document-understanding-ecosystem>

⁶<https://customers.microsoft.com/en-gb/story/chevron-mining-oil-gas-azure-cognitive-services>

⁷Everest Group Research Playbook

⁸Everest Group Research Playbook page 155

⁹Everest Group Research Playbook

¹⁰Everest Group Research Playbook page 155



With over 25 years of business transformation experience, Auxis partners with leading organizations to accelerate their end-to-end intelligent automation journey with UiPath. Through strategic roadmaps, Auxis aligns efforts with client priorities to maximize ROI and targeted outcomes. Auxis' unique delivery model leverages onshore and nearshore talent from Latin America for seamless collaboration. Auxis is a proud UiPath Gold Partner, UiPath Platinum Competency Partner (based on number of certified automation professionals), and UiPath Professional Services Certified. To learn more, visit www.auxis.com

