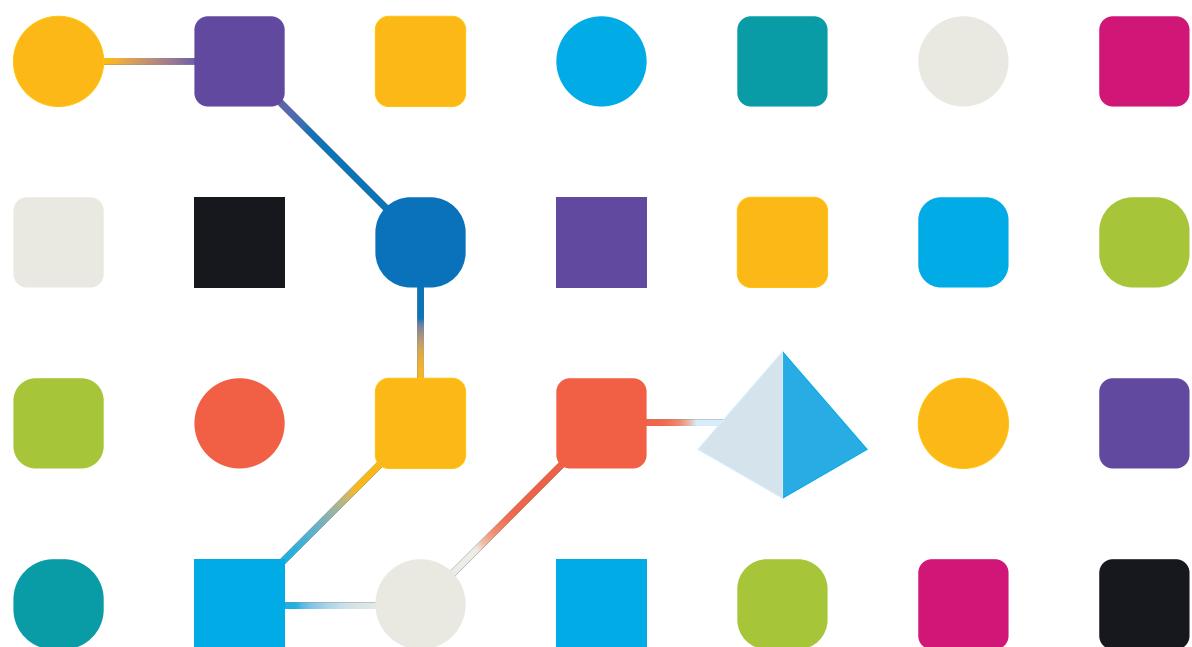


blueprism®

Blue Prism 6.2

Google Intelligent Services User Guide

Document Revision: 1.0



Trademarks and Copyright

The information contained in this document is the proprietary and confidential information of Blue Prism Limited and should not be disclosed to a third-party without the written consent of an authorized Blue Prism representative. No part of this document may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying without the written permission of Blue Prism Limited.

© Blue Prism Limited, 2001 – 2021

© “Blue Prism”, the “Blue Prism” logo and Prism device are either trademarks or registered trademarks of Blue Prism Limited and its affiliates. All Rights Reserved.

All trademarks are hereby acknowledged and are used to the benefit of their respective owners.
Blue Prism is not responsible for the content of external websites referenced by this document.

Blue Prism Limited, 2 Cinnamon Park, Crab Lane, Warrington, WA2 0XP, United Kingdom.
Registered in England: Reg. No. 4260035. Tel: +44 370 879 3000. Web: www.blueprism.com

Contents

Introduction	4
Google Intelligent Services Solution Overview	5
Limitations	5
Pre-Requisites and Environment Configuration	6
Cloud Services Prerequisites	6
Blue Prism Configuration	6
Using the VBOs	10
Natural Language Processing	10
Translation	14
Vision	16

Introduction

Robotic Process Automation (RPA) is an ever-expanding market which allows for back end office automation via intelligent robots. With the addition of a Blue Prism® Artificial Intelligence offering, the list of use cases for RPA will increase dramatically.

This document focuses on the design of the integration between Blue Prism and Google's Intelligent Services, implemented through REST APIs. These APIs include:

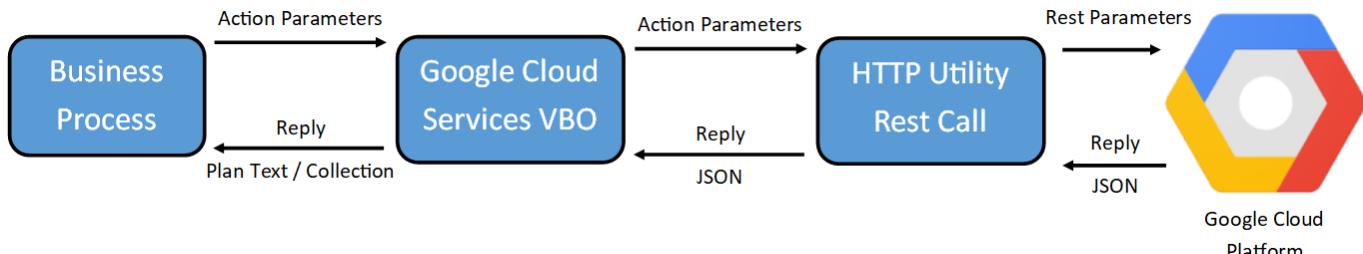
- Text Analytics (Natural Language Understanding)
- Visual Recognition
- Language Translation

Google Intelligent Services Solution Overview

The basic design is to provide one or more VBOs that encapsulate the different Google Intelligent services. These VBOs can then be offered as an easy bridge to connect a client's processes to the different services provided by Google.



The provided VBOs are only supported with Blue Prism 6.2 and above.



The Google Cloud Services VBO currently offers three different business objects. Natural Language Processing (NLP), Translation and Computer Vision (CV). These objects have been created to allow for an intuitive, and easy interface into Googles Cloud Services. This section shall show the architecture of these VBOs, while giving an in-depth explanation of how they function.

Limitations

The following limitations should be understood before attempting to use these integrations:

- The customer or partner is responsible for the configuration and maintenance of the relevant cloud subscriptions and services. Blue Prism can not provide any support on the configuration of the cloud environment itself.
- Use of the APIs may incur additional costs, depending on usage
- There is always a possibility with external services that the APIs will change. The VBOs are provided as-is without warranties and support is provided by Blue Prism on a best endeavours basis and are not subject to formal SLAs

Pre-Requisites and Environment Configuration

This section outlines the pre-requisites that are required in order to use the integrations. Note that Blue Prism is not able to provide any support in configuring the Cloud Services themselves.

Cloud Services Prerequisites

The following is required, before configuring the Blue Prism environment and attempting to use the VBOs:

- A Google Cloud Subscription
- An API Key
- Google Cloud Vision API
- Google Cloud Natural Language Processing API
- Google Cloud Translation API

Blue Prism Configuration

First, import the bprelease file. Before using the VBOs, it is necessary that the following environment variables be defined, as well as their appropriate values set. These values link up to Credentials in BluePrism which hold the subscription keys.

Environmental Variables

VBO	Variable Name	Description
Computer Vision	Google Services	This one Variable contains the API Key for all three of the Google Cloud Services VBO; Computer Vision, Natural Language Processing and Computer Vision.
Natural Language Processing		
Computer Vision		

 All the above are of the type, Text.

Credentials

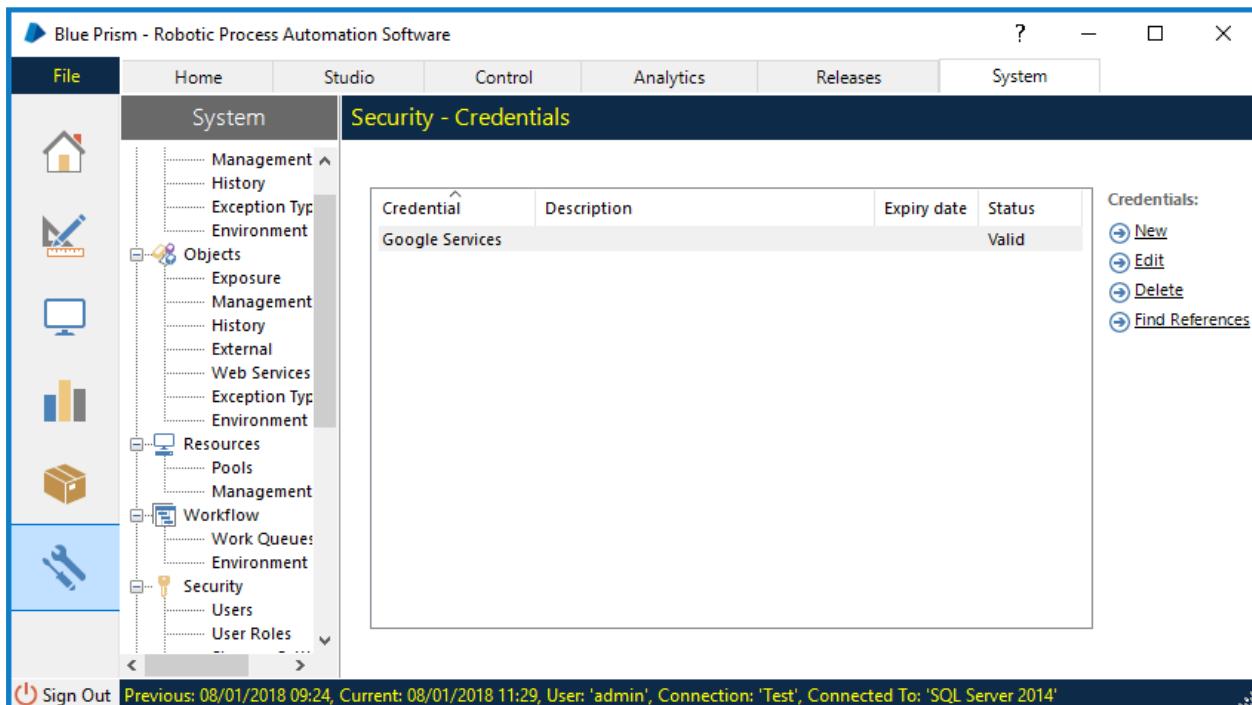
The credential manager holds the API key in relation to each Google Cloud Platform API. Each credential name will be the same as the environmental name. This section shall explain how to add a new credential for all three of the above VBOs.

Creating a Credential

Before being able to send a request to the NLP, Vision or Translation API, you will need to generate an API key in your Google Cloud Platform Developer Console.

1. Once an API key has been created, this needs to be saved in a Blue Prism Credential.
2. Open up BluePrism and navigate to the System tab, then, go to the Credentials tab under the Security section.

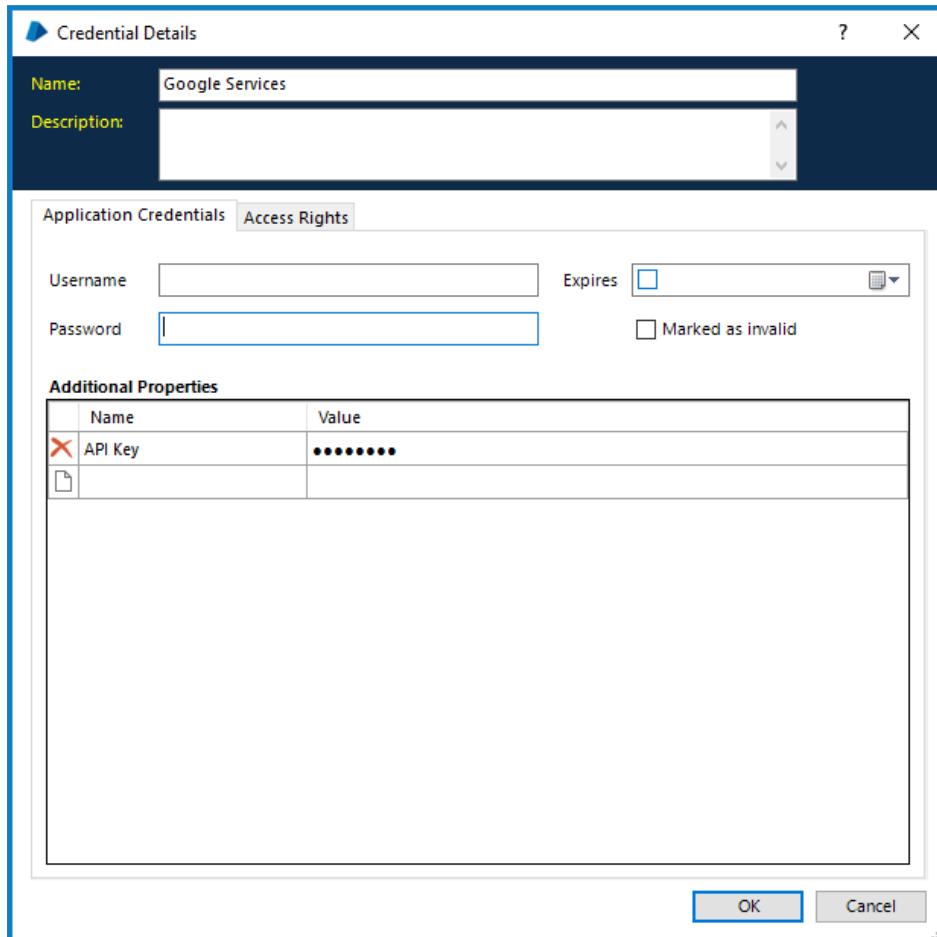
There should already be a credential present, if not continue to step 3 (If there is, head straight to step 6).



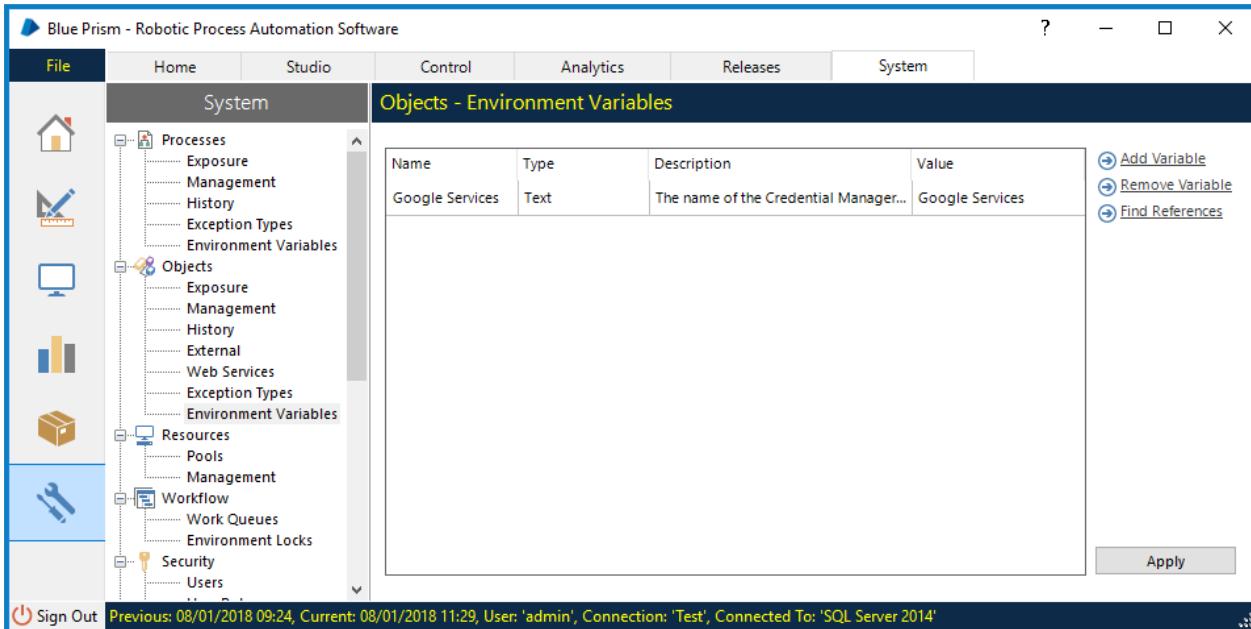
3. If a credential is not already present, click the "New" button to the right. Now, go ahead and call the Credential "Google Services". You may leave the Username and Password field empty (We aren't storing any credentials at this time)

4. Create an additional property named “API Key” and copy and paste the API Key in the value text box (This was generated in step 3). Click OK.

Your Credential should look the same as the image below.



5. The final step is to now link an environmental variable to the credential – this is so the Google Cloud Service VBO can access and use the API key in a secure manner. Go to the Environmental Variable tab under the Objects section. Create a new variable named Google Services with the value of Google Services if one is not already present.



6. The Google Cloud Service VBO has now been correctly configured.

Using the VBOs

The following section outlines the individual configuration and usage of each Google API and the supported actions within the VBO

Natural Language Processing

The natural language processing VBO consists of five different actions:

- Sentiment Analysis
- Entity Analysis
- Sentiment Entity Analysis
- Syntax Analysis
- Content Classification

Sentiment analysis

Sentiment Analysis allows your organisation to analyse the sentiment of the provided text. Once the call has been sent, the returned data is formatted into specific, easy to use outputs:

- **Document Sentiment** - This collection is organised into two fields; Magnitude and Score, this represents the overall sentiment of the provided text.
 - **Magnitude** is a non-negative number in the range of [0, +inf] which represents the amount of “passion” in the provided text.
 - **Score** is a number between -1.0 and 1.0. Anything below 0 has a negative sentiment, and anything above 0 is positive sentiment. Though, a sentiment of 1.0 is more positive than a sentiment value of 0.2, for example - this works the same for a negative sentiment value.
- **Language** - This is stored as a text output and corresponds to the [ISO language format](#).
- **Sentence** - This last output is a collection that contains the Document Sentiment information fields, as well as a sentence field which states what sentence the sentiment analysis belongs to.

Outputs

Document Sentiment (Collection)	Language (Text)	Sentence (Collection)
Field 1: Magnitude	Language Code	Field 1: Sentence
Field 2: Score		Field 2: Magnitude
		Field 3: Score

Entity analysis

Entity Analysis allows your organisation to analyse and find entities in the given text. An entity can be classed as a proper name, or common nouns. This function will also find entity types, such as salience, mentions and other properties. This action has two outputs:

- **Entities** - This collection is organised into four fields:
 - **Name** - The representative name for the entity
 - **Entity Type** - The entity type (Unknown, Person, Location, Organisation, Event, Art, Consumer or other)
 - **Metadata** - Contains a URL link if present
 - **Salience** - The salience score for an entity provides information about the importance or centrality of that entity to the entire document text. Scores closer to 0 are less salient, while scores closer to 1.0 are highly salient.
- **Language** - This is stored as a text output and corresponds to the [ISO language format](#).

Outputs

Entities (Collection)	Language (Text)
Field 1: Name	Language Code
Field 2: Entity Type	
Field 3: Metadata	
Field 4: Salience	

Sentiment Entity Analysis

Sentiment Entity Analysis allows your organisation to find entities and analyse their sentiment values in the given text. An entity can be classed as a proper name, or common nouns. This action has two outputs:

- **Entities** - This collection is organised into seven fields:
 - **Name** - The representative name for the entity
 - **Entity Type** - The entity type (Unknown, Proper Name or Common / Compound Noun).
 - **Sentiment Magnitude** - A non-negative number in the range of [0, +inf] which represents the amount of “passion” in the provided text.
 - **Sentiment Score** - A number between -1.0 and 1.0. Anything below 0 has a negative sentiment, and anything above 0 is positive sentiment. Though, a sentiment of 1.0 is more positive than a sentiment value of 0.2, for example - this works the same for a negative sentiment value.
- **Language** - This is stored as a text output and corresponds to the [ISO language format](#).

Outputs

Entities (Collection)	Language (Text)
Field 1: Name	Language Code
Field 2: Entity Type	
Field 3: Metadata	
Field 4: Salience	
Field 5: Mentions	
Field 6: Sentiment Score	
Field 7: Sentiment Magnitude	

Syntax Analysis

Syntax Analysis analyses the syntax of the text and provides sentence tokenisation along with part of speech tags, dependency trees, and other properties. This action has two outputs:

- **Tokens** - This collection is organised into fourteen fields:
 - **Tag** - Speech tag (Adjective, Adverb etc.)
 - **Aspect** - The characteristics of the verb.
 - **Case** - The grammatical function performed by a noun or pronoun in a phrase, clause, or sentence.
 - **Form** - Depending on the language, Form can be categorizing different forms of verbs, adjectives, adverbs, etc.
 - **Gender** - Gender classes of nouns reflected in the behavior of associated words.
 - **Mood** - The grammatical feature of verbs, used for showing modality and attitude.
 - **Number** - Count distinctions.
 - **Person** - The distinction between the speaker, second person, third person, etc.
 - **Proper** - This category shows if the token is part of a proper name.
 - **Reciprocity** - Reciprocal features of a pronoun.
 - **Tense** - Time reference (past, future etc.).
 - **Voice** - The relationship between the action that a verb expresses and the participants identified by its arguments.
 - **Label** - [The parse label for the token](#) (Abbreviation modifier, Adjectival complement etc.).
 - **Lemma** - The word being analysed
- **Language** - This is stored as a text output and corresponds to the [ISO language format](#).

Outputs

Syntax (Collection)	Language (Text)
Field 1: Tag	Language Code
Field 2: Aspect	
Field 3: Case	
Field 4: Form	
Field 5: Gender	
Field 6: Mood	
Field 7: Number	
Field 8: Person	
Field 9: Proper	
Field 10: Reciprocity	
Field 11: Tense	
Field 12: Voice	
Field 13: Label	
Field 14: Lemma	

Content Classification

Content Classification classifies the given text into specific categories with a confidence score. This action has one output:

- **Classification** - Categories representing the input document. This collection has two fields:
 - **Name** - The name of the category representing the document.
 - **Confidence** - The classifier's confidence of the category. Number represents how certain the classifier is that this category represents the given text (0.0 – 1.0).

See table below for the outputs.

Classification (Collection)
Field 1: Name
Field 2: Confidence

Translation

The Translation VBO consists of four separate actions:

- Detect Language
- Detect Multiple
- Translate Text
- Translate Multiple

Detect Language

Detect Language allows your organisation to detect the language of a piece of text. Once the action has been computed, the returned data is organised into outputs:

- **Detected Source Language** - This output represents the actual detected language of the given text, and corresponds to the [ISO language format](#).
- **Confidence Score** - This score is between 0, and 1, where 0 corresponds to 0%, and 1 corresponds to 100%.

Outputs

Detected Source Language (Text)	Confidence Score (Text)
Language Code	Confidence Score

Detect Multiple

Detect Multiple allows your organisation to detect an unlimited amount of text formatted in a wide range of languages. Once the action has been computed, the returned data is organised into the following outputs:

- **Detected Source Language and Confidence Collection** - This Collection represents the actual detected language of the given text with a confidence score for input. It contains two fields:
 - **Detected Source Language** - This output represents the actual detected language of the given text, and corresponds to the [ISO language format](#).
 - **Confidence Score** - This score is between 0, and 1, where 0 corresponds to 0%, and 1 corresponds to 100%.

Outputs

Detected Language and Confidence Score (Collection)
Field 1: Detected Source Language
Field 2: Confidence Score

Translate Text

Translate Text allows your organisation to translate the provided text into any of the supported languages. Once the action has been computed, the returned data is organised into the following outputs:

- **Translated Text** - The provided text translated to whichever language was stated.
- **Detected Source Language** - This output represents the actual detected language of the given text, and corresponds to the [ISO language format](#).

Outputs

Translated Text (Text)	Detected Source Language (Text)
Translated Text	Detected Source Language

Translate Multiple

Translate Multiple allows your organisation to translate a Collection of provided text into any of the supported languages. Once the action has been computed, the returned data is organised into the following outputs:

- **Translated Text Collection** - This Collection contains two fields with represent the translated text, and also the detected source language.
 - **Translated Text** - The provided text translated to whichever language was stated.
 - **Detected Source Language** - This output represents the actual detected language of the given text, and corresponds to the [ISO language format](#).

Outputs

Detected Language and Confidence Score (Collection)
Field 1: Translated Text
Field 2: Detected Source Language

Vision

The Computer Vision VBO consists of ten different actions:

- Crop Hints
- Face Detection
- Read Document Text
- Read Text
- Get Image Properties
- Label Detection
- Landmark Detection
- Logo Detection
- Safe Search
- Web Search

Crop Hints

Crop hints requests return the coordinates of a bounding box that surrounds the dominant object or face in an image. These coordinates can be used to crop the image to feature that dominant object. Once the action has been computed, the returned data is organised into the following outputs:

- **Translated Text Collection** - This Collection contains two fields with represent the translated text, and also the detected source language.
 - **Translated Text** - The provided text translated to whichever language was stated.
 - **Detected Source Language** - This output represents the actual detected language of the given text, and corresponds to the [ISO language format](#).

Outputs

Detected Language and Confidence Score (Collection)
Field 1: Translated Text
Field 2: Detected Source Language

Face Detection

Face Detection detects multiple faces within an image along with the associated key facial attributes such as emotional state or wearing head wear. The returned data is organised into the following outputs:

- **Facial Attributes** - This Collection contains three fields representing the facial feature detected with its location formatted as a coordinate.
 - **Type** - The type of facial feature detected (Eye, cheek etc.)
 - **Position (Collection)** - The location of the facial feature as an X, Y and Z Coordinate.
- **Image Properties** - This collection contains 10 fields which represents the emotion of the persons face, any potential rotations and other properties
 - Roll Angle
 - Pan Angle
 - Tilt Angle
 - Joy Likelihood
 - Sorrow Likelihood
 - Anger Likelihood
 - Surprise Likelihood
 - Under Exposed Likelihood
 - Blurred Likelihood
 - Headwear Likelihood
- **Detection Confidence** - This output is single text output with a confidence score between 0, and 1.
- **Landmarking Confidence** - This output is single text output with a confidence score between 0, and 1.

Outputs

Facial Attributes (Collection)	Image Properties (Collection)	Detection Confidence (Text)	Landmarking Confidence (Text)
Field 1: Type	Field 1: Roll Angle	Detection Confidence	Landmarking Confidence
Field 2: Position (Collection)	Field 2: Pan Angle		
	Field 3: Tilt Angle		
	Field 4: Joy Likelihood		
	Field 5: Sorrow Likelihood		
	Field 6: Anger Likelihood		
	Field 7: Surprise Likelihood		
	Field 8: Under Exposed Likelihood		
	Field 9: Blurred Likelihood		
	Field 10: Headwear Likelihood		

Read Document Text

Read Document Text is an OCR function which reads textual content from an image. Once the action has been computed, the returned data is organised into the following outputs:

- **Text** - This is the main overall output of the OCR method, it contains the read text.
- **Detected Source Language** - This output represents the actual detected language of the given text, and corresponds to the [ISO language format](#).

Outputs

Text	Detected Source Language
Read Text	Detected Source Language

Read Text

Read Text is an OCR function which reads textual content from an image. This action is better suited to imagery which isn't in a document format (Such as a receipt). Once the action has been computed, the returned data is organised into the following outputs:

- **Text** - This is the main overall output of the OCR method, it contains the
- **Detected Source Language** - This output represents the actual detected language of the given text, and corresponds to the [ISO language format](#).

Outputs

Text	Detected Source Language
Read Text	Detected Source Language

Get Image Properties

An image properties request returns the dominant colours in the image as RGB values and percent of the total pixel count. Once the action has been computed, the returned data is organised into the following outputs:

- **Image Properties Collection** - This Collection contains three fields with represent the overall image properties
 - **RGB** - This field holds the colour information formatted as Red, Green and Blue.
 - **Score** - Image-specific score for this colour. Value in range [0, 1].
 - **Pixel Fraction** - The fraction of pixels the colour occupies in the image. Value in range [0, 1].

Outputs

Image Properties (Collection)
Field 1: RGB
Field 2: Score
Field 3: Pixel Fraction

Label Detection

The label detection process can detect and extract information about entities within an image, across a broad group of categories. Labels can identify objects, locations, activities, animal species, products, and more. Labels are returned in English only, but can be translated into another language by using the Translation Process.. Once the action has been computed, the returned data is organised into the following outputs:

- **Label Annotation Collection** - This Collection contains two fields with represent a label with a confidence score.
 - **Description** - A description of the classified label.
 - **Confidence Score** - Is the confidence score, which ranges from 0 (no confidence) to 1 (very high confidence).

Outputs

Label Annotation (Collection)
Field 1: Description
Field 2: Confidence Score

Landmark Detection

Landmark requests detect well-known natural and human-made landmarks, and return identifying information such as the landmark's name and location. Once the action has been computed, the returned data is organised into the following outputs:

- **Landmark Collection** - This Collection contains three fields with represent the description, location and confidence score of the landmark found.
 - **Description** - A description of the landmark
 - **Confidence Score** - The overall confidence in the classification (0 to 1)
 - **Location** - The longitude and Latitude coordinates.

Outputs

Landmark (Collection)
Field 1: Description
Field 2: Confidence Score
Field 3: Location

Logo Detection

Logo Detection allows your organisation to detect popular product and corporate logos within an image. Once the action has been computed, the returned data is organised into the following outputs:

- **Logo Collection** - This Collection contains two fields which represent the description of the logo and overall confidence score in its classification decision.
 - **Description** - The description of the logo
 - **Confidence Score** - The overall confidence in the classification (0 to 1)

Outputs

Logo Classification (Collection)
Field 1: Description
Field 2: Confidence Score

Safe Search

Safe Search requests examines an image for potentially unsafe or undesirable content. Likelihood of such imagery is returned in 4 categories:

- Adult indicates content generally suited for 18 years plus, such as nudity, sexual activity, and pornography (including cartoons or anime).
- Spoof indicates content that has been modified from the original to make it funny or offensive.
- Medical indicates content such as surgeries or MRIs.
- Violent indicates violent content, including but not limited to the presence of blood, war images, weapons, injuries, or car crashes.

Likelihood levels are bucketed into 4 categories, ranging from VERY_UNLIKELY to VERY_LIKELY. Once the action has been computed, the returned data is organised into the following outputs:

- **Safe Search Collection** - This Collection contains four fields which indicate which category the image belongs too
 - Adult
 - Spoof
 - Medical
 - Violent

Outputs

Safe Search Classification (Collection)
Field 1: Adult
Field 2: Spoof
Field 3: Medical
Field 4: Violent

Web Search

This action allows your organisation to do a web search based upon the given image. This process returns information to do with the description of the given image, URL links to full matching and similar imagery. Once the action has been computed, the returned data is organised into the following outputs:

- **Web Entities Collection** - This Collection contains two fields with represent the description and score of the classified Entity.
 - **Description** - A description of each classified entity.
 - **Score** - The confidence score related to the overall classification accuracy.
- **Visually Similar Imagery** - This collection contains four fields related to similar matching imagery:
 - **Full Matching Images** - Fully matching images from the Internet. Can include resized copies of the query image.
 - **Partial Matching Images** - Partial matching images from the Internet. Those images are similar enough to share some key-point features. For example, an original image will likely have partial matching for its crops.
 - **Pages with Matching Images** - Web pages containing the matching images from the Internet.
 - **Visually Similar Images** - The visually similar image results.

Outputs

Web Entities (Collection)	Visually Similar Imagery (Collection)
Field 1: Description	Field 1: Full Matching Images
Field 2: Score	Field 2: Partial Matching Images
	Field 3: Pages with Matching Images
	Field 4: Visually Similar Images