



Leveraging RPA and OCR to Automate a District Attorney's Office

Client

A decorative horizontal bar with a gradient from orange to light orange.

The Client is a District Attorney's office for one of the most populous counties in the United States.

Challenges

A decorative horizontal bar with a gradient from red to light pink.

The Client receives information on defendants via scanned documents sent by email in PDF format. Apart from the defendant's personal information, these PDFs also include a stamped date and an alphanumeric case number. These 10-90-page PDF documents are then downloaded, split into single-page documents, renamed, and moved into separate folders depending on their case type.

The main challenge centered around the sheer size of the documents, which often took 30-40 minutes to sift through and reorganize, slowing the entire process down. This time-consuming process meant that the business user often could not complete the activity within the stipulated timeline. Another major drawback to this process was the presence of alphanumeric data, which was challenging to read manually and led to errors. The read data was also used to locate the case within a legacy application to insert a record into the database; this proved highly time-consuming and labor-intensive.

The Client was looking to solve two significant challenges in this process: a solution based on Optical Character Recognition (OCR) for extracting data from the PDFs and a logistics application to minimize the processing time, reduce operational expenses and eliminate human errors.

Solution

A decorative horizontal bar with a gradient from blue to light blue.

Since the data that had to be extracted was located at different positions on the various documents, we used UiPath's OmniPage OCR tool and our custom-developed programming to include additional logic to automate the location and recognition of the data. Different templates were created for the various document layouts, with the text regions specified for each template.

The software then automatically found the corresponding template for each processed image. In this way, the applications could automatically perform OCR and extract the necessary information accurately without the need to specify the location of the data on the form manually. In addition, the bot performed all the required validation, including cross-checking against other systems and ultimately deciding which category the file goes into.

To solve logistical issues related to human resources, the data is read from the database and captured in a defined template and shared with the stakeholders automatically without extracting email ids from a different browser. The BOT was specially designed to make API calls to determine the email address of the recipient to whom the final attachments are sent to.

Value Delivered

Opteamix developers worked collaboratively with the Client to achieve the best results and implement the OCR functionality to achieve an error-free, time-saving, and cost-effective process. Some of the benefits the Client achieved by engaging with Opteamix are listed below:

- ▶ The entire process that earlier took 30 minutes to complete now takes 2 minutes from start to finish increasing operational efficiency by 93%.
- ▶ Automating the login, search, pdf read, and excel read functionalities negated any human intervention, promising accuracy of 96%.
- ▶ In addition to creating a seamless process, we also designed a source control database where files could be stored and accessed from one place.
- ▶ Achieved efficient invoice processing and reduced payment cycle duration.

About Opteamix

Opteamix is a digital automation technology consulting firm with deep expertise in Application Development, Robotic Process Automation, Artificial Intelligence, DevOps, Enterprise Mobility, and Test Automation Services. We are headquartered in Denver, Colorado with a wholly-owned delivery center in Bangalore, India.