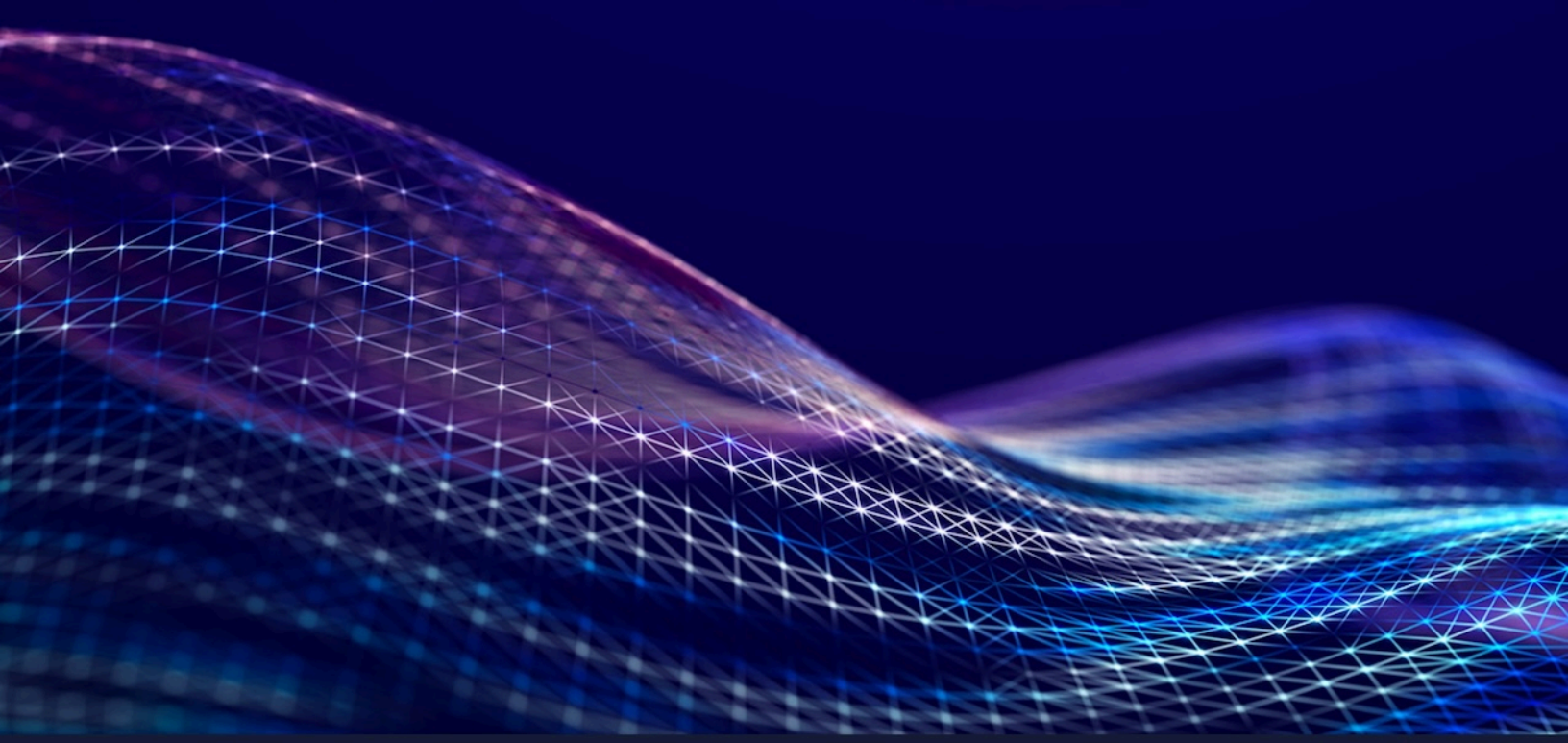

# Tungsten Extract Service
## Getting Started Guide

Version: 1.0.0

Date: 2024-06-13

TUNGSTEN AUTOMATION

# Table of Contents

# Preface

The Tungsten Extract Service is a license-based API service that can extract text from various types of images and PDF documents using Optical Character Recognition (OCR). This guide describes how to get started with Tungsten Extract Service.

## Training

Tungsten Automation offers both on-demand and instructor-led training to help you make the most of your product. To learn more about training courses and schedules, visit the Tungsten Automation Learning Cloud.

## Getting help with Tungsten Automation products

The Tungsten Automation Knowledge Portal repository contains articles that are updated on a regular basis to keep you informed about Tungsten Automation products. We encourage you to use the Knowledge Portal to obtain answers to your product questions.

To access the Tungsten Automation Knowledge Portal, go to https://knowledge.tungstenautomation.com/.

The Tungsten Automation Knowledge Portal is optimized for use with Google Chrome, Mozilla Firefox, or Microsoft Edge.

The Tungsten Automation Knowledge Portal provides:

- Powerful search capabilities to help you quickly locate the information you need.
  Type your search terms or phrase into the **Search** box, and then click the search icon.
- Product information, configuration details and documentation, including release news.
  To locate articles, go to the Knowledge Portal home page and select the applicable Solution Family for your product, or click the View All Products button.

From the Knowledge Portal home page, you can:

- Access the Tungsten Automation Community (for all customers).
  On the Resources menu, click the **Community** link.
- Access the Tungsten Automation Customer Portal (for eligible customers).
  Go to the Support Portal Information page and click **Log in to the Customer Portal**.
- Access the Tungsten Automation Partner Portal (for eligible partners).
  Go to the Support Portal Information page and click **Log in to the Partner Portal**.

- Access Tungsten Automation support commitments, lifecycle policies, electronic fulfillment details, and self-service tools.

  Go to the Support Details page and select the appropriate article.

# About Tungsten Extract Service

The Tungsten Extract Service recognizes text on images and PDF documents using Optical Character Recognition (OCR). OCR is a machine-learning technique used to recognize machine-printed text on images, such as forms, contracts, invoices, or other business documents. The text is extracted as words with coordinates relative to the page, essentially creating a digital version of the text that is searchable and selectable. With Tungsten Extract Service, you can eliminate or significantly reduce the need for manual data entry.

# Using OCR

Tungsten Extract Service provides a synchronous and asynchronous RESTful web service API to send images of business documents and retrieve the resulting text as JSON. This service API can be embedded in client applications or called from a no-code application that supports RESTful services.

To call the APIs, use the API Key. The API Key is provided via email after you purchase the service.

## Recommendations

- For the best possible text recognition, submit good-quality images to the API. The ideal documents are scanned business documents at 300 DPI in TIFF format.
- Results may vary if a document is less than 300 DPI; or if it represents only one region within a larger image, such as a photograph that includes a desk or other background behind the document.

## Supported languages

Tungsten Extract Service OCR supports a wide range of languages for the printed text including support for Western languages, Chinese, Arabic, and many others. See Set the language parameter for more information.

## Tungsten Extract Service features

The key features of Tungsten Extract Service include:

- Converting any supported image format to a searchable PDF with selectable text.
- Selecting speed or accuracy depending on the document image quality.
- Perfecting image before text recognition.

### Image conversion to a searchable PDF

The OCR APIs return the recognized text as JSON. Optionally, you can include a searchable PDF in the JSON reply that allows you to convert any supported image format to a searchable PDF with selectable text. We recommend that you request the PDF only if needed, to avoid unnecessarily large payloads that may cause delays.

## Speed versus accuracy

The OCR APIs support bias towards speed or accuracy. If your document images are of good quality, the fast mode may just work perfectly for you. If the images are more challenging, try the accurate mode.

## Image perfection

Tungsten Automation has industry-leading image perfection technology. This technology is available as an option in the OCR APIs. If you activate the image perfection option, the submitted images are perfected before the text is recognized. Optionally, you can return the perfected image in the JSON payload. The image perfection option can also detect the orientation of the image. For example, if an image is scanned upside down, the image perfection feature would detect it and rotate the image before recognizing the text.

When you select the image perfection option, it is recommended to also select the option to return the perfected image, because the coordinates (bounding boxes) of the recognized words are provided in relation to the perfected image, not the original image.

# Getting started with OCR

This section describes how to use the Tungsten Extract Service OCR APIs and provides code examples for reference.

## Prerequisites

The prerequisites for calling the OCR APIs are:

- An API Key, which you receive after purchasing the OCR page volume from Tungsten Automation
- TLS 1.2 or later

## Example: Call the synchronous OCR API

You can call the RESTful APIs from any client application using your preferred programming language, or from code-free applications that support RESTful services.

```
https://eu.cex.tungstenautomationcloud.com/services/sdk/ocr/v1/performSync[?language]
 [?createPdf]
[?ocrMode] [?imagePerfection] [?returnPerfectedImage] [?orientationDetection]
[?returnOCRData]
```

## Image constraints

The supported file formats are JPEG, PNG, PDF, and TIFF.

Limitations for synchronous and asynchronous APIs are listed in the following table.

| Limitation | Synchronous APIs | Asynchronous APIs |
|---|---|---|
| Number of pages | • Maximum number of pages for multipage type: 5 pages | • Maximum number of pages supported for OCR including PDF generation:<br>  • Bitonal A4/Letter/Legal 200 x 200 DPI or lower: 2000 pages<br>  • Bitonal A4/Letter/Legal 300 x 300 DPI: 1000 pages<br>  • Bitonal A4/Letter/Legal 600 x 600 DPI: 250 pages<br>  • Color/Grayscale A4/Letter/Legal 200 x 200 DPI or lower: 200 pages<br>  • Color/Grayscale A4/Letter/Legal 300 x 300 DPI: 100 pages<br>  • Color/Grayscale A4/Letter/Legal 600 x 600 DPI: 25 pages<br>• Maximum number of pages for A3 document: 10 pages |
| Size of submitted file | • Maximum: 10 MB | • Maximum: 200 MB |
| Image size | • Maximum: 4300 x 4300 pixels<br>If the input color format is 1-bit per pixel, or if Image Perfection is enabled, the Maximum is 8500 x 8500 pixels.<br>• Minimum: 200 x 200 pixels | • Maximum: 10000 x 10000 pixels<br>• Minimum: 200 x 200 pixels |
| Resolution | • 96 to 600 DPI | • 96 to 600 DPI |

# Set the language parameter

The required input language parameter to the OCR APIs is `[?language]`.

You must set this parameter to a supported language, such as German. You can retrieve the available languages and their exact names for the `[?language]` parameter by using the following API:

`/services/sdk/ocr/v1/languages`

If Tungsten Automation adds new languages, the preceding API gives you the ability to display the most current list of selectable languages. Currently, the following languages are supported.

Make sure you use the names exactly as displayed in this list with the `[?language]` parameter.

| | | | | | |
|---|---|---|---|---|---|
| Afrikaans | Dutch | Hungarian | Malay | Romanian | Swahili |
| Albanian | English | Icelandic | Malinke | Romany | Swazi |
| Arabic | Eskimo | Ido | Maltese | Rundi | Swedish |
| Aymara | Esperanto | Indonesian | Maori | Russian | Tagalog |
| Basque | Estonian | Interlingua | Mayan | Rwanda | Tahitian |
| Bemba | Faeroese | Italian | Miao | Sami | Thai |
| Blackfoot | Fijian | Japanese | Minangkabau | Sami(Lule) | Tongan |
| Brazalian | Finnish | Kabardian | Mohawk | Sami(Northern) | Tswana(Chuana) |
| Breton | French | Kasub | Moldavian | Sami(Southern) | Tun |
| Bugotu | Frisian | Kawa | Nahuatl | Samoan | Turkish |
| Bulgarian | Friulian | Kikuyu | Norwegian | Sardinian | Ukrainian |
| Byelorussian | Scottish | Kongo | Nyanja | Serbian | Vietnamese |
| Catalan | Gaelic(Irish) | Korean | Occidental | Serbian(Latin) | Visayan |
| Chamorro | Gaelic(Scottish) | Kpelle | Ojibway | Shona | Welsh |
| Chechen | Galician | Kurdish | Papiamento | Sioux | Wolof |
| Chinese(S) | Ganda(Luganda) | Latin | PidginEnglish | Slovak | Xhosa |
| Chinese(T) | German | Latvian | Pirez | Slovenian | Zapotec |
| Corsican | Greek | Lithuanian | Polish | Somali | Zulu |
| Croatian | Guarani | Luba | Portuguese | Sorbian(Wend) | |
| Crow | Hani | Luxembourgish | Provencal | Sotho | |
| Czech | Hawaiian | Macedonian | Quechua | Spanish | |
| Danish | Hebrew | Malagasy | Rhaetic | Sundanese | |

## Include a searchable PDF in the reply

You can select the **[?createPdf]** option to include a searchable PDF in the reply. In this case, a new PDF is always created even if the submitted document is a PDF.

## Example JSON

The OCR text is returned from the API calls as JSON. Each page object in the JSON has a width and height in pixels, DPI information, and a list of words. The words are roughly recognized and ordered from the top left to the bottom right of the document. Each word shows pixel coordinates relative to the top left of the page, which is [0,0]. For example, a word with Left = 100 starts at 100 pixels from the left edge of the page. Words also have a confidence property indicating how confident the OCR was about the Text result.

For a small two-page document with two words per page, the result looks similar to the following code.


```
{
    "Pages": [
        {
            "Size": {
                "Width": 3601,
                "Height": 2291
            },
            "HorizontalDPI": 300,
            "VerticalDPI": 300,
            "WordCollection": [
                {
                    "PageIndex": 0,
                    "Left": 163,
                    "Top": 167,
                    "Width": 75,
                    "Height": 64,
                    "Confidence": 0.1,
                    "Text": "Hello"
                },
                {
                    "PageIndex": 0,
                    "Left": 120,
                    "Top": 218,
                    "Width": 169,
                    "Height": 92,
                    "Confidence": 0.9,
                    "Text": "World"
                }
            ]
        },
            {
            "Size": {
                "Width": 3602,
                "Height": 2290
            },
            "HorizontalDPI": 300,
            "VerticalDPI": 300,
            "WordCollection": [
                {
                    "PageIndex": 0,
                    "Left": 163,
                    "Top": 167,
                    "Width": 75,
                    "Height": 64,
                    "Confidence": 0.99,
                    "Text": "Hi"
                },
                {
                    "PageIndex": 0,
                    "Left": 120,
                    "Top": 218,
                    "Width": 169,
                    "Height": 92,
                    "Confidence": 1,
                    "Text": "Again"
                }
            ]
        }
    ]
}
```

## Get results

The `/services/sdk/ocr/v1/performAsync` API processes the document immediately. Your client application needs to wait for the reply, which contains the JSON as payload along with multi-part content if you opted to have a PDF generated or the perfected image to be returned.

The `/services/sdk/ocr/v1/performAsync` API processes the document asynchronously. The call returns immediately with an `operationID` required to retrieve the result. This API is designed for larger documents that may take few seconds or few minutes to process. The client application needs to frequently check if the OCR operation is completed by calling the `/services/sdk/ocr/v1/getResult` API with the `operationID` as a parameter.

## Example code

The following is an example in C# code that describes how to call the `/services/sdk/ocr/v1/performSync` API and process the results by writing the OCR text of the input document to the console.

```
using System.Net.Http.Headers;

namespace QuickCloudOCR
{
    internal class Program
    {
        // Add your api key and URI to the service endpoint
        static readonly string ApiKey = @"PASTE_YOUR_OCR_API_KEY_HERE";
        static readonly string URI = @"https://SERVICEHOST.tungstenautomationcloud.com/
Services/Sdk/Ocr/v1/PerformSync";

        private static readonly string File = @"test_image.png";
        private static readonly string Mimetype = @"image/png";

        static void Main(string[] args)
        {
            // Image file will be the body of the request
            var httpContent = new ByteArrayContent(System.IO.File.ReadAllBytes(File));

            // Specify the mimetype of the image file
            httpContent.Headers.ContentType = MediaTypeHeaderValue.Parse(Mimetype);

            // Build the request uri with the endpoint plus parameters
            var uriBuilder = new UriBuilder(URI)
            {
                Query = @"language=English" // Add optional parameters here separated
 by &
            };

            // Create the client; the default timeout might not be long enough.
            using var client = new HttpClient() { Timeout = TimeSpan.FromMinutes(10) };

            // Put the api key in the headers
            client.DefaultRequestHeaders.Add(@"x-api-key", ApiKey);

            // Send the request
```

```
            var responseTask = client.PostAsync(uriBuilder.Uri, httpContent);

            // Wait for the response
            responseTask.Wait();

            var result = responseTask.Result;

            if (result.IsSuccessStatusCode == false)
            {
                var stringTask = result.Content.ReadAsStringAsync();
                stringTask.Wait();
                Console.Error.WriteLine($"Error: {result.ReasonPhrase}:
 {stringTask.Result}");
                Environment.Exit(1);
            }

            if (result.Content.Headers.ContentType?.MediaType == @"application/json")
            {
                var task = result.Content.ReadAsStringAsync();
                task.Wait();

                // Here are the OCR results in JSON
                Console.WriteLine(task.Result);
            }
        }
    }
}
```

The following is an example request to call the service from other languages.

```
curl --location --
request POST https://eu.cex.tungstenautomationcloud.com/Services/Sdk/Ocr/v1/
PerformSync ?language=English&createPdf=true&ImagePerfection=true&returnPerfectedImage
=true&returnOCRData=true&orientationDetection=true&ocrMode=1' \
--header 'x-api-key: PASTE_YOUR_OCR_API_KEY_HERE' \
--header 'Accept-Language: en-IN' \
--header 'Content-Type: image/png' \
--data-binary '@/C:/fakepath/test_image.png'
```