Ominitive Extract - Invoices

About Taiger

Taiger is a B2B Artificial Intelligence (AI) Software company that focuses on Enterprise automation. Our niche is in automating and solving complex business problems where dependency on manpower and tedious unstructured data causes delays, errors and high costs. Unlike competitors, we need very few training samples and we also contractually guarantee accuracy. Since being founded in 2009 in Europe based on our Founder's PhD thesis, Taiger has grown to 130+ employees across 4 offices in Singapore, UAE, Spain and Mexico. We work on >50, multi-year, enterprise projects across Banking, Insurance, Government and other industries. No customer that's licensed Taiger's technology has ever stopped it since.

About The Product - Extract

Extract is an Intelligent Reading tool that can read, understand and derive information from complex, unstructured data in any language or format. Using our proprietary Cleansing, Classification, Optical Character Recognition and Natural Language Processing capabilities, Ominitive Extract works well with both Hard-Copy (Scanned or Photo) and Soft-Copy documents. It processes heaps of document types to extract information that is most relevant to your business. In the past, we have processed simple, Structured Documents (Ex. Passports), Semi-Structured Documents (Ex. Invoices), and Unstructured Documents (Ex. Annual Reports or Legal Documents) with >80% accuracy.

What is this solution?

Processing of invoices has always been a rush for the finance team at the end of every month. Over the years, organizations have tried to adapt their processes to speed up the process with little success. It was a challenge for RPA solutions to address this need due to the countless types of invoices from suppliers. With Ominitive Extract, the model is trained to extract the data points for backend processing.

How does this Demo work?

1. Click "Add Documents"
2. Select the correct model:
    a. Click "Options" to expand the menu
    b. Click "Select Type" to choose the model
    c. Select "tradedocument" in the first drop-down menu
    d. Select "InvoiceDemo" from the second drop-down menu

   Note: This step is only required for this trial/demonstration purposes. In production, the system will automatically detect and classify the document accordingly.

3. Click the box to browse, or drag-and-drop into the box, to upload your sample invoices.
   Note: For trial/demonstration purposes, the system only accepts one invoice per file.
4. Click "Upload" to start processing the document.
5. Click on each document (5a) or "Search" (5b) to browse for the documents to review the extracted results.
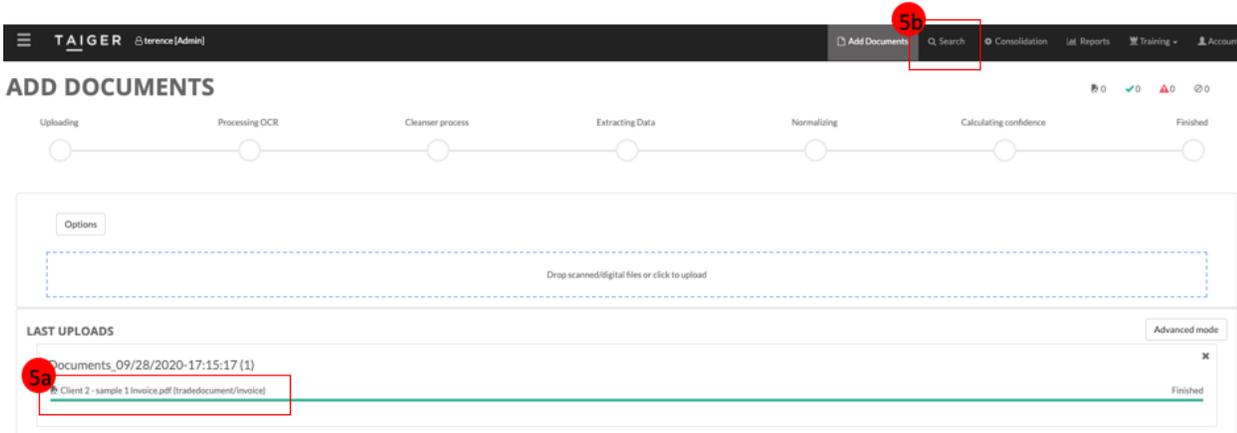
What should I expect?

The client whom we have implemented this project for has experienced* the following:

72%	**48.8%**	**>92%**
Reduction in total processing time	Successful completion of payments, 1.5x higher than the competitor benchmark	Extraction accuracy

* Note that every client's experience differs, and the above figures are for references and not a commitment in any form.

The system is trained to extract the data points from invoices, if the data is present on the document. The following are the data points that the system is trained to extract:

- Invoice Number
- Start Date
- Total Amount
- Currency

- Payment Terms
- Company Name
- To Date

Should you encounter any empty fields or "NA", the following are some of the possibilities:

1. Data point is not present on the document, masked/covered, noise (caused by handwritten text, stamps, etc).
2. The model is not trained based on your sample document
   Note: Unlike most capturing software out there, we do not use templates. The system is trained based on variances - the more varying documents the system is trained on, the more accurately it can extract any varying types of the document.

Reach out to the TAIGER team if you encounter any issues with the extraction or if you wish to understand how our technology works.

General Notes:

- Each file should meet the following requirements:
  - Documents are expected to be original digital copies in white background.
    - For the purpose of trial/demonstration, scanned copies are not acceptable. For specific use-cases, the models will be trained based on each client's documents.
  - Size of the document should be less than 10MB.
  - Documents will be in PDF formats only.
  - DPI threshold for each document should be at least 300 dpi.
  - Documents with noise, including but not limited to misaligned text, overlapping rows/characters/tables, will affect the OCR results.
  - Documents with stamps and handwritten text cannot be processed.