UiPath Document Understanding

Every company in the world has documents and they create more of them every day. An efficient way to deal with high volumes of documents is to automate. Finance departments alone can save 25,000 hours of rework caused by human errors at a cost of $878,000 per year for an organization with 40 full-time accounting staff, according to a Gartner press release. That's just one department. Imagine how much can be saved if you also automate document processing in areas like human resources (HR). For organizations in document-intensive industries like insurance, banking, manufacturing, healthcare, or the public sector the savings are even larger.

Intelligent document processing is not just about accelerating operational efficiency. It's also about improving accuracy, customer experience, compliance, and employee satisfaction. It's about making your organization successful and contributing to your company's growth. This is why we are so excited to introduce General Availability of a new product – UiPath Document Understanding.

How robots can learn to understand documents

How do you imagine a robot processes documents? With Robotic Process Automation (RPA), software robots help humans automate highly manual, mundane chores giving them more time for higher-value tasks and business goals. These invisible little helpers make our lives easier and more interesting as we can delegate our routine paperwork to them and concentrate on more creative and important work. These are not physical robots sitting at a desktop and pushing buttons instead of humans – these are digital assistants who have been trained to process documents by reading and using applications like you do.

Thanks to UiPath Document Understanding, robots can read, extract, interpret, and act upon data from the documents using artificial intelligence (AI). For documents with fixed structure—like forms, passports, or licenses—it's enough to create rules or templates which will work for thousands of similar documents with no need for AI. At the same time, documents with varying layouts or with no fixed structure—like receipts, bills, or resumes—require advanced AI skills which can automatically determine the location of data even if the layout changes. Machine learning (ML) models are continually improving robots' skills to make them fast and accurate at document processing.

How Document Understanding works

UiPath Document Understanding pairs RPA and AI to automatically process your documents. It enables intelligent document processing within automation workflows, thus, allowing automation of complex and cognitive processes which are usually highly manual. In particular, Document Understanding can cope with tricky details like:

- Various structured documents like forms
- A wide range of less structured documents with varying layouts, for example, bills, invoices, receipts (including the ones with tables)
- Handwriting, signatures, and checkboxes
- Different file formats such as PDF, PNG, GIF, JPEG, TIFF, BMP, etc.
- Skewed, rotated, unrelated, or low-resolution scanned documents

Our team is excited to finally offer these capabilities to you, so let's dive a bit deeper into how all this works. UiPath Document Understanding includes several infrastructures integrated into the UiPath Platform which make it a unique solution currently available on the market.

Document Understanding framework

A composable Document Understanding framework is available as drag-and-drop activities used to build automation workflows in UiPath Studio.

1. Load taxonomy – define documents and data to be processed using the Taxonomy Manager.
2. Digitize – use one of the available optical character recognition (OCR) engines to digitize the text and determine its location.
3. Classify – classify the documents from the user-specified list.
4. Extract – choose the most suitable extractors according to your document type (see more details on extraction techniques in the following section).
5. Validate – if needed, a human can check in to confirm or correct the extracted data or handle exceptions.
6. Export – send the extracted info for further usage, for example, to email, spreadsheet, or SAP software.

In addition to the native capabilities, you can leverage a dynamic document understanding ecosystem complemented by the leading partner technologies for intelligent document processing. This can extend the UiPath Document Understanding framework to diverse document types and business needs.

Data extractors – from rules to ML models

Data extraction is where all the magic happens. In one of our previous blogs, we've already discussed how important it is to have an opportunity to combine rule-based and model-based approaches to data extraction. You have the option to use:

- Rules and templates
- ML models
- A hybrid approach (multi approach) combining both

With UiPath Document Understanding, you can easily configure which extractor to use for a specific document or even individual fields. All you need to do is open the Data Extraction Scope, set minimal confidence scores for each extractor, and put a tick wherever needed (see the screenshot below). As a result, this flexibility allows you to apply a hybrid approach. This means that the data can be extracted with a few techniques tried one-by-one unless the confidence score is met.

Now, let's look at the capabilities of each extractor:

- RegEx Based Extractor – relies on regular expressions to extract the information and works for data that appears in very similar textual contexts across documents.
- Form Extractor – functions based on the rules and templates and can process structured documents, tables, checkboxes.
- Intelligent Form Extractor – applies intelligent templates and rules for data extraction, and can process mostly structured documents including tables, handwriting, signatures, and checkboxes.
- ML Extractors for invoices, receipts, purchase orders, etc. – are powered by ML models. They are perfect for processing less structured documents with varying layouts, that are semi-structured

documents. The name of the extractor indicates the type of documents it is trained for. For now, we offer pre-trained ML models for invoices, receipts, purchase orders, and utility bills.

For unstructured documents like contracts or emails, we recommend trying a hybrid approach configured separately for each field or opting for our partner integrations.

It's crucial that you can also retrain the existing pre-trained ML models based on the customer data or use custom ML models for specific business needs. To enable model training, together with Document Understanding you get access to UiPath AI Fabric which can be used for model deployment and management.

Human validation

To help robots deal with low extraction accuracy and exceptions, employees can validate the extracted data using Validation Station in the UiPath Action Center. An employee receives a notification with a request to validate data or handle exceptions and can solve any uncertainties in a matter of clicks. They can use the Validation Station to review the extracted fields if needed to make sure the data is extracted accurately.

Document Understanding roadmap

In the upcoming releases, Document Understanding extractors will automatically learn from the custom data filled in by employees in the Validation Station. This is probably one of the most appealing features as it will allow ongoing model retraining at the touch of a button. The more you work with these models, the more trained and accurate they will become.

Apart from the currently available ML models for invoices, receipts, purchase orders, and utility bills, we're also working on some other common document types to ensure fast and smooth implementation of UiPath Document Understanding in your organization. Another high priority for us is to provide support for documents in some other languages like Japanese and Chinese. Finally, the solution will be available both on-premises and cloud, so you can choose the most favourable and compliant option for your company.