aws

API Reference

Amazon Textract

API Version 2018-06-27

Amazon Textract: API Reference

Table of Contents

Welcome

Amazon Textract detects and analyzes text in documents and converts it into machine-readable text. This is the API reference documentation for Amazon Textract.

This document was last published on May 29, 2026.

Actions

The following actions are supported:

- AnalyzeDocument
- AnalyzeExpense
- AnalyzeID
- CreateAdapter
- CreateAdapterVersion
- DeleteAdapter
- DeleteAdapterVersion
- DetectDocumentText
- GetAdapter
- GetAdapterVersion
- GetDocumentAnalysis
- GetDocumentTextDetection
- GetExpenseAnalysis
- GetLendingAnalysis
- GetLendingAnalysisSummary
- ListAdapters
- ListAdapterVersions
- ListTagsForResource
- StartDocumentAnalysis
- StartDocumentTextDetection
- StartExpenseAnalysis
- StartLendingAnalysis
- TagResource
- UntagResource
- UpdateAdapter

AnalyzeDocument

Analyzes an input document for relationships between detected items.

The types of information returned are as follows:

- Form data (key-value pairs). The related information is returned in two Block objects, each of type KEY_VALUE_SET: a KEY Block object and a VALUE Block object. For example, *Name: Ana Silva Carolina* contains a key and value. *Name:* is the key. *Ana Silva Carolina* is the value.
- Table and table cell data. A TABLE Block object contains information about a detected table. A CELL Block object is returned for each cell in a table.
- Lines and words of text. A LINE Block object contains one or more WORD Block objects. All lines and words that are detected in the document are returned (including text that doesn't have a relationship with the value of FeatureTypes).
- Signatures. A SIGNATURE Block object contains the location information of a signature in a document. If used in conjunction with forms or tables, a signature can be given a Key-Value pairing or be detected in the cell of a table.
- Query. A QUERY Block object contains the query text, alias and link to the associated Query results block object.
- Query Result. A QUERY_RESULT Block object contains the answer to the query and an ID that connects it to the query asked. This Block also contains a confidence score.

Selection elements such as check boxes and option buttons (radio buttons) can be detected in form data and in tables. A SELECTION_ELEMENT Block object contains information about a selection element, including the selection status.

You can choose which type of analysis to perform by specifying the FeatureTypes list.

The output is returned in a list of Block objects.

AnalyzeDocument is a synchronous operation. To analyze documents asynchronously, use StartDocumentAnalysis.

For more information, see Document Text Analysis.

Request Syntax

```
{
```

```
   "AdaptersConfig": {
      "Adapters": [
         {
            "AdapterId": "string",
            "Pages": [ "string" ],
            "Version": "string"
         }
      ]
   },
   "Document": {
      "Bytes": blob,
      "S3Object": {
         "Bucket": "string",
         "Name": "string",
         "Version": "string"
      }
   },
   "FeatureTypes": [ "string" ],
   "HumanLoopConfig": {
      "DataAttributes": {
         "ContentClassifiers": [ "string" ]
      },
      "FlowDefinitionArn": "string",
      "HumanLoopName": "string"
   },
   "QueriesConfig": {
      "Queries": [
         {
            "Alias": "string",
            "Pages": [ "string" ],
            "Text": "string"
         }
      ]
   }
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters.

The request accepts the following data in JSON format.

AdaptersConfig

Specifies the adapter to be used when analyzing a document.

Type: AdaptersConfig object

Required: No

Document

The input document as base64-encoded bytes or an Amazon S3 object. If you use the AWS CLI to call Amazon Textract operations, you can't pass image bytes. The document must be an image in JPEG, PNG, PDF, or TIFF format.

If you're using an AWS SDK to call Amazon Textract, you might not need to base64-encode image bytes that are passed using the `Bytes` field.

Type: Document object

Required: Yes

FeatureTypes

A list of the types of analysis to perform. Add TABLES to the list to return information about the tables that are detected in the input document. Add FORMS to return detected form data. Add SIGNATURES to return the locations of detected signatures. Add LAYOUT to the list to return information about the layout of the document. All lines and words detected in the document are included in the response (including text that isn't related to the value of `FeatureTypes`).

Type: Array of strings

Valid Values: `TABLES | FORMS | QUERIES | SIGNATURES | LAYOUT`

Required: Yes

HumanLoopConfig

Sets the configuration for the human in the loop workflow for analyzing documents.

Type: HumanLoopConfig object

Required: No

QueriesConfig

Contains Queries and the alias for those Queries, as determined by the input.

Type: QueriesConfig object

Required: No

Response Syntax

```
{
   "AnalyzeDocumentModelVersion": "string",
   "Blocks": [
      {
         "BlockType": "string",
         "ColumnIndex": number,
         "ColumnSpan": number,
         "Confidence": number,
         "EntityTypes": [ "string" ],
         "Geometry": {
            "BoundingBox": {
               "Height": number,
               "Left": number,
               "Top": number,
               "Width": number
            },
            "Polygon": [
               {
                  "X": number,
                  "Y": number
               }
            ],
            "RotationAngle": number
         },
         "Id": "string",
         "Page": number,
         "Query": {
            "Alias": "string",
            "Pages": [ "string" ],
            "Text": "string"
         },
         "Relationships": [
            {
```

```
                "Ids": [ "string" ],
                "Type": "string"
            }
         ],
         "RowIndex": number,
         "RowSpan": number,
         "SelectionStatus": "string",
         "Text": "string",
         "TextType": "string"
      }
   ],
   "DocumentMetadata": {
      "Pages": number
   },
   "HumanLoopActivationOutput": {
      "HumanLoopActivationConditionsEvaluationResults": "string",
      "HumanLoopActivationReasons": [ "string" ],
      "HumanLoopArn": "string"
   }
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

AnalyzeDocumentModelVersion

The version of the model used to analyze the document.

Type: String

Blocks

The items that are detected and analyzed by `AnalyzeDocument`.

Type: Array of Block objects

DocumentMetadata

Metadata about the analyzed document. An example is the number of pages.

Type: DocumentMetadata object

HumanLoopActivationOutput

Shows the results of the human in the loop evaluation.

Type: HumanLoopActivationOutput object

Errors

For information about the errors that are common to all actions, see Common Error Types.

AccessDeniedException

You aren't authorized to perform the action. Use the Amazon Resource Name (ARN) of an authorized user or IAM role to perform the operation.

HTTP Status Code: 400

BadDocumentException

Amazon Textract isn't able to read the document. For more information on the document limits in Amazon Textract, see Hard limits.

HTTP Status Code: 400

DocumentTooLargeException

The document can't be processed because it's too large. The maximum document size for synchronous operations 10 MB. The maximum document size for asynchronous operations is 500 MB for PDF files.

HTTP Status Code: 400

HumanLoopQuotaExceededException

Indicates you have exceeded the maximum number of active human in the loop workflows available

QuotaCode

The quota code.

ResourceType

The resource type.

ServiceCode

The service code.

HTTP Status Code: 400

InternalServerError

Amazon Textract experienced a service issue. Try your call again.

HTTP Status Code: 500

InvalidParameterException

An input parameter violated a constraint. For example, in synchronous operations, an `InvalidParameterException` exception occurs when neither of the `S3Object` or `Bytes` values are supplied in the `Document` request parameter. Validate your parameter before calling the API operation again.

HTTP Status Code: 400

InvalidS3ObjectException

Amazon Textract is unable to access the S3 object that's specified in the request. for more information, Configure Access to Amazon S3 For troubleshooting information, see Troubleshooting Amazon S3

HTTP Status Code: 400

ProvisionedThroughputExceededException

The number of requests exceeded your throughput limit. If you want to increase this limit, contact Amazon Textract.

HTTP Status Code: 400

ThrottlingException

Amazon Textract is temporarily unable to process the request. Try your call again.

HTTP Status Code: 500

UnsupportedDocumentException

The format of the input document isn't supported. Documents for operations can be in PNG, JPEG, PDF, or TIFF format.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface V2
- AWS SDK for .NET V4
- AWS SDK for C++
- AWS SDK for Go v2
- AWS SDK for Java V2
- AWS SDK for JavaScript V3
- AWS SDK for Kotlin
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3

AnalyzeExpense

`AnalyzeExpense` synchronously analyzes an input document for financially related relationships between text.

Information is returned as `ExpenseDocuments` and seperated as follows:

- `LineItemGroups`- A data set containing `LineItems` which store information about the lines of text, such as an item purchased and its price on a receipt.
- `SummaryFields`- Contains all other information a receipt, such as header information or the vendors name.

Request Syntax

```
{
   "Document": {
      "Bytes": blob,
      "S3Object": {
         "Bucket": "string",
         "Name": "string",
         "Version": "string"
      }
   }
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters.

The request accepts the following data in JSON format.

Document

The input document, either as bytes or as an S3 object.

You pass image bytes to an Amazon Textract API operation by using the `Bytes` property. For example, you would use the `Bytes` property to pass a document loaded from a local file system. Image bytes passed by using the `Bytes` property must be base64 encoded. Your code might not need to encode document file bytes if you're using an AWS SDK to call Amazon Textract API operations.

You pass images stored in an S3 bucket to an Amazon Textract API operation by using the `S3Object` property. Documents stored in an S3 bucket don't need to be base64 encoded.

The AWS Region for the S3 bucket that contains the S3 object must match the AWS Region that you use for Amazon Textract operations.

If you use the AWS CLI to call Amazon Textract operations, passing image bytes using the Bytes property isn't supported. You must first upload the document to an Amazon S3 bucket, and then call the operation using the S3Object property.

For Amazon Textract to process an S3 object, the user must have permission to access the S3 object.

Type: Document object

Required: Yes

Response Syntax

```
{
   "DocumentMetadata": {
      "Pages": number
   },
   "ExpenseDocuments": [
      {
         "Blocks": [
            {
               "BlockType": "string",
               "ColumnIndex": number,
               "ColumnSpan": number,
               "Confidence": number,
               "EntityTypes": [ "string" ],
               "Geometry": {
                  "BoundingBox": {
                     "Height": number,
                     "Left": number,
                     "Top": number,
                     "Width": number
                  },
                  "Polygon": [
                     {
                        "X": number,
```

```
                        "Y": number
                     }
                  ],
                  "RotationAngle": number
               },
               "Id": "string",
               "Page": number,
               "Query": {
                  "Alias": "string",
                  "Pages": [ "string" ],
                  "Text": "string"
               },
               "Relationships": [
                  {
                     "Ids": [ "string" ],
                     "Type": "string"
                  }
               ],
               "RowIndex": number,
               "RowSpan": number,
               "SelectionStatus": "string",
               "Text": "string",
               "TextType": "string"
            }
         ],
         "ExpenseIndex": number,
         "LineItemGroups": [
            {
               "LineItemGroupIndex": number,
               "LineItems": [
                  {
                     "LineItemExpenseFields": [
                        {
                           "Currency": {
                              "Code": "string",
                              "Confidence": number
                           },
                           "GroupProperties": [
                              {
                                 "Id": "string",
                                 "Types": [ "string" ]
                              }
                           ],
                           "LabelDetection": {
```

```
                        "Confidence": number,
                        "Geometry": {
                           "BoundingBox": {
                              "Height": number,
                              "Left": number,
                              "Top": number,
                              "Width": number
                           },
                           "Polygon": [
                              {
                                 "X": number,
                                 "Y": number
                              }
                           ],
                           "RotationAngle": number
                        },
                        "Text": "string"
                     },
                     "PageNumber": number,
                     "Type": {
                        "Confidence": number,
                        "Text": "string"
                     },
                     "ValueDetection": {
                        "Confidence": number,
                        "Geometry": {
                           "BoundingBox": {
                              "Height": number,
                              "Left": number,
                              "Top": number,
                              "Width": number
                           },
                           "Polygon": [
                              {
                                 "X": number,
                                 "Y": number
                              }
                           ],
                           "RotationAngle": number
                        },
                        "Text": "string"
                     }
                  }
               ]
```

```
                }
            ]
        }
    ],
    "SummaryFields": [
        {
            "Currency": {
                "Code": "string",
                "Confidence": number
            },
            "GroupProperties": [
                {
                    "Id": "string",
                    "Types": [ "string" ]
                }
            ],
            "LabelDetection": {
                "Confidence": number,
                "Geometry": {
                    "BoundingBox": {
                        "Height": number,
                        "Left": number,
                        "Top": number,
                        "Width": number
                    },
                    "Polygon": [
                        {
                            "X": number,
                            "Y": number
                        }
                    ],
                    "RotationAngle": number
                },
                "Text": "string"
            },
            "PageNumber": number,
            "Type": {
                "Confidence": number,
                "Text": "string"
            },
            "ValueDetection": {
                "Confidence": number,
                "Geometry": {
                    "BoundingBox": {
```

```
                        "Height": number,
                        "Left": number,
                        "Top": number,
                        "Width": number
                     },
                     "Polygon": [
                        {
                           "X": number,
                           "Y": number
                        }
                     ],
                     "RotationAngle": number
                  },
                  "Text": "string"
               }
            }
         ]
      }
   ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

DocumentMetadata

Information about the input document.

Type: DocumentMetadata object

ExpenseDocuments

The expenses detected by Amazon Textract.

Type: Array of ExpenseDocument objects

Errors

For information about the errors that are common to all actions, see Common Error Types.

AccessDeniedException

You aren't authorized to perform the action. Use the Amazon Resource Name (ARN) of an authorized user or IAM role to perform the operation.

HTTP Status Code: 400

BadDocumentException

Amazon Textract isn't able to read the document. For more information on the document limits in Amazon Textract, see Hard limits.

HTTP Status Code: 400

DocumentTooLargeException

The document can't be processed because it's too large. The maximum document size for synchronous operations 10 MB. The maximum document size for asynchronous operations is 500 MB for PDF files.

HTTP Status Code: 400

InternalServerError

Amazon Textract experienced a service issue. Try your call again.

HTTP Status Code: 500

InvalidParameterException

An input parameter violated a constraint. For example, in synchronous operations, an `InvalidParameterException` exception occurs when neither of the `S3Object` or `Bytes` values are supplied in the `Document` request parameter. Validate your parameter before calling the API operation again.

HTTP Status Code: 400

InvalidS3ObjectException

Amazon Textract is unable to access the S3 object that's specified in the request. for more information, Configure Access to Amazon S3 For troubleshooting information, see Troubleshooting Amazon S3

HTTP Status Code: 400

ProvisionedThroughputExceededException

The number of requests exceeded your throughput limit. If you want to increase this limit, contact Amazon Textract.

HTTP Status Code: 400

ThrottlingException

Amazon Textract is temporarily unable to process the request. Try your call again.

HTTP Status Code: 500

UnsupportedDocumentException

The format of the input document isn't supported. Documents for operations can be in PNG, JPEG, PDF, or TIFF format.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface V2
- AWS SDK for .NET V4
- AWS SDK for C++
- AWS SDK for Go v2
- AWS SDK for Java V2
- AWS SDK for JavaScript V3
- AWS SDK for Kotlin
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3

AnalyzeID

Analyzes identity documents for relevant information. This information is extracted and returned as `IdentityDocumentFields`, which records both the normalized field and value of the extracted text. Unlike other Amazon Textract operations, `AnalyzeID` doesn't return any Geometry data.

Request Syntax

```
{
   "DocumentPages": [
      {
         "Bytes": blob,
         "S3Object": {
            "Bucket": "string",
            "Name": "string",
            "Version": "string"
         }
      }
   ]
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters.

The request accepts the following data in JSON format.

DocumentPages

The document being passed to AnalyzeID.

Type: Array of Document objects

Array Members: Minimum number of 1 item. Maximum number of 2 items.

Required: Yes

Response Syntax

```
{
```

```
   "AnalyzeIDModelVersion": "string",
   "DocumentMetadata": {
      "Pages": number
   },
   "IdentityDocuments": [
      {
         "Blocks": [
            {
               "BlockType": "string",
               "ColumnIndex": number,
               "ColumnSpan": number,
               "Confidence": number,
               "EntityTypes": [ "string" ],
               "Geometry": {
                  "BoundingBox": {
                     "Height": number,
                     "Left": number,
                     "Top": number,
                     "Width": number
                  },
                  "Polygon": [
                     {
                        "X": number,
                        "Y": number
                     }
                  ],
                  "RotationAngle": number
               },
               "Id": "string",
               "Page": number,
               "Query": {
                  "Alias": "string",
                  "Pages": [ "string" ],
                  "Text": "string"
               },
               "Relationships": [
                  {
                     "Ids": [ "string" ],
                     "Type": "string"
                  }
               ],
               "RowIndex": number,
               "RowSpan": number,
               "SelectionStatus": "string",
```

```
                "Text": "string",
                "TextType": "string"
             }
          ],
          "DocumentIndex": number,
          "IdentityDocumentFields": [
             {
                "Type": {
                   "Confidence": number,
                   "NormalizedValue": {
                      "Value": "string",
                      "ValueType": "string"
                   },
                   "Text": "string"
                },
                "ValueDetection": {
                   "Confidence": number,
                   "NormalizedValue": {
                      "Value": "string",
                      "ValueType": "string"
                   },
                   "Text": "string"
                }
             }
          ]
       }
    ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

AnalyzeIDModelVersion

The version of the AnalyzeIdentity API being used to process documents.

Type: String

DocumentMetadata

Information about the input document.

Type: DocumentMetadata object

IdentityDocuments

The list of documents processed by AnalyzeID. Includes a number denoting their place in the list and the response structure for the document.

Type: Array of IdentityDocument objects

Errors

For information about the errors that are common to all actions, see Common Error Types.

AccessDeniedException

You aren't authorized to perform the action. Use the Amazon Resource Name (ARN) of an authorized user or IAM role to perform the operation.

HTTP Status Code: 400

BadDocumentException

Amazon Textract isn't able to read the document. For more information on the document limits in Amazon Textract, see Hard limits.

HTTP Status Code: 400

DocumentTooLargeException

The document can't be processed because it's too large. The maximum document size for synchronous operations 10 MB. The maximum document size for asynchronous operations is 500 MB for PDF files.

HTTP Status Code: 400

InternalServerError

Amazon Textract experienced a service issue. Try your call again.

HTTP Status Code: 500

InvalidParameterException

An input parameter violated a constraint. For example, in synchronous operations, an `InvalidParameterException` exception occurs when neither of the `S3Object` or `Bytes`

values are supplied in the `Document` request parameter. Validate your parameter before calling the API operation again.

HTTP Status Code: 400

InvalidS3ObjectException

Amazon Textract is unable to access the S3 object that's specified in the request. for more information, Configure Access to Amazon S3 For troubleshooting information, see Troubleshooting Amazon S3

HTTP Status Code: 400

ProvisionedThroughputExceededException

The number of requests exceeded your throughput limit. If you want to increase this limit, contact Amazon Textract.

HTTP Status Code: 400

ThrottlingException

Amazon Textract is temporarily unable to process the request. Try your call again.

HTTP Status Code: 500

UnsupportedDocumentException

The format of the input document isn't supported. Documents for operations can be in PNG, JPEG, PDF, or TIFF format.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface V2
- AWS SDK for .NET V4
- AWS SDK for C++
- AWS SDK for Go v2

- AWS SDK for Java V2
- AWS SDK for JavaScript V3
- AWS SDK for Kotlin
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3

CreateAdapter

Creates an adapter, which can be fine-tuned for enhanced performance on user provided documents. Takes an AdapterName and FeatureType. Currently the only supported feature type is QUERIES. You can also provide a Description, Tags, and a ClientRequestToken. You can choose whether or not the adapter should be AutoUpdated with the AutoUpdate argument. By default, AutoUpdate is set to DISABLED.

Request Syntax

```
{
   "AdapterName": "string",
   "AutoUpdate": "string",
   "ClientRequestToken": "string",
   "Description": "string",
   "FeatureTypes": [ "string" ],
   "Tags": {
      "string" : "string"
   }
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters.

The request accepts the following data in JSON format.

AdapterName

The name to be assigned to the adapter being created.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Pattern: `[a-zA-Z0-9-_]+`

Required: Yes

AutoUpdate

Controls whether or not the adapter should automatically update.

Type: String

Valid Values: `ENABLED | DISABLED`

Required: No

ClientRequestToken

Idempotent token is used to recognize the request. If the same token is used with multiple CreateAdapter requests, the same session is returned. This token is employed to avoid unintentionally creating the same session multiple times.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: `^[a-zA-Z0-9-_]+$`

Required: No

Description

The description to be assigned to the adapter being created.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: `^[a-zA-Z0-9\s!"\#\$%'&\(\)\*\+\,\-\./:;=\?@\[\\\]\^_`\{\|\}~><]+$`

Required: No

FeatureTypes

The type of feature that the adapter is being trained on. Currrenly, supported feature types are: `QUERIES`

Type: Array of strings

Valid Values: `TABLES | FORMS | QUERIES | SIGNATURES | LAYOUT`

Required: Yes

Tags

A list of tags to be added to the adapter.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 200 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Key Pattern: `^(?!aws:)[\p{L}\p{Z}\p{N}_.:/=+\-@]*$`

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Value Pattern: `^([\p{L}\p{Z}\p{N}_.:/=+\-@]*)$`

Required: No

Response Syntax

```
{
   "AdapterId": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

AdapterId

A string containing the unique ID for the adapter that has been created.

Type: String

Length Constraints: Minimum length of 12. Maximum length of 1011.

Errors

For information about the errors that are common to all actions, see Common Error Types.

AccessDeniedException

You aren't authorized to perform the action. Use the Amazon Resource Name (ARN) of an authorized user or IAM role to perform the operation.

HTTP Status Code: 400

ConflictException

Updating or deleting a resource can cause an inconsistent state.

HTTP Status Code: 400

IdempotentParameterMismatchException

A `ClientRequestToken` input parameter was reused with an operation, but at least one of the other input parameters is different from the previous call to the operation.

HTTP Status Code: 400

InternalServerError

Amazon Textract experienced a service issue. Try your call again.

HTTP Status Code: 500

InvalidParameterException

An input parameter violated a constraint. For example, in synchronous operations, an `InvalidParameterException` exception occurs when neither of the `S3Object` or `Bytes` values are supplied in the `Document` request parameter. Validate your parameter before calling the API operation again.

HTTP Status Code: 400

LimitExceededException

An Amazon Textract service limit was exceeded. For example, if you start too many asynchronous jobs concurrently, calls to start operations (`StartDocumentTextDetection`, for example) raise a LimitExceededException exception (HTTP status code: 400) until the number of concurrently running jobs is below the Amazon Textract service limit.

HTTP Status Code: 400

ProvisionedThroughputExceededException

The number of requests exceeded your throughput limit. If you want to increase this limit, contact Amazon Textract.

HTTP Status Code: 400

ServiceQuotaExceededException

Returned when a request cannot be completed as it would exceed a maximum service quota.

HTTP Status Code: 400

ThrottlingException

Amazon Textract is temporarily unable to process the request. Try your call again.

HTTP Status Code: 500

ValidationException

Indicates that a request was not valid. Check request for proper formatting.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface V2
- AWS SDK for .NET V4
- AWS SDK for C++
- AWS SDK for Go v2
- AWS SDK for Java V2
- AWS SDK for JavaScript V3
- AWS SDK for Kotlin
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3

CreateAdapterVersion

Creates a new version of an adapter. Operates on a provided AdapterId and a specified dataset provided via the DatasetConfig argument. Requires that you specify an Amazon S3 bucket with the OutputConfig argument. You can provide an optional KMSKeyId, an optional ClientRequestToken, and optional tags.

Request Syntax

```
{
   "AdapterId": "string",
   "ClientRequestToken": "string",
   "DatasetConfig": {
      "ManifestS3Object": {
         "Bucket": "string",
         "Name": "string",
         "Version": "string"
      }
   },
   "KMSKeyId": "string",
   "OutputConfig": {
      "S3Bucket": "string",
      "S3Prefix": "string"
   },
   "Tags": {
      "string" : "string"
   }
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters.

The request accepts the following data in JSON format.

AdapterId

A string containing a unique ID for the adapter that will receive a new version.

Type: String

Length Constraints: Minimum length of 12. Maximum length of 1011.

Required: Yes

ClientRequestToken

Idempotent token is used to recognize the request. If the same token is used with multiple CreateAdapterVersion requests, the same session is returned. This token is employed to avoid unintentionally creating the same session multiple times.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: `^[a-zA-Z0-9-_]+$`

Required: No

DatasetConfig

Specifies a dataset used to train a new adapter version. Takes a ManifestS3Object as the value.

Type: AdapterVersionDatasetConfig object

Required: Yes

KMSKeyId

The identifier for your AWS Key Management Service key (AWS KMS key). Used to encrypt your documents.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `^[A-Za-z0-9][A-Za-z0-9:_/+=,@.-]{0,2048}$`

Required: No

OutputConfig

Sets whether or not your output will go to a user created bucket. Used to set the name of the bucket, and the prefix on the output file.

`OutputConfig` is an optional parameter which lets you adjust where your output will be placed. By default, Amazon Textract will store the results internally and can only be accessed by the Get API operations. With `OutputConfig` enabled, you can set the name of the bucket

the output will be sent to the file prefix of the results where you can download your results. Additionally, you can set the KMSKeyID parameter to a customer master key (CMK) to encrypt your output. Without this parameter set Amazon Textract will encrypt server-side using the AWS managed CMK for Amazon S3.

Decryption of Customer Content is necessary for processing of the documents by Amazon Textract. If your account is opted out under an AI services opt out policy then all unencrypted Customer Content is immediately and permanently deleted after the Customer Content has been processed by the service. No copy of of the output is retained by Amazon Textract. For information about how to opt out, see Managing AI services opt-out policy.

For more information on data privacy, see the Data Privacy FAQ.

Type: OutputConfig object

Required: Yes

Tags

A set of tags (key-value pairs) that you want to attach to the adapter version.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 200 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Key Pattern: `^(?!aws:)[\p{L}\p{Z}\p{N}_.:/=+\-@]*$`

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Value Pattern: `^([\p{L}\p{Z}\p{N}_.:/=+\-@]*)$`

Required: No

Response Syntax

```
{
   "AdapterId": "string",
   "AdapterVersion": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

AdapterId

A string containing the unique ID for the adapter that has received a new version.

Type: String

Length Constraints: Minimum length of 12. Maximum length of 1011.

AdapterVersion

A string describing the new version of the adapter.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Errors

For information about the errors that are common to all actions, see Common Error Types.

AccessDeniedException

You aren't authorized to perform the action. Use the Amazon Resource Name (ARN) of an authorized user or IAM role to perform the operation.

HTTP Status Code: 400

ConflictException

Updating or deleting a resource can cause an inconsistent state.

HTTP Status Code: 400

IdempotentParameterMismatchException

A `ClientRequestToken` input parameter was reused with an operation, but at least one of the other input parameters is different from the previous call to the operation.

HTTP Status Code: 400

InternalServerError

Amazon Textract experienced a service issue. Try your call again.

HTTP Status Code: 500

InvalidKMSKeyException

Indicates you do not have decrypt permissions with the KMS key entered, or the KMS key was entered incorrectly.

HTTP Status Code: 400

InvalidParameterException

An input parameter violated a constraint. For example, in synchronous operations, an `InvalidParameterException` exception occurs when neither of the `S3Object` or `Bytes` values are supplied in the `Document` request parameter. Validate your parameter before calling the API operation again.

HTTP Status Code: 400

InvalidS3ObjectException

Amazon Textract is unable to access the S3 object that's specified in the request. for more information, Configure Access to Amazon S3 For troubleshooting information, see Troubleshooting Amazon S3

HTTP Status Code: 400

LimitExceededException

An Amazon Textract service limit was exceeded. For example, if you start too many asynchronous jobs concurrently, calls to start operations (`StartDocumentTextDetection`, for example) raise a LimitExceededException exception (HTTP status code: 400) until the number of concurrently running jobs is below the Amazon Textract service limit.

HTTP Status Code: 400

ProvisionedThroughputExceededException

The number of requests exceeded your throughput limit. If you want to increase this limit, contact Amazon Textract.

HTTP Status Code: 400

ResourceNotFoundException

Returned when an operation tried to access a nonexistent resource.

HTTP Status Code: 400

ServiceQuotaExceededException

Returned when a request cannot be completed as it would exceed a maximum service quota.

HTTP Status Code: 400

ThrottlingException

Amazon Textract is temporarily unable to process the request. Try your call again.

HTTP Status Code: 500

ValidationException

Indicates that a request was not valid. Check request for proper formatting.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface V2
- AWS SDK for .NET V4
- AWS SDK for C++
- AWS SDK for Go v2
- AWS SDK for Java V2
- AWS SDK for JavaScript V3
- AWS SDK for Kotlin
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3

DeleteAdapter

Deletes an Amazon Textract adapter. Takes an AdapterId and deletes the adapter specified by the ID.

Request Syntax

```
{
   "AdapterId": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters.

The request accepts the following data in JSON format.

AdapterId

A string containing a unique ID for the adapter to be deleted.

Type: String

Length Constraints: Minimum length of 12. Maximum length of 1011.

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Error Types.

AccessDeniedException

You aren't authorized to perform the action. Use the Amazon Resource Name (ARN) of an authorized user or IAM role to perform the operation.

HTTP Status Code: 400

ConflictException

Updating or deleting a resource can cause an inconsistent state.

HTTP Status Code: 400

InternalServerError

Amazon Textract experienced a service issue. Try your call again.

HTTP Status Code: 500

InvalidParameterException

An input parameter violated a constraint. For example, in synchronous operations, an `InvalidParameterException` exception occurs when neither of the `S3Object` or `Bytes` values are supplied in the `Document` request parameter. Validate your parameter before calling the API operation again.

HTTP Status Code: 400

ProvisionedThroughputExceededException

The number of requests exceeded your throughput limit. If you want to increase this limit, contact Amazon Textract.

HTTP Status Code: 400

ResourceNotFoundException

Returned when an operation tried to access a nonexistent resource.

HTTP Status Code: 400

ThrottlingException

Amazon Textract is temporarily unable to process the request. Try your call again.

HTTP Status Code: 500

ValidationException

Indicates that a request was not valid. Check request for proper formatting.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface V2
- AWS SDK for .NET V4
- AWS SDK for C++
- AWS SDK for Go v2
- AWS SDK for Java V2
- AWS SDK for JavaScript V3
- AWS SDK for Kotlin
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3

DeleteAdapterVersion

Deletes an Amazon Textract adapter version. Requires that you specify both an AdapterId and a AdapterVersion. Deletes the adapter version specified by the AdapterId and the AdapterVersion.

Request Syntax

```
{
   "AdapterId": "string",
   "AdapterVersion": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters.

The request accepts the following data in JSON format.

AdapterId

A string containing a unique ID for the adapter version that will be deleted.

Type: String

Length Constraints: Minimum length of 12. Maximum length of 1011.

Required: Yes

AdapterVersion

Specifies the adapter version to be deleted.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Error Types.

AccessDeniedException

You aren't authorized to perform the action. Use the Amazon Resource Name (ARN) of an authorized user or IAM role to perform the operation.

HTTP Status Code: 400

ConflictException

Updating or deleting a resource can cause an inconsistent state.

HTTP Status Code: 400

InternalServerError

Amazon Textract experienced a service issue. Try your call again.

HTTP Status Code: 500

InvalidParameterException

An input parameter violated a constraint. For example, in synchronous operations, an `InvalidParameterException` exception occurs when neither of the `S3Object` or `Bytes` values are supplied in the `Document` request parameter. Validate your parameter before calling the API operation again.

HTTP Status Code: 400

ProvisionedThroughputExceededException

The number of requests exceeded your throughput limit. If you want to increase this limit, contact Amazon Textract.

HTTP Status Code: 400

ResourceNotFoundException

Returned when an operation tried to access a nonexistent resource.

HTTP Status Code: 400

ThrottlingException

Amazon Textract is temporarily unable to process the request. Try your call again.

HTTP Status Code: 500

ValidationException

Indicates that a request was not valid. Check request for proper formatting.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface V2
- AWS SDK for .NET V4
- AWS SDK for C++
- AWS SDK for Go v2
- AWS SDK for Java V2
- AWS SDK for JavaScript V3
- AWS SDK for Kotlin
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3

DetectDocumentText

Detects text in the input document. Amazon Textract can detect lines of text and the words that make up a line of text. The input document must be in one of the following image formats: JPEG, PNG, PDF, or TIFF. `DetectDocumentText` returns the detected text in an array of Block objects.

Each document page has as an associated `Block` of type PAGE. Each PAGE `Block` object is the parent of LINE `Block` objects that represent the lines of detected text on a page. A LINE `Block` object is a parent for each word that makes up the line. Words are represented by `Block` objects of type WORD.

`DetectDocumentText` is a synchronous operation. To analyze documents asynchronously, use StartDocumentTextDetection.

For more information, see Document Text Detection.

Request Syntax

```
{
   "Document": {
      "Bytes": blob,
      "S3Object": {
         "Bucket": "string",
         "Name": "string",
         "Version": "string"
      }
   }
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters.

The request accepts the following data in JSON format.

Document

The input document as base64-encoded bytes or an Amazon S3 object. If you use the AWS CLI to call Amazon Textract operations, you can't pass image bytes. The document must be an image in JPEG or PNG format.

If you're using an AWS SDK to call Amazon Textract, you might not need to base64-encode image bytes that are passed using the `Bytes` field.

Type: Document object

Required: Yes

Response Syntax

```
{
   "Blocks": [
      {
         "BlockType": "string",
         "ColumnIndex": number,
         "ColumnSpan": number,
         "Confidence": number,
         "EntityTypes": [ "string" ],
         "Geometry": {
            "BoundingBox": {
               "Height": number,
               "Left": number,
               "Top": number,
               "Width": number
            },
            "Polygon": [
               {
                  "X": number,
                  "Y": number
               }
            ],
            "RotationAngle": number
         },
         "Id": "string",
         "Page": number,
         "Query": {
            "Alias": "string",
            "Pages": [ "string" ],
            "Text": "string"
         },
         "Relationships": [
            {
               "Ids": [ "string" ],
```

```
                "Type": "string"
             }
          ],
          "RowIndex": number,
          "RowSpan": number,
          "SelectionStatus": "string",
          "Text": "string",
          "TextType": "string"
       }
    ],
    "DetectDocumentTextModelVersion": "string",
    "DocumentMetadata": {
       "Pages": number
    }
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Blocks

An array of `Block` objects that contain the text that's detected in the document.

Type: Array of Block objects

DetectDocumentTextModelVersion

Type: String

DocumentMetadata

Metadata about the document. It contains the number of pages that are detected in the document.

Type: DocumentMetadata object

Errors

For information about the errors that are common to all actions, see Common Error Types.

AccessDeniedException

You aren't authorized to perform the action. Use the Amazon Resource Name (ARN) of an authorized user or IAM role to perform the operation.

HTTP Status Code: 400

BadDocumentException

Amazon Textract isn't able to read the document. For more information on the document limits in Amazon Textract, see Hard limits.

HTTP Status Code: 400

DocumentTooLargeException

The document can't be processed because it's too large. The maximum document size for synchronous operations 10 MB. The maximum document size for asynchronous operations is 500 MB for PDF files.

HTTP Status Code: 400

InternalServerError

Amazon Textract experienced a service issue. Try your call again.

HTTP Status Code: 500

InvalidParameterException

An input parameter violated a constraint. For example, in synchronous operations, an `InvalidParameterException` exception occurs when neither of the `S3Object` or `Bytes` values are supplied in the `Document` request parameter. Validate your parameter before calling the API operation again.

HTTP Status Code: 400

InvalidS3ObjectException

Amazon Textract is unable to access the S3 object that's specified in the request. for more information, Configure Access to Amazon S3 For troubleshooting information, see Troubleshooting Amazon S3

HTTP Status Code: 400

ProvisionedThroughputExceededException

The number of requests exceeded your throughput limit. If you want to increase this limit, contact Amazon Textract.

HTTP Status Code: 400

ThrottlingException

Amazon Textract is temporarily unable to process the request. Try your call again.

HTTP Status Code: 500

UnsupportedDocumentException

The format of the input document isn't supported. Documents for operations can be in PNG, JPEG, PDF, or TIFF format.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface V2
- AWS SDK for .NET V4
- AWS SDK for C++
- AWS SDK for Go v2
- AWS SDK for Java V2
- AWS SDK for JavaScript V3
- AWS SDK for Kotlin
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3

GetAdapter

Gets configuration information for an adapter specified by an AdapterId, returning information on AdapterName, Description, CreationTime, AutoUpdate status, and FeatureTypes.

Request Syntax

```
{
   "AdapterId": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters.

The request accepts the following data in JSON format.

AdapterId

A string containing a unique ID for the adapter.

Type: String

Length Constraints: Minimum length of 12. Maximum length of 1011.

Required: Yes

Response Syntax

```
{
   "AdapterId": "string",
   "AdapterName": "string",
   "AutoUpdate": "string",
   "CreationTime": number,
   "Description": "string",
   "FeatureTypes": [ "string" ],
   "Tags": {
      "string" : "string"
   }
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

AdapterId

A string identifying the adapter that information has been retrieved for.

Type: String

Length Constraints: Minimum length of 12. Maximum length of 1011.

AdapterName

The name of the requested adapter.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Pattern: `[a-zA-Z0-9-_]+`

AutoUpdate

Binary value indicating if the adapter is being automatically updated or not.

Type: String

Valid Values: `ENABLED | DISABLED`

CreationTime

The date and time the requested adapter was created at.

Type: Timestamp

Description

The description for the requested adapter.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: `^[a-zA-Z0-9\s!"\#\$%'&\(\)\*\+\,\-\./:;=\?@\[\\\]\^_`\{\|\}~><]+$`

FeatureTypes

List of the targeted feature types for the requested adapter.

Type: Array of strings

Valid Values: `TABLES | FORMS | QUERIES | SIGNATURES | LAYOUT`

Tags

A set of tags (key-value pairs) associated with the adapter that has been retrieved.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 200 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Key Pattern: `^(?!aws:)[\p{L}\p{Z}\p{N}_.:/=+\-@]*$`

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Value Pattern: `^([\p{L}\p{Z}\p{N}_.:/=+\-@]*)$`

Errors

For information about the errors that are common to all actions, see Common Error Types.

AccessDeniedException

You aren't authorized to perform the action. Use the Amazon Resource Name (ARN) of an authorized user or IAM role to perform the operation.

HTTP Status Code: 400

InternalServerError

Amazon Textract experienced a service issue. Try your call again.

HTTP Status Code: 500

InvalidParameterException

An input parameter violated a constraint. For example, in synchronous operations, an `InvalidParameterException` exception occurs when neither of the `S3Object` or `Bytes`

values are supplied in the `Document` request parameter. Validate your parameter before calling the API operation again.

HTTP Status Code: 400

ProvisionedThroughputExceededException

The number of requests exceeded your throughput limit. If you want to increase this limit, contact Amazon Textract.

HTTP Status Code: 400

ResourceNotFoundException

Returned when an operation tried to access a nonexistent resource.

HTTP Status Code: 400

ThrottlingException

Amazon Textract is temporarily unable to process the request. Try your call again.

HTTP Status Code: 500

ValidationException

Indicates that a request was not valid. Check request for proper formatting.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface V2
- AWS SDK for .NET V4
- AWS SDK for C++
- AWS SDK for Go v2
- AWS SDK for Java V2
- AWS SDK for JavaScript V3
- AWS SDK for Kotlin

- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3

GetAdapterVersion

Gets configuration information for the specified adapter version, including: AdapterId, AdapterVersion, FeatureTypes, Status, StatusMessage, DatasetConfig, KMSKeyId, OutputConfig, Tags and EvaluationMetrics.

Request Syntax

```
{
   "AdapterId": "string",
   "AdapterVersion": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters.

The request accepts the following data in JSON format.

AdapterId

A string specifying a unique ID for the adapter version you want to retrieve information for.

Type: String

Length Constraints: Minimum length of 12. Maximum length of 1011.

Required: Yes

AdapterVersion

A string specifying the adapter version you want to retrieve information for.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Required: Yes

Response Syntax

```
{
```

```
   "AdapterId": "string",
   "AdapterVersion": "string",
   "CreationTime": number,
   "DatasetConfig": {
      "ManifestS3Object": {
         "Bucket": "string",
         "Name": "string",
         "Version": "string"
      }
   },
   "EvaluationMetrics": [
      {
         "AdapterVersion": {
            "F1Score": number,
            "Precision": number,
            "Recall": number
         },
         "Baseline": {
            "F1Score": number,
            "Precision": number,
            "Recall": number
         },
         "FeatureType": "string"
      }
   ],
   "FeatureTypes": [ "string" ],
   "KMSKeyId": "string",
   "OutputConfig": {
      "S3Bucket": "string",
      "S3Prefix": "string"
   },
   "Status": "string",
   "StatusMessage": "string",
   "Tags": {
      "string" : "string"
   }
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

AdapterId

A string containing a unique ID for the adapter version being retrieved.

Type: String

Length Constraints: Minimum length of 12. Maximum length of 1011.

AdapterVersion

A string containing the adapter version that has been retrieved.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

CreationTime

The time that the adapter version was created.

Type: Timestamp

DatasetConfig

Specifies a dataset used to train a new adapter version. Takes a ManifestS3Objec as the value.

Type: AdapterVersionDatasetConfig object

EvaluationMetrics

The evaluation metrics (F1 score, Precision, and Recall) for the requested version, grouped by baseline metrics and adapter version.

Type: Array of AdapterVersionEvaluationMetric objects

FeatureTypes

List of the targeted feature types for the requested adapter version.

Type: Array of strings

Valid Values: `TABLES | FORMS | QUERIES | SIGNATURES | LAYOUT`

KMSKeyId

The identifier for your AWS Key Management Service key (AWS KMS key). Used to encrypt your documents.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `^[A-Za-z0-9][A-Za-z0-9:_/+=,@.-]{0,2048}$`

OutputConfig

Sets whether or not your output will go to a user created bucket. Used to set the name of the bucket, and the prefix on the output file.

`OutputConfig` is an optional parameter which lets you adjust where your output will be placed. By default, Amazon Textract will store the results internally and can only be accessed by the Get API operations. With `OutputConfig` enabled, you can set the name of the bucket the output will be sent to the file prefix of the results where you can download your results. Additionally, you can set the `KMSKeyID` parameter to a customer master key (CMK) to encrypt your output. Without this parameter set Amazon Textract will encrypt server-side using the AWS managed CMK for Amazon S3.

Decryption of Customer Content is necessary for processing of the documents by Amazon Textract. If your account is opted out under an AI services opt out policy then all unencrypted Customer Content is immediately and permanently deleted after the Customer Content has been processed by the service. No copy of of the output is retained by Amazon Textract. For information about how to opt out, see Managing AI services opt-out policy.

For more information on data privacy, see the Data Privacy FAQ.

Type: OutputConfig object

Status

The status of the adapter version that has been requested.

Type: String

Valid Values: `ACTIVE | AT_RISK | DEPRECATED | CREATION_ERROR | CREATION_IN_PROGRESS`

StatusMessage

A message that describes the status of the requested adapter version.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: `^[a-zA-Z0-9\s!"\#\$%'&\(\)\*\+\,\-\./:;=\?@\[\\\]\^_`\{\|\}~><]+$`

Tags

A set of tags (key-value pairs) that are associated with the adapter version.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 200 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Key Pattern: `^(?!aws:)[\p{L}\p{Z}\p{N}_.:/=+\-@]*$`

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Value Pattern: `^([\p{L}\p{Z}\p{N}_.:/=+\-@]*)$`

Errors

For information about the errors that are common to all actions, see Common Error Types.

AccessDeniedException

You aren't authorized to perform the action. Use the Amazon Resource Name (ARN) of an authorized user or IAM role to perform the operation.

HTTP Status Code: 400

InternalServerError

Amazon Textract experienced a service issue. Try your call again.

HTTP Status Code: 500

InvalidParameterException

An input parameter violated a constraint. For example, in synchronous operations, an `InvalidParameterException` exception occurs when neither of the `S3Object` or `Bytes` values are supplied in the `Document` request parameter. Validate your parameter before calling the API operation again.

HTTP Status Code: 400

ProvisionedThroughputExceededException

The number of requests exceeded your throughput limit. If you want to increase this limit, contact Amazon Textract.

HTTP Status Code: 400

ResourceNotFoundException

Returned when an operation tried to access a nonexistent resource.

HTTP Status Code: 400

ThrottlingException

Amazon Textract is temporarily unable to process the request. Try your call again.

HTTP Status Code: 500

ValidationException

Indicates that a request was not valid. Check request for proper formatting.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface V2
- AWS SDK for .NET V4
- AWS SDK for C++
- AWS SDK for Go v2
- AWS SDK for Java V2
- AWS SDK for JavaScript V3
- AWS SDK for Kotlin
- AWS SDK for PHP V3
- AWS SDK for Python

- AWS SDK for Ruby V3

GetDocumentAnalysis

Gets the results for an Amazon Textract asynchronous operation that analyzes text in a document.

You start asynchronous text analysis by calling StartDocumentAnalysis, which returns a job identifier (`JobId`). When the text analysis operation finishes, Amazon Textract publishes a completion status to the Amazon Simple Notification Service (Amazon SNS) topic that's registered in the initial call to `StartDocumentAnalysis`. To get the results of the text-detection operation, first check that the status value published to the Amazon SNS topic is SUCCEEDED. If so, call `GetDocumentAnalysis`, and pass the job identifier (`JobId`) from the initial call to `StartDocumentAnalysis`.

`GetDocumentAnalysis` returns an array of Block objects. The following types of information are returned:

- Form data (key-value pairs). The related information is returned in two Block objects, each of type `KEY_VALUE_SET`: a KEY `Block` object and a VALUE `Block` object. For example, *Name: Ana Silva Carolina* contains a key and value. *Name:* is the key. *Ana Silva Carolina* is the value.
- Table and table cell data. A TABLE `Block` object contains information about a detected table. A CELL `Block` object is returned for each cell in a table.
- Lines and words of text. A LINE `Block` object contains one or more WORD `Block` objects. All lines and words that are detected in the document are returned (including text that doesn't have a relationship with the value of the `StartDocumentAnalysis` `FeatureTypes` input parameter).
- Query. A QUERY Block object contains the query text, alias and link to the associated Query results block object.
- Query Results. A QUERY_RESULT Block object contains the answer to the query and an ID that connects it to the query asked. This Block also contains a confidence score.

> **Note**
>
> While processing a document with queries, look out for `INVALID_REQUEST_PARAMETERS` output. This indicates that either the per page query limit has been exceeded or that the operation is trying to query a page in the document which doesn’t exist.

Selection elements such as check boxes and option buttons (radio buttons) can be detected in form data and in tables. A SELECTION_ELEMENT `Block` object contains information about a selection element, including the selection status.

Use the `MaxResults` parameter to limit the number of blocks that are returned. If there are more results than specified in `MaxResults`, the value of `NextToken` in the operation response contains a pagination token for getting the next set of results. To get the next page of results, call `GetDocumentAnalysis`, and populate the `NextToken` request parameter with the token value that's returned from the previous call to `GetDocumentAnalysis`.

For more information, see Document Text Analysis.

Request Syntax

```
{
   "JobId": "string",
   "MaxResults": number,
   "NextToken": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters.

The request accepts the following data in JSON format.

JobId

A unique identifier for the text-detection job. The `JobId` is returned from `StartDocumentAnalysis`. A `JobId` value is only valid for 7 days.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: `^[a-zA-Z0-9-_]+$`

Required: Yes

MaxResults

The maximum number of results to return per paginated call. The largest value that you can specify is 1,000. If you specify a value greater than 1,000, a maximum of 1,000 results is returned. The default value is 1,000.

Type: Integer

Valid Range: Minimum value of 1.

Required: No

NextToken

If the previous response was incomplete (because there are more blocks to retrieve), Amazon Textract returns a pagination token in the response. You can use this pagination token to retrieve the next set of blocks.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1024.

Pattern: `.*\S.*`

Required: No

Response Syntax

```
{
   "AnalyzeDocumentModelVersion": "string",
   "Blocks": [
      {
         "BlockType": "string",
         "ColumnIndex": number,
         "ColumnSpan": number,
         "Confidence": number,
         "EntityTypes": [ "string" ],
         "Geometry": {
            "BoundingBox": {
               "Height": number,
               "Left": number,
               "Top": number,
               "Width": number
            },
```

```
            "Polygon": [
               {
                  "X": number,
                  "Y": number
               }
            ],
            "RotationAngle": number
         },
         "Id": "string",
         "Page": number,
         "Query": {
            "Alias": "string",
            "Pages": [ "string" ],
            "Text": "string"
         },
         "Relationships": [
            {
               "Ids": [ "string" ],
               "Type": "string"
            }
         ],
         "RowIndex": number,
         "RowSpan": number,
         "SelectionStatus": "string",
         "Text": "string",
         "TextType": "string"
      }
   ],
   "DocumentMetadata": {
      "Pages": number
   },
   "JobStatus": "string",
   "NextToken": "string",
   "StatusMessage": "string",
   "Warnings": [
      {
         "ErrorCode": "string",
         "Pages": [ number ]
      }
   ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

AnalyzeDocumentModelVersion

Type: String

Blocks

The results of the text-analysis operation.

Type: Array of Block objects

DocumentMetadata

Information about a document that Amazon Textract processed. `DocumentMetadata` is returned in every page of paginated responses from an Amazon Textract video operation.

Type: DocumentMetadata object

JobStatus

The current status of the text detection job.

Type: String

Valid Values: `IN_PROGRESS | SUCCEEDED | FAILED | PARTIAL_SUCCESS`

NextToken

If the response is truncated, Amazon Textract returns this token. You can use this token in the subsequent request to retrieve the next set of text detection results.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1024.

Pattern: `.*\S.*`

StatusMessage

Returns if the detection job could not be completed. Contains explanation for what error occured.

Type: String

Warnings

A list of warnings that occurred during the document-analysis operation.

Type: Array of Warning objects

Errors

For information about the errors that are common to all actions, see Common Error Types.

AccessDeniedException

You aren't authorized to perform the action. Use the Amazon Resource Name (ARN) of an authorized user or IAM role to perform the operation.

HTTP Status Code: 400

InternalServerError

Amazon Textract experienced a service issue. Try your call again.

HTTP Status Code: 500

InvalidJobIdException

An invalid job identifier was passed to an asynchronous analysis operation.

HTTP Status Code: 400

InvalidKMSKeyException

Indicates you do not have decrypt permissions with the KMS key entered, or the KMS key was entered incorrectly.

HTTP Status Code: 400

InvalidParameterException

An input parameter violated a constraint. For example, in synchronous operations, an `InvalidParameterException` exception occurs when neither of the `S3Object` or `Bytes` values are supplied in the `Document` request parameter. Validate your parameter before calling the API operation again.

HTTP Status Code: 400

InvalidS3ObjectException

Amazon Textract is unable to access the S3 object that's specified in the request. for more information, Configure Access to Amazon S3 For troubleshooting information, see Troubleshooting Amazon S3

HTTP Status Code: 400

ProvisionedThroughputExceededException

The number of requests exceeded your throughput limit. If you want to increase this limit, contact Amazon Textract.

HTTP Status Code: 400

ThrottlingException

Amazon Textract is temporarily unable to process the request. Try your call again.

HTTP Status Code: 500

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface V2
- AWS SDK for .NET V4
- AWS SDK for C++
- AWS SDK for Go v2
- AWS SDK for Java V2
- AWS SDK for JavaScript V3
- AWS SDK for Kotlin
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3

GetDocumentTextDetection

Gets the results for an Amazon Textract asynchronous operation that detects text in a document. Amazon Textract can detect lines of text and the words that make up a line of text.

You start asynchronous text detection by calling StartDocumentTextDetection, which returns a job identifier (`JobId`). When the text detection operation finishes, Amazon Textract publishes a completion status to the Amazon Simple Notification Service (Amazon SNS) topic that's registered in the initial call to `StartDocumentTextDetection`. To get the results of the text-detection operation, first check that the status value published to the Amazon SNS topic is SUCCEEDED. If so, call `GetDocumentTextDetection`, and pass the job identifier (`JobId`) from the initial call to `StartDocumentTextDetection`.

`GetDocumentTextDetection` returns an array of Block objects.

Each document page has as an associated `Block` of type PAGE. Each PAGE `Block` object is the parent of LINE `Block` objects that represent the lines of detected text on a page. A LINE `Block` object is a parent for each word that makes up the line. Words are represented by `Block` objects of type WORD.

Use the MaxResults parameter to limit the number of blocks that are returned. If there are more results than specified in `MaxResults`, the value of `NextToken` in the operation response contains a pagination token for getting the next set of results. To get the next page of results, call `GetDocumentTextDetection`, and populate the `NextToken` request parameter with the token value that's returned from the previous call to `GetDocumentTextDetection`.

For more information, see Document Text Detection.

Request Syntax

```
{
   "JobId": "string",
   "MaxResults": number,
   "NextToken": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters.

The request accepts the following data in JSON format.

JobId

A unique identifier for the text detection job. The `JobId` is returned from `StartDocumentTextDetection`. A `JobId` value is only valid for 7 days.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: `^[a-zA-Z0-9-_]+$`

Required: Yes

MaxResults

The maximum number of results to return per paginated call. The largest value you can specify is 1,000. If you specify a value greater than 1,000, a maximum of 1,000 results is returned. The default value is 1,000.

Type: Integer

Valid Range: Minimum value of 1.

Required: No

NextToken

If the previous response was incomplete (because there are more blocks to retrieve), Amazon Textract returns a pagination token in the response. You can use this pagination token to retrieve the next set of blocks.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1024.

Pattern: `.*\S.*`

Required: No

Response Syntax

```
{
   "Blocks": [
```

```
      {
         "BlockType": "string",
         "ColumnIndex": number,
         "ColumnSpan": number,
         "Confidence": number,
         "EntityTypes": [ "string" ],
         "Geometry": {
            "BoundingBox": {
               "Height": number,
               "Left": number,
               "Top": number,
               "Width": number
            },
            "Polygon": [
               {
                  "X": number,
                  "Y": number
               }
            ],
            "RotationAngle": number
         },
         "Id": "string",
         "Page": number,
         "Query": {
            "Alias": "string",
            "Pages": [ "string" ],
            "Text": "string"
         },
         "Relationships": [
            {
               "Ids": [ "string" ],
               "Type": "string"
            }
         ],
         "RowIndex": number,
         "RowSpan": number,
         "SelectionStatus": "string",
         "Text": "string",
         "TextType": "string"
      }
   ],
   "DetectDocumentTextModelVersion": "string",
   "DocumentMetadata": {
      "Pages": number
```

```
   },
   "JobStatus": "string",
   "NextToken": "string",
   "StatusMessage": "string",
   "Warnings": [
      {
         "ErrorCode": "string",
         "Pages": [ number ]
      }
   ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Blocks

The results of the text-detection operation.

Type: Array of Block objects

DetectDocumentTextModelVersion

Type: String

DocumentMetadata

Information about a document that Amazon Textract processed. `DocumentMetadata` is returned in every page of paginated responses from an Amazon Textract video operation.

Type: DocumentMetadata object

JobStatus

The current status of the text detection job.

Type: String

Valid Values: `IN_PROGRESS | SUCCEEDED | FAILED | PARTIAL_SUCCESS`

NextToken

If the response is truncated, Amazon Textract returns this token. You can use this token in the subsequent request to retrieve the next set of text-detection results.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1024.

Pattern: `.*\S.*`

StatusMessage

Returns if the detection job could not be completed. Contains explanation for what error occured.

Type: String

Warnings

A list of warnings that occurred during the text-detection operation for the document.

Type: Array of Warning objects

Errors

For information about the errors that are common to all actions, see Common Error Types.

AccessDeniedException

You aren't authorized to perform the action. Use the Amazon Resource Name (ARN) of an authorized user or IAM role to perform the operation.

HTTP Status Code: 400

InternalServerError

Amazon Textract experienced a service issue. Try your call again.

HTTP Status Code: 500

InvalidJobIdException

An invalid job identifier was passed to an asynchronous analysis operation.

HTTP Status Code: 400

InvalidKMSKeyException

Indicates you do not have decrypt permissions with the KMS key entered, or the KMS key was entered incorrectly.

HTTP Status Code: 400

InvalidParameterException

An input parameter violated a constraint. For example, in synchronous operations, an `InvalidParameterException` exception occurs when neither of the `S3Object` or `Bytes` values are supplied in the `Document` request parameter. Validate your parameter before calling the API operation again.

HTTP Status Code: 400

InvalidS3ObjectException

Amazon Textract is unable to access the S3 object that's specified in the request. for more information, Configure Access to Amazon S3 For troubleshooting information, see Troubleshooting Amazon S3

HTTP Status Code: 400

ProvisionedThroughputExceededException

The number of requests exceeded your throughput limit. If you want to increase this limit, contact Amazon Textract.

HTTP Status Code: 400

ThrottlingException

Amazon Textract is temporarily unable to process the request. Try your call again.

HTTP Status Code: 500

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface V2
- AWS SDK for .NET V4
- AWS SDK for C++
- AWS SDK for Go v2
- AWS SDK for Java V2
- AWS SDK for JavaScript V3
- AWS SDK for Kotlin
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3

GetExpenseAnalysis

Gets the results for an Amazon Textract asynchronous operation that analyzes invoices and receipts. Amazon Textract finds contact information, items purchased, and vendor name, from input invoices and receipts.

You start asynchronous invoice/receipt analysis by calling StartExpenseAnalysis, which returns a job identifier (`JobId`). Upon completion of the invoice/receipt analysis, Amazon Textract publishes the completion status to the Amazon Simple Notification Service (Amazon SNS) topic. This topic must be registered in the initial call to `StartExpenseAnalysis`. To get the results of the invoice/receipt analysis operation, first ensure that the status value published to the Amazon SNS topic is SUCCEEDED. If so, call `GetExpenseAnalysis`, and pass the job identifier (`JobId`) from the initial call to `StartExpenseAnalysis`.

Use the MaxResults parameter to limit the number of blocks that are returned. If there are more results than specified in `MaxResults`, the value of `NextToken` in the operation response contains a pagination token for getting the next set of results. To get the next page of results, call `GetExpenseAnalysis`, and populate the `NextToken` request parameter with the token value that's returned from the previous call to `GetExpenseAnalysis`.

For more information, see Analyzing Invoices and Receipts.

Request Syntax

```
{
   "JobId": "string",
   "MaxResults": number,
   "NextToken": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters.

The request accepts the following data in JSON format.

JobId

A unique identifier for the text detection job. The `JobId` is returned from `StartExpenseAnalysis`. A `JobId` value is only valid for 7 days.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: `^[a-zA-Z0-9-_]+$`

Required: Yes

MaxResults

The maximum number of results to return per paginated call. The largest value you can specify is 20. If you specify a value greater than 20, a maximum of 20 results is returned. The default value is 20.

Type: Integer

Valid Range: Minimum value of 1.

Required: No

NextToken

If the previous response was incomplete (because there are more blocks to retrieve), Amazon Textract returns a pagination token in the response. You can use this pagination token to retrieve the next set of blocks.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1024.

Pattern: `.*\S.*`

Required: No

Response Syntax

```
{
   "AnalyzeExpenseModelVersion": "string",
   "DocumentMetadata": {
      "Pages": number
   },
   "ExpenseDocuments": [
      {
         "Blocks": [
            {
```

```
            "BlockType": "string",
            "ColumnIndex": number,
            "ColumnSpan": number,
            "Confidence": number,
            "EntityTypes": [ "string" ],
            "Geometry": {
               "BoundingBox": {
                  "Height": number,
                  "Left": number,
                  "Top": number,
                  "Width": number
               },
               "Polygon": [
                  {
                     "X": number,
                     "Y": number
                  }
               ],
               "RotationAngle": number
            },
            "Id": "string",
            "Page": number,
            "Query": {
               "Alias": "string",
               "Pages": [ "string" ],
               "Text": "string"
            },
            "Relationships": [
               {
                  "Ids": [ "string" ],
                  "Type": "string"
               }
            ],
            "RowIndex": number,
            "RowSpan": number,
            "SelectionStatus": "string",
            "Text": "string",
            "TextType": "string"
         }
      ],
      "ExpenseIndex": number,
      "LineItemGroups": [
         {
            "LineItemGroupIndex": number,
```

```
            "LineItems": [
               {
                  "LineItemExpenseFields": [
                     {
                        "Currency": {
                           "Code": "string",
                           "Confidence": number
                        },
                        "GroupProperties": [
                           {
                              "Id": "string",
                              "Types": [ "string" ]
                           }
                        ],
                        "LabelDetection": {
                           "Confidence": number,
                           "Geometry": {
                              "BoundingBox": {
                                 "Height": number,
                                 "Left": number,
                                 "Top": number,
                                 "Width": number
                              },
                              "Polygon": [
                                 {
                                    "X": number,
                                    "Y": number
                                 }
                              ],
                              "RotationAngle": number
                           },
                           "Text": "string"
                        },
                        "PageNumber": number,
                        "Type": {
                           "Confidence": number,
                           "Text": "string"
                        },
                        "ValueDetection": {
                           "Confidence": number,
                           "Geometry": {
                              "BoundingBox": {
                                 "Height": number,
                                 "Left": number,
```

```
                                    "Top": number,
                                    "Width": number
                                 },
                                 "Polygon": [
                                    {
                                       "X": number,
                                       "Y": number
                                    }
                                 ],
                                 "RotationAngle": number
                              },
                              "Text": "string"
                           }
                        }
                     ]
                  }
               ]
            }
         ],
         "SummaryFields": [
            {
               "Currency": {
                  "Code": "string",
                  "Confidence": number
               },
               "GroupProperties": [
                  {
                     "Id": "string",
                     "Types": [ "string" ]
                  }
               ],
               "LabelDetection": {
                  "Confidence": number,
                  "Geometry": {
                     "BoundingBox": {
                        "Height": number,
                        "Left": number,
                        "Top": number,
                        "Width": number
                     },
                     "Polygon": [
                        {
                           "X": number,
                           "Y": number
```

```
                        }
                     ],
                     "RotationAngle": number
                  },
                  "Text": "string"
               },
               "PageNumber": number,
               "Type": {
                  "Confidence": number,
                  "Text": "string"
               },
               "ValueDetection": {
                  "Confidence": number,
                  "Geometry": {
                     "BoundingBox": {
                        "Height": number,
                        "Left": number,
                        "Top": number,
                        "Width": number
                     },
                     "Polygon": [
                        {
                           "X": number,
                           "Y": number
                        }
                     ],
                     "RotationAngle": number
                  },
                  "Text": "string"
               }
            }
         ]
      }
   ],
   "JobStatus": "string",
   "NextToken": "string",
   "StatusMessage": "string",
   "Warnings": [
      {
         "ErrorCode": "string",
         "Pages": [ number ]
      }
   ]
```

```
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

AnalyzeExpenseModelVersion

The current model version of AnalyzeExpense.

Type: String

DocumentMetadata

Information about a document that Amazon Textract processed. `DocumentMetadata` is returned in every page of paginated responses from an Amazon Textract operation.

Type: DocumentMetadata object

ExpenseDocuments

The expenses detected by Amazon Textract.

Type: Array of ExpenseDocument objects

JobStatus

The current status of the text detection job.

Type: String

Valid Values: `IN_PROGRESS | SUCCEEDED | FAILED | PARTIAL_SUCCESS`

NextToken

If the response is truncated, Amazon Textract returns this token. You can use this token in the subsequent request to retrieve the next set of text-detection results.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1024.

Pattern: `.*\S.*`

StatusMessage

Returns if the detection job could not be completed. Contains explanation for what error occured.

Type: String

Warnings

A list of warnings that occurred during the text-detection operation for the document.

Type: Array of Warning objects

Errors

For information about the errors that are common to all actions, see Common Error Types.

AccessDeniedException

You aren't authorized to perform the action. Use the Amazon Resource Name (ARN) of an authorized user or IAM role to perform the operation.

HTTP Status Code: 400

InternalServerError

Amazon Textract experienced a service issue. Try your call again.

HTTP Status Code: 500

InvalidJobIdException

An invalid job identifier was passed to an asynchronous analysis operation.

HTTP Status Code: 400

InvalidKMSKeyException

Indicates you do not have decrypt permissions with the KMS key entered, or the KMS key was entered incorrectly.

HTTP Status Code: 400

InvalidParameterException

An input parameter violated a constraint. For example, in synchronous operations, an `InvalidParameterException` exception occurs when neither of the `S3Object` or `Bytes` values are supplied in the `Document` request parameter. Validate your parameter before calling the API operation again.

HTTP Status Code: 400

InvalidS3ObjectException

Amazon Textract is unable to access the S3 object that's specified in the request. for more information, Configure Access to Amazon S3 For troubleshooting information, see Troubleshooting Amazon S3

HTTP Status Code: 400

ProvisionedThroughputExceededException

The number of requests exceeded your throughput limit. If you want to increase this limit, contact Amazon Textract.

HTTP Status Code: 400

ThrottlingException

Amazon Textract is temporarily unable to process the request. Try your call again.

HTTP Status Code: 500

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface V2
- AWS SDK for .NET V4
- AWS SDK for C++
- AWS SDK for Go v2
- AWS SDK for Java V2
- AWS SDK for JavaScript V3

- AWS SDK for Kotlin
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3

GetLendingAnalysis

Gets the results for an Amazon Textract asynchronous operation that analyzes text in a lending document.

You start asynchronous text analysis by calling `StartLendingAnalysis`, which returns a job identifier (`JobId`). When the text analysis operation finishes, Amazon Textract publishes a completion status to the Amazon Simple Notification Service (Amazon SNS) topic that's registered in the initial call to `StartLendingAnalysis`.

To get the results of the text analysis operation, first check that the status value published to the Amazon SNS topic is SUCCEEDED. If so, call GetLendingAnalysis, and pass the job identifier (`JobId`) from the initial call to `StartLendingAnalysis`.

Request Syntax

```
{
   "JobId": "string",
   "MaxResults": number,
   "NextToken": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters.

The request accepts the following data in JSON format.

JobId

A unique identifier for the lending or text-detection job. The `JobId` is returned from `StartLendingAnalysis`. A `JobId` value is only valid for 7 days.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: `^[a-zA-Z0-9-_]+$`

Required: Yes

MaxResults

The maximum number of results to return per paginated call. The largest value that you can specify is 30. If you specify a value greater than 30, a maximum of 30 results is returned. The default value is 30.

Type: Integer

Valid Range: Minimum value of 1.

Required: No

NextToken

If the previous response was incomplete, Amazon Textract returns a pagination token in the response. You can use this pagination token to retrieve the next set of lending results.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1024.

Pattern: `.*\S.*`

Required: No

Response Syntax

```
{
   "AnalyzeLendingModelVersion": "string",
   "DocumentMetadata": {
      "Pages": number
   },
   "JobStatus": "string",
   "NextToken": "string",
   "Results": [
      {
         "Extractions": [
            {
               "ExpenseDocument": {
                  "Blocks": [
                     {
                        "BlockType": "string",
                        "ColumnIndex": number,
                        "ColumnSpan": number,
```

```
                        "Confidence": number,
                        "EntityTypes": [ "string" ],
                        "Geometry": {
                           "BoundingBox": {
                              "Height": number,
                              "Left": number,
                              "Top": number,
                              "Width": number
                           },
                           "Polygon": [
                              {
                                 "X": number,
                                 "Y": number
                              }
                           ],
                           "RotationAngle": number
                        },
                        "Id": "string",
                        "Page": number,
                        "Query": {
                           "Alias": "string",
                           "Pages": [ "string" ],
                           "Text": "string"
                        },
                        "Relationships": [
                           {
                              "Ids": [ "string" ],
                              "Type": "string"
                           }
                        ],
                        "RowIndex": number,
                        "RowSpan": number,
                        "SelectionStatus": "string",
                        "Text": "string",
                        "TextType": "string"
                     }
                  ],
                  "ExpenseIndex": number,
                  "LineItemGroups": [
                     {
                        "LineItemGroupIndex": number,
                        "LineItems": [
                           {
                              "LineItemExpenseFields": [
```

```
{
   "Currency": {
      "Code": "string",
      "Confidence": number
   },
   "GroupProperties": [
      {
         "Id": "string",
         "Types": [ "string" ]
      }
   ],
   "LabelDetection": {
      "Confidence": number,
      "Geometry": {
         "BoundingBox": {
            "Height": number,
            "Left": number,
            "Top": number,
            "Width": number
         },
         "Polygon": [
            {
               "X": number,
               "Y": number
            }
         ],
         "RotationAngle": number
      },
      "Text": "string"
   },
   "PageNumber": number,
   "Type": {
      "Confidence": number,
      "Text": "string"
   },
   "ValueDetection": {
      "Confidence": number,
      "Geometry": {
         "BoundingBox": {
            "Height": number,
            "Left": number,
            "Top": number,
            "Width": number
         },
```

```
                                    "Polygon": [
                                       {
                                          "X": number,
                                          "Y": number
                                       }
                                    ],
                                    "RotationAngle": number
                                 },
                                 "Text": "string"
                              }
                           }
                        ]
                     }
                  ]
               }
            ],
            "SummaryFields": [
               {
                  "Currency": {
                     "Code": "string",
                     "Confidence": number
                  },
                  "GroupProperties": [
                     {
                        "Id": "string",
                        "Types": [ "string" ]
                     }
                  ],
                  "LabelDetection": {
                     "Confidence": number,
                     "Geometry": {
                        "BoundingBox": {
                           "Height": number,
                           "Left": number,
                           "Top": number,
                           "Width": number
                        },
                        "Polygon": [
                           {
                              "X": number,
                              "Y": number
                           }
                        ],
                        "RotationAngle": number
```

```
                        },
                        "Text": "string"
                     },
                     "PageNumber": number,
                     "Type": {
                        "Confidence": number,
                        "Text": "string"
                     },
                     "ValueDetection": {
                        "Confidence": number,
                        "Geometry": {
                           "BoundingBox": {
                              "Height": number,
                              "Left": number,
                              "Top": number,
                              "Width": number
                           },
                           "Polygon": [
                              {
                                 "X": number,
                                 "Y": number
                              }
                           ],
                           "RotationAngle": number
                        },
                        "Text": "string"
                     }
                  }
               ]
            },
            "IdentityDocument": {
               "Blocks": [
                  {
                     "BlockType": "string",
                     "ColumnIndex": number,
                     "ColumnSpan": number,
                     "Confidence": number,
                     "EntityTypes": [ "string" ],
                     "Geometry": {
                        "BoundingBox": {
                           "Height": number,
                           "Left": number,
                           "Top": number,
                           "Width": number
```

```
                        },
                        "Polygon": [
                           {
                              "X": number,
                              "Y": number
                           }
                        ],
                        "RotationAngle": number
                     },
                     "Id": "string",
                     "Page": number,
                     "Query": {
                        "Alias": "string",
                        "Pages": [ "string" ],
                        "Text": "string"
                     },
                     "Relationships": [
                        {
                           "Ids": [ "string" ],
                           "Type": "string"
                        }
                     ],
                     "RowIndex": number,
                     "RowSpan": number,
                     "SelectionStatus": "string",
                     "Text": "string",
                     "TextType": "string"
                  }
               ],
               "DocumentIndex": number,
               "IdentityDocumentFields": [
                  {
                     "Type": {
                        "Confidence": number,
                        "NormalizedValue": {
                           "Value": "string",
                           "ValueType": "string"
                        },
                        "Text": "string"
                     },
                     "ValueDetection": {
                        "Confidence": number,
                        "NormalizedValue": {
                           "Value": "string",
```

```
                        "ValueType": "string"
                     },
                     "Text": "string"
                  }
               }
            ]
         },
         "LendingDocument": {
            "LendingFields": [
               {
                  "KeyDetection": {
                     "Confidence": number,
                     "Geometry": {
                        "BoundingBox": {
                           "Height": number,
                           "Left": number,
                           "Top": number,
                           "Width": number
                        },
                        "Polygon": [
                           {
                              "X": number,
                              "Y": number
                           }
                        ],
                        "RotationAngle": number
                     },
                     "SelectionStatus": "string",
                     "Text": "string"
                  },
                  "Type": "string",
                  "ValueDetections": [
                     {
                        "Confidence": number,
                        "Geometry": {
                           "BoundingBox": {
                              "Height": number,
                              "Left": number,
                              "Top": number,
                              "Width": number
                           },
                           "Polygon": [
                              {
                                 "X": number,
```

```
                                    "Y": number
                                 }
                              ],
                              "RotationAngle": number
                           },
                           "SelectionStatus": "string",
                           "Text": "string"
                        }
                     ]
                  }
               ],
               "SignatureDetections": [
                  {
                     "Confidence": number,
                     "Geometry": {
                        "BoundingBox": {
                           "Height": number,
                           "Left": number,
                           "Top": number,
                           "Width": number
                        },
                        "Polygon": [
                           {
                              "X": number,
                              "Y": number
                           }
                        ],
                        "RotationAngle": number
                     }
                  }
               ]
            }
         }
      ],
      "Page": number,
      "PageClassification": {
         "PageNumber": [
            {
               "Confidence": number,
               "Value": "string"
            }
         ],
         "PageType": [
            {
```

```
                  "Confidence": number,
                  "Value": "string"
               }
            ]
         }
      }
   ],
   "StatusMessage": "string",
   "Warnings": [
      {
         "ErrorCode": "string",
         "Pages": [ number ]
      }
   ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

AnalyzeLendingModelVersion

The current model version of the Analyze Lending API.

Type: String

DocumentMetadata

Information about the input document.

Type: DocumentMetadata object

JobStatus

The current status of the lending analysis job.

Type: String

Valid Values: `IN_PROGRESS | SUCCEEDED | FAILED | PARTIAL_SUCCESS`

NextToken

If the response is truncated, Amazon Textract returns this token. You can use this token in the subsequent request to retrieve the next set of lending results.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1024.

Pattern: `.*\S.*`

Results

Holds the information returned by one of AmazonTextract's document analysis operations for the pinstripe.

Type: Array of LendingResult objects

StatusMessage

Returns if the lending analysis job could not be completed. Contains explanation for what error occurred.

Type: String

Warnings

A list of warnings that occurred during the lending analysis operation.

Type: Array of Warning objects

Errors

For information about the errors that are common to all actions, see Common Error Types.

AccessDeniedException

You aren't authorized to perform the action. Use the Amazon Resource Name (ARN) of an authorized user or IAM role to perform the operation.

HTTP Status Code: 400

InternalServerError

Amazon Textract experienced a service issue. Try your call again.

HTTP Status Code: 500

InvalidJobIdException

An invalid job identifier was passed to an asynchronous analysis operation.

HTTP Status Code: 400

InvalidKMSKeyException

Indicates you do not have decrypt permissions with the KMS key entered, or the KMS key was entered incorrectly.

HTTP Status Code: 400

InvalidParameterException

An input parameter violated a constraint. For example, in synchronous operations, an `InvalidParameterException` exception occurs when neither of the `S3Object` or `Bytes` values are supplied in the `Document` request parameter. Validate your parameter before calling the API operation again.

HTTP Status Code: 400

InvalidS3ObjectException

Amazon Textract is unable to access the S3 object that's specified in the request. for more information, Configure Access to Amazon S3 For troubleshooting information, see Troubleshooting Amazon S3

HTTP Status Code: 400

ProvisionedThroughputExceededException

The number of requests exceeded your throughput limit. If you want to increase this limit, contact Amazon Textract.

HTTP Status Code: 400

ThrottlingException

Amazon Textract is temporarily unable to process the request. Try your call again.

HTTP Status Code: 500

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface V2
- AWS SDK for .NET V4
- AWS SDK for C++
- AWS SDK for Go v2
- AWS SDK for Java V2
- AWS SDK for JavaScript V3
- AWS SDK for Kotlin
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3

GetLendingAnalysisSummary

Gets summarized results for the `StartLendingAnalysis` operation, which analyzes text in a lending document. The returned summary consists of information about documents grouped together by a common document type. Information like detected signatures, page numbers, and split documents is returned with respect to the type of grouped document.

You start asynchronous text analysis by calling `StartLendingAnalysis`, which returns a job identifier (`JobId`). When the text analysis operation finishes, Amazon Textract publishes a completion status to the Amazon Simple Notification Service (Amazon SNS) topic that's registered in the initial call to `StartLendingAnalysis`.

To get the results of the text analysis operation, first check that the status value published to the Amazon SNS topic is SUCCEEDED. If so, call `GetLendingAnalysisSummary`, and pass the job identifier (`JobId`) from the initial call to `StartLendingAnalysis`.

Request Syntax

```
{
   "JobId": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters.

The request accepts the following data in JSON format.

JobId

A unique identifier for the lending or text-detection job. The `JobId` is returned from StartLendingAnalysis. A `JobId` value is only valid for 7 days.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: `^[a-zA-Z0-9-_]+$`

Required: Yes

Response Syntax

```
{
   "AnalyzeLendingModelVersion": "string",
   "DocumentMetadata": {
      "Pages": number
   },
   "JobStatus": "string",
   "StatusMessage": "string",
   "Summary": {
      "DocumentGroups": [
         {
            "DetectedSignatures": [
               {
                  "Page": number
               }
            ],
            "SplitDocuments": [
               {
                  "Index": number,
                  "Pages": [ number ]
               }
            ],
            "Type": "string",
            "UndetectedSignatures": [
               {
                  "Page": number
               }
            ]
         }
      ],
      "UndetectedDocumentTypes": [ "string" ]
   },
   "Warnings": [
      {
         "ErrorCode": "string",
         "Pages": [ number ]
      }
   ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

AnalyzeLendingModelVersion

The current model version of the Analyze Lending API.

Type: String

DocumentMetadata

Information about the input document.

Type: DocumentMetadata object

JobStatus

The current status of the lending analysis job.

Type: String

Valid Values: `IN_PROGRESS | SUCCEEDED | FAILED | PARTIAL_SUCCESS`

StatusMessage

Returns if the lending analysis could not be completed. Contains explanation for what error occurred.

Type: String

Summary

Contains summary information for documents grouped by type.

Type: LendingSummary object

Warnings

A list of warnings that occurred during the lending analysis operation.

Type: Array of Warning objects

Errors

For information about the errors that are common to all actions, see Common Error Types.

AccessDeniedException

You aren't authorized to perform the action. Use the Amazon Resource Name (ARN) of an authorized user or IAM role to perform the operation.

HTTP Status Code: 400

InternalServerError

Amazon Textract experienced a service issue. Try your call again.

HTTP Status Code: 500

InvalidJobIdException

An invalid job identifier was passed to an asynchronous analysis operation.

HTTP Status Code: 400

InvalidKMSKeyException

Indicates you do not have decrypt permissions with the KMS key entered, or the KMS key was entered incorrectly.

HTTP Status Code: 400

InvalidParameterException

An input parameter violated a constraint. For example, in synchronous operations, an `InvalidParameterException` exception occurs when neither of the `S3Object` or `Bytes` values are supplied in the `Document` request parameter. Validate your parameter before calling the API operation again.

HTTP Status Code: 400

InvalidS3ObjectException

Amazon Textract is unable to access the S3 object that's specified in the request. for more information, Configure Access to Amazon S3 For troubleshooting information, see Troubleshooting Amazon S3

HTTP Status Code: 400

ProvisionedThroughputExceededException

The number of requests exceeded your throughput limit. If you want to increase this limit, contact Amazon Textract.

HTTP Status Code: 400

ThrottlingException

Amazon Textract is temporarily unable to process the request. Try your call again.

HTTP Status Code: 500

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface V2
- AWS SDK for .NET V4
- AWS SDK for C++
- AWS SDK for Go v2
- AWS SDK for Java V2
- AWS SDK for JavaScript V3
- AWS SDK for Kotlin
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3

ListAdapters

Lists all adapters that match the specified filtration criteria.

Request Syntax

```
{
   "AfterCreationTime": number,
   "BeforeCreationTime": number,
   "MaxResults": number,
   "NextToken": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters.

The request accepts the following data in JSON format.

AfterCreationTime

Specifies the lower bound for the ListAdapters operation. Ensures ListAdapters returns only adapters created after the specified creation time.

Type: Timestamp

Required: No

BeforeCreationTime

Specifies the upper bound for the ListAdapters operation. Ensures ListAdapters returns only adapters created before the specified creation time.

Type: Timestamp

Required: No

MaxResults

The maximum number of results to return when listing adapters.

Type: Integer

Valid Range: Minimum value of 1.

Required: No

NextToken

Identifies the next page of results to return when listing adapters.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1024.

Pattern: `.*\S.*`

Required: No

Response Syntax

```
{
   "Adapters": [
      {
         "AdapterId": "string",
         "AdapterName": "string",
         "CreationTime": number,
         "FeatureTypes": [ "string" ]
      }
   ],
   "NextToken": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Adapters

A list of adapters that matches the filtering criteria specified when calling ListAdapters.

Type: Array of AdapterOverview objects

NextToken

Identifies the next page of results to return when listing adapters.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1024.

Pattern: `.*\S.*`

Errors

For information about the errors that are common to all actions, see Common Error Types.

AccessDeniedException

You aren't authorized to perform the action. Use the Amazon Resource Name (ARN) of an authorized user or IAM role to perform the operation.

HTTP Status Code: 400

InternalServerError

Amazon Textract experienced a service issue. Try your call again.

HTTP Status Code: 500

InvalidParameterException

An input parameter violated a constraint. For example, in synchronous operations, an `InvalidParameterException` exception occurs when neither of the `S3Object` or `Bytes` values are supplied in the `Document` request parameter. Validate your parameter before calling the API operation again.

HTTP Status Code: 400

ProvisionedThroughputExceededException

The number of requests exceeded your throughput limit. If you want to increase this limit, contact Amazon Textract.

HTTP Status Code: 400

ThrottlingException

Amazon Textract is temporarily unable to process the request. Try your call again.

HTTP Status Code: 500

ValidationException

Indicates that a request was not valid. Check request for proper formatting.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface V2
- AWS SDK for .NET V4
- AWS SDK for C++
- AWS SDK for Go v2
- AWS SDK for Java V2
- AWS SDK for JavaScript V3
- AWS SDK for Kotlin
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3

ListAdapterVersions

List all version of an adapter that meet the specified filtration criteria.

Request Syntax

```
{
   "AdapterId": "string",
   "AfterCreationTime": number,
   "BeforeCreationTime": number,
   "MaxResults": number,
   "NextToken": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters.

The request accepts the following data in JSON format.

AdapterId

A string containing a unique ID for the adapter to match for when listing adapter versions.

Type: String

Length Constraints: Minimum length of 12. Maximum length of 1011.

Required: No

AfterCreationTime

Specifies the lower bound for the ListAdapterVersions operation. Ensures ListAdapterVersions returns only adapter versions created after the specified creation time.

Type: Timestamp

Required: No

BeforeCreationTime

Specifies the upper bound for the ListAdapterVersions operation. Ensures ListAdapterVersions returns only adapter versions created after the specified creation time.

Type: Timestamp

Required: No

MaxResults

The maximum number of results to return when listing adapter versions.

Type: Integer

Valid Range: Minimum value of 1.

Required: No

NextToken

Identifies the next page of results to return when listing adapter versions.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1024.

Pattern: `.*\S.*`

Required: No

Response Syntax

```
{
   "AdapterVersions": [
      {
         "AdapterId": "string",
         "AdapterVersion": "string",
         "CreationTime": number,
         "FeatureTypes": [ "string" ],
         "Status": "string",
         "StatusMessage": "string"
      }
   ],
   "NextToken": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

AdapterVersions

Adapter versions that match the filtering criteria specified when calling ListAdapters.

Type: Array of AdapterVersionOverview objects

NextToken

Identifies the next page of results to return when listing adapter versions.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1024.

Pattern: `.*\S.*`

Errors

For information about the errors that are common to all actions, see Common Error Types.

AccessDeniedException

You aren't authorized to perform the action. Use the Amazon Resource Name (ARN) of an authorized user or IAM role to perform the operation.

HTTP Status Code: 400

InternalServerError

Amazon Textract experienced a service issue. Try your call again.

HTTP Status Code: 500

InvalidParameterException

An input parameter violated a constraint. For example, in synchronous operations, an `InvalidParameterException` exception occurs when neither of the `S3Object` or `Bytes` values are supplied in the `Document` request parameter. Validate your parameter before calling the API operation again.

HTTP Status Code: 400

ProvisionedThroughputExceededException

The number of requests exceeded your throughput limit. If you want to increase this limit, contact Amazon Textract.

HTTP Status Code: 400

ResourceNotFoundException

Returned when an operation tried to access a nonexistent resource.

HTTP Status Code: 400

ThrottlingException

Amazon Textract is temporarily unable to process the request. Try your call again.

HTTP Status Code: 500

ValidationException

Indicates that a request was not valid. Check request for proper formatting.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface V2
- AWS SDK for .NET V4
- AWS SDK for C++
- AWS SDK for Go v2
- AWS SDK for Java V2
- AWS SDK for JavaScript V3
- AWS SDK for Kotlin
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3

ListTagsForResource

Lists all tags for an Amazon Textract resource.

Request Syntax

```
{
   "ResourceARN": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters.

The request accepts the following data in JSON format.

ResourceARN

The Amazon Resource Name (ARN) that specifies the resource to list tags for.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1011.

Required: Yes

Response Syntax

```
{
   "Tags": {
      "string" : "string"
   }
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

Tags

A set of tags (key-value pairs) that are part of the requested resource.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 200 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Key Pattern: `^(?!aws:)[\p{L}\p{Z}\p{N}_.:/=+\-@]*$`

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Value Pattern: `^([\p{L}\p{Z}\p{N}_.:/=+\-@]*)$`

Errors

For information about the errors that are common to all actions, see Common Error Types.

AccessDeniedException

You aren't authorized to perform the action. Use the Amazon Resource Name (ARN) of an authorized user or IAM role to perform the operation.

HTTP Status Code: 400

InternalServerError

Amazon Textract experienced a service issue. Try your call again.

HTTP Status Code: 500

InvalidParameterException

An input parameter violated a constraint. For example, in synchronous operations, an `InvalidParameterException` exception occurs when neither of the `S3Object` or `Bytes` values are supplied in the `Document` request parameter. Validate your parameter before calling the API operation again.

HTTP Status Code: 400

ProvisionedThroughputExceededException

The number of requests exceeded your throughput limit. If you want to increase this limit, contact Amazon Textract.

HTTP Status Code: 400

ResourceNotFoundException

Returned when an operation tried to access a nonexistent resource.

HTTP Status Code: 400

ThrottlingException

Amazon Textract is temporarily unable to process the request. Try your call again.

HTTP Status Code: 500

ValidationException

Indicates that a request was not valid. Check request for proper formatting.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface V2
- AWS SDK for .NET V4
- AWS SDK for C++
- AWS SDK for Go v2
- AWS SDK for Java V2
- AWS SDK for JavaScript V3
- AWS SDK for Kotlin
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3

StartDocumentAnalysis

Starts the asynchronous analysis of an input document for relationships between detected items such as key-value pairs, tables, and selection elements.

`StartDocumentAnalysis` can analyze text in documents that are in JPEG, PNG, TIFF, and PDF format. The documents are stored in an Amazon S3 bucket. Use DocumentLocation to specify the bucket name and file name of the document.

`StartDocumentAnalysis` returns a job identifier (`JobId`) that you use to get the results of the operation. When text analysis is finished, Amazon Textract publishes a completion status to the Amazon Simple Notification Service (Amazon SNS) topic that you specify in `NotificationChannel`. To get the results of the text analysis operation, first check that the status value published to the Amazon SNS topic is SUCCEEDED. If so, call GetDocumentAnalysis, and pass the job identifier (`JobId`) from the initial call to `StartDocumentAnalysis`.

For more information, see Document Text Analysis.

Request Syntax

```
{
   "AdaptersConfig": {
      "Adapters": [
         {
            "AdapterId": "string",
            "Pages": [ "string" ],
            "Version": "string"
         }
      ]
   },
   "ClientRequestToken": "string",
   "DocumentLocation": {
      "S3Object": {
         "Bucket": "string",
         "Name": "string",
         "Version": "string"
      }
   },
   "FeatureTypes": [ "string" ],
   "JobTag": "string",
   "KMSKeyId": "string",
```

```
   "NotificationChannel": {
      "RoleArn": "string",
      "SNSTopicArn": "string"
   },
   "OutputConfig": {
      "S3Bucket": "string",
      "S3Prefix": "string"
   },
   "QueriesConfig": {
      "Queries": [
         {
            "Alias": "string",
            "Pages": [ "string" ],
            "Text": "string"
         }
      ]
   }
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters.

The request accepts the following data in JSON format.

AdaptersConfig

Specifies the adapter to be used when analyzing a document.

Type: AdaptersConfig object

Required: No

ClientRequestToken

The idempotent token that you use to identify the start request. If you use the same token with multiple `StartDocumentAnalysis` requests, the same `JobId` is returned. Use `ClientRequestToken` to prevent the same job from being accidentally started more than once. For more information, see Calling Amazon Textract Asynchronous Operations.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: ^[a-zA-Z0-9-_]+$

Required: No

DocumentLocation

The location of the document to be processed.

Type: DocumentLocation object

Required: Yes

FeatureTypes

A list of the types of analysis to perform. Add TABLES to the list to return information about the tables that are detected in the input document. Add FORMS to return detected form data. To perform both types of analysis, add TABLES and FORMS to FeatureTypes. All lines and words detected in the document are included in the response (including text that isn't related to the value of FeatureTypes).

Type: Array of strings

Valid Values: TABLES | FORMS | QUERIES | SIGNATURES | LAYOUT

Required: Yes

JobTag

An identifier that you specify that's included in the completion notification published to the Amazon SNS topic. For example, you can use JobTag to identify the type of document that the completion notification corresponds to (such as a tax form or a receipt).

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: [a-zA-Z0-9_.\-:]+

Required: No

KMSKeyId

The KMS key used to encrypt the inference results. This can be in either Key ID or Key Alias format. When a KMS key is provided, the KMS key will be used for server-side encryption of the objects in the customer bucket. When this parameter is not enabled, the result will be encrypted server side,using SSE-S3.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `^[A-Za-z0-9][A-Za-z0-9:_/+=,@.-]{0,2048}$`

Required: No

NotificationChannel

The Amazon SNS topic ARN that you want Amazon Textract to publish the completion status of the operation to.

Type: NotificationChannel object

Required: No

OutputConfig

Sets if the output will go to a customer defined bucket. By default, Amazon Textract will save the results internally to be accessed by the GetDocumentAnalysis operation.

Type: OutputConfig object

Required: No

QueriesConfig

Type: QueriesConfig object

Required: No

Response Syntax

```
{
   "JobId": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

JobId

The identifier for the document text detection job. Use `JobId` to identify the job in a subsequent call to `GetDocumentAnalysis`. A `JobId` value is only valid for 7 days.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: `^[a-zA-Z0-9-_]+$`

Errors

For information about the errors that are common to all actions, see Common Error Types.

AccessDeniedException

You aren't authorized to perform the action. Use the Amazon Resource Name (ARN) of an authorized user or IAM role to perform the operation.

HTTP Status Code: 400

BadDocumentException

Amazon Textract isn't able to read the document. For more information on the document limits in Amazon Textract, see Hard limits.

HTTP Status Code: 400

DocumentTooLargeException

The document can't be processed because it's too large. The maximum document size for synchronous operations 10 MB. The maximum document size for asynchronous operations is 500 MB for PDF files.

HTTP Status Code: 400

IdempotentParameterMismatchException

A `ClientRequestToken` input parameter was reused with an operation, but at least one of the other input parameters is different from the previous call to the operation.

HTTP Status Code: 400

InternalServerError

Amazon Textract experienced a service issue. Try your call again.

HTTP Status Code: 500

InvalidKMSKeyException

Indicates you do not have decrypt permissions with the KMS key entered, or the KMS key was entered incorrectly.

HTTP Status Code: 400

InvalidParameterException

An input parameter violated a constraint. For example, in synchronous operations, an `InvalidParameterException` exception occurs when neither of the `S3Object` or `Bytes` values are supplied in the `Document` request parameter. Validate your parameter before calling the API operation again.

HTTP Status Code: 400

InvalidS3ObjectException

Amazon Textract is unable to access the S3 object that's specified in the request. for more information, Configure Access to Amazon S3 For troubleshooting information, see Troubleshooting Amazon S3

HTTP Status Code: 400

LimitExceededException

An Amazon Textract service limit was exceeded. For example, if you start too many asynchronous jobs concurrently, calls to start operations (`StartDocumentTextDetection`, for example) raise a LimitExceededException exception (HTTP status code: 400) until the number of concurrently running jobs is below the Amazon Textract service limit.

HTTP Status Code: 400

ProvisionedThroughputExceededException

The number of requests exceeded your throughput limit. If you want to increase this limit, contact Amazon Textract.

HTTP Status Code: 400

ThrottlingException

Amazon Textract is temporarily unable to process the request. Try your call again.

HTTP Status Code: 500

UnsupportedDocumentException

The format of the input document isn't supported. Documents for operations can be in PNG, JPEG, PDF, or TIFF format.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface V2
- AWS SDK for .NET V4
- AWS SDK for C++
- AWS SDK for Go v2
- AWS SDK for Java V2
- AWS SDK for JavaScript V3
- AWS SDK for Kotlin
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3

StartDocumentTextDetection

Starts the asynchronous detection of text in a document. Amazon Textract can detect lines of text and the words that make up a line of text.

`StartDocumentTextDetection` can analyze text in documents that are in JPEG, PNG, TIFF, and PDF format. The documents are stored in an Amazon S3 bucket. Use DocumentLocation to specify the bucket name and file name of the document.

`StartDocumentTextDetection` returns a job identifier (`JobId`) that you use to get the results of the operation. When text detection is finished, Amazon Textract publishes a completion status to the Amazon Simple Notification Service (Amazon SNS) topic that you specify in `NotificationChannel`. To get the results of the text detection operation, first check that the status value published to the Amazon SNS topic is SUCCEEDED. If so, call GetDocumentTextDetection, and pass the job identifier (`JobId`) from the initial call to `StartDocumentTextDetection`.

For more information, see Document Text Detection.

Request Syntax

```
{
   "ClientRequestToken": "string",
   "DocumentLocation": {
      "S3Object": {
         "Bucket": "string",
         "Name": "string",
         "Version": "string"
      }
   },
   "JobTag": "string",
   "KMSKeyId": "string",
   "NotificationChannel": {
      "RoleArn": "string",
      "SNSTopicArn": "string"
   },
   "OutputConfig": {
      "S3Bucket": "string",
      "S3Prefix": "string"
   }
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters.

The request accepts the following data in JSON format.

ClientRequestToken

The idempotent token that's used to identify the start request. If you use the same token with multiple `StartDocumentTextDetection` requests, the same `JobId` is returned. Use `ClientRequestToken` to prevent the same job from being accidentally started more than once. For more information, see Calling Amazon Textract Asynchronous Operations.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: `^[a-zA-Z0-9-_]+$`

Required: No

DocumentLocation

The location of the document to be processed.

Type: DocumentLocation object

Required: Yes

JobTag

An identifier that you specify that's included in the completion notification published to the Amazon SNS topic. For example, you can use `JobTag` to identify the type of document that the completion notification corresponds to (such as a tax form or a receipt).

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: `[a-zA-Z0-9_.\-:]+`

Required: No

KMSKeyId

The KMS key used to encrypt the inference results. This can be in either Key ID or Key Alias format. When a KMS key is provided, the KMS key will be used for server-side encryption of the objects in the customer bucket. When this parameter is not enabled, the result will be encrypted server side,using SSE-S3.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `^[A-Za-z0-9][A-Za-z0-9:_/+=,@.-]{0,2048}$`

Required: No

NotificationChannel

The Amazon SNS topic ARN that you want Amazon Textract to publish the completion status of the operation to.

Type: NotificationChannel object

Required: No

OutputConfig

Sets if the output will go to a customer defined bucket. By default Amazon Textract will save the results internally to be accessed with the GetDocumentTextDetection operation.

Type: OutputConfig object

Required: No

Response Syntax

```
{
   "JobId": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

JobId

The identifier of the text detection job for the document. Use `JobId` to identify the job in a subsequent call to `GetDocumentTextDetection`. A `JobId` value is only valid for 7 days.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: `^[a-zA-Z0-9-_]+$`

Errors

For information about the errors that are common to all actions, see Common Error Types.

AccessDeniedException

You aren't authorized to perform the action. Use the Amazon Resource Name (ARN) of an authorized user or IAM role to perform the operation.

HTTP Status Code: 400

BadDocumentException

Amazon Textract isn't able to read the document. For more information on the document limits in Amazon Textract, see Hard limits.

HTTP Status Code: 400

DocumentTooLargeException

The document can't be processed because it's too large. The maximum document size for synchronous operations 10 MB. The maximum document size for asynchronous operations is 500 MB for PDF files.

HTTP Status Code: 400

IdempotentParameterMismatchException

A `ClientRequestToken` input parameter was reused with an operation, but at least one of the other input parameters is different from the previous call to the operation.

HTTP Status Code: 400

InternalServerError

Amazon Textract experienced a service issue. Try your call again.

HTTP Status Code: 500

InvalidKMSKeyException

Indicates you do not have decrypt permissions with the KMS key entered, or the KMS key was entered incorrectly.

HTTP Status Code: 400

InvalidParameterException

An input parameter violated a constraint. For example, in synchronous operations, an `InvalidParameterException` exception occurs when neither of the `S3Object` or `Bytes` values are supplied in the `Document` request parameter. Validate your parameter before calling the API operation again.

HTTP Status Code: 400

InvalidS3ObjectException

Amazon Textract is unable to access the S3 object that's specified in the request. for more information, Configure Access to Amazon S3 For troubleshooting information, see Troubleshooting Amazon S3

HTTP Status Code: 400

LimitExceededException

An Amazon Textract service limit was exceeded. For example, if you start too many asynchronous jobs concurrently, calls to start operations (`StartDocumentTextDetection`, for example) raise a LimitExceededException exception (HTTP status code: 400) until the number of concurrently running jobs is below the Amazon Textract service limit.

HTTP Status Code: 400

ProvisionedThroughputExceededException

The number of requests exceeded your throughput limit. If you want to increase this limit, contact Amazon Textract.

HTTP Status Code: 400

ThrottlingException

Amazon Textract is temporarily unable to process the request. Try your call again.

HTTP Status Code: 500

UnsupportedDocumentException

The format of the input document isn't supported. Documents for operations can be in PNG, JPEG, PDF, or TIFF format.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface V2
- AWS SDK for .NET V4
- AWS SDK for C++
- AWS SDK for Go v2
- AWS SDK for Java V2
- AWS SDK for JavaScript V3
- AWS SDK for Kotlin
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3

StartExpenseAnalysis

Starts the asynchronous analysis of invoices or receipts for data like contact information, items purchased, and vendor names.

`StartExpenseAnalysis` can analyze text in documents that are in JPEG, PNG, and PDF format. The documents must be stored in an Amazon S3 bucket. Use the DocumentLocation parameter to specify the name of your S3 bucket and the name of the document in that bucket.

`StartExpenseAnalysis` returns a job identifier (`JobId`) that you will provide to `GetExpenseAnalysis` to retrieve the results of the operation. When the analysis of the input invoices/receipts is finished, Amazon Textract publishes a completion status to the Amazon Simple Notification Service (Amazon SNS) topic that you provide to the `NotificationChannel`. To obtain the results of the invoice and receipt analysis operation, ensure that the status value published to the Amazon SNS topic is SUCCEEDED. If so, call GetExpenseAnalysis, and pass the job identifier (`JobId`) that was returned by your call to `StartExpenseAnalysis`.

For more information, see Analyzing Invoices and Receipts.

Request Syntax

```
{
   "ClientRequestToken": "string",
   "DocumentLocation": {
      "S3Object": {
         "Bucket": "string",
         "Name": "string",
         "Version": "string"
      }
   },
   "JobTag": "string",
   "KMSKeyId": "string",
   "NotificationChannel": {
      "RoleArn": "string",
      "SNSTopicArn": "string"
   },
   "OutputConfig": {
      "S3Bucket": "string",
      "S3Prefix": "string"
   }
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters.

The request accepts the following data in JSON format.

ClientRequestToken

The idempotent token that's used to identify the start request. If you use the same token with multiple `StartDocumentTextDetection` requests, the same `JobId` is returned. Use `ClientRequestToken` to prevent the same job from being accidentally started more than once. For more information, see Calling Amazon Textract Asynchronous Operations

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: `^[a-zA-Z0-9-_]+$`

Required: No

DocumentLocation

The location of the document to be processed.

Type: DocumentLocation object

Required: Yes

JobTag

An identifier you specify that's included in the completion notification published to the Amazon SNS topic. For example, you can use `JobTag` to identify the type of document that the completion notification corresponds to (such as a tax form or a receipt).

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: `[a-zA-Z0-9_.\-:]+`

Required: No

KMSKeyId

The KMS key used to encrypt the inference results. This can be in either Key ID or Key Alias format. When a KMS key is provided, the KMS key will be used for server-side encryption of the objects in the customer bucket. When this parameter is not enabled, the result will be encrypted server side,using SSE-S3.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `^[A-Za-z0-9][A-Za-z0-9:_/+=,@.-]{0,2048}$`

Required: No

NotificationChannel

The Amazon SNS topic ARN that you want Amazon Textract to publish the completion status of the operation to.

Type: NotificationChannel object

Required: No

OutputConfig

Sets if the output will go to a customer defined bucket. By default, Amazon Textract will save the results internally to be accessed by the `GetExpenseAnalysis` operation.

Type: OutputConfig object

Required: No

Response Syntax

```
{
   "JobId": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

JobId

A unique identifier for the text detection job. The `JobId` is returned from `StartExpenseAnalysis`. A `JobId` value is only valid for 7 days.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: `^[a-zA-Z0-9-_]+$`

Errors

For information about the errors that are common to all actions, see Common Error Types.

AccessDeniedException

You aren't authorized to perform the action. Use the Amazon Resource Name (ARN) of an authorized user or IAM role to perform the operation.

HTTP Status Code: 400

BadDocumentException

Amazon Textract isn't able to read the document. For more information on the document limits in Amazon Textract, see Hard limits.

HTTP Status Code: 400

DocumentTooLargeException

The document can't be processed because it's too large. The maximum document size for synchronous operations 10 MB. The maximum document size for asynchronous operations is 500 MB for PDF files.

HTTP Status Code: 400

IdempotentParameterMismatchException

A `ClientRequestToken` input parameter was reused with an operation, but at least one of the other input parameters is different from the previous call to the operation.

HTTP Status Code: 400

InternalServerError

Amazon Textract experienced a service issue. Try your call again.

HTTP Status Code: 500

InvalidKMSKeyException

Indicates you do not have decrypt permissions with the KMS key entered, or the KMS key was entered incorrectly.

HTTP Status Code: 400

InvalidParameterException

An input parameter violated a constraint. For example, in synchronous operations, an `InvalidParameterException` exception occurs when neither of the `S3Object` or `Bytes` values are supplied in the `Document` request parameter. Validate your parameter before calling the API operation again.

HTTP Status Code: 400

InvalidS3ObjectException

Amazon Textract is unable to access the S3 object that's specified in the request. for more information, Configure Access to Amazon S3 For troubleshooting information, see Troubleshooting Amazon S3

HTTP Status Code: 400

LimitExceededException

An Amazon Textract service limit was exceeded. For example, if you start too many asynchronous jobs concurrently, calls to start operations (`StartDocumentTextDetection`, for example) raise a LimitExceededException exception (HTTP status code: 400) until the number of concurrently running jobs is below the Amazon Textract service limit.

HTTP Status Code: 400

ProvisionedThroughputExceededException

The number of requests exceeded your throughput limit. If you want to increase this limit, contact Amazon Textract.

HTTP Status Code: 400

ThrottlingException

Amazon Textract is temporarily unable to process the request. Try your call again.

HTTP Status Code: 500

UnsupportedDocumentException

The format of the input document isn't supported. Documents for operations can be in PNG, JPEG, PDF, or TIFF format.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface V2
- AWS SDK for .NET V4
- AWS SDK for C++
- AWS SDK for Go v2
- AWS SDK for Java V2
- AWS SDK for JavaScript V3
- AWS SDK for Kotlin
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3

StartLendingAnalysis

Starts the classification and analysis of an input document. `StartLendingAnalysis` initiates the classification and analysis of a packet of lending documents. `StartLendingAnalysis` operates on a document file located in an Amazon S3 bucket.

`StartLendingAnalysis` can analyze text in documents that are in one of the following formats: JPEG, PNG, TIFF, PDF. Use `DocumentLocation` to specify the bucket name and the file name of the document.

`StartLendingAnalysis` returns a job identifier (`JobId`) that you use to get the results of the operation. When the text analysis is finished, Amazon Textract publishes a completion status to the Amazon Simple Notification Service (Amazon SNS) topic that you specify in `NotificationChannel`. To get the results of the text analysis operation, first check that the status value published to the Amazon SNS topic is SUCCEEDED. If the status is SUCCEEDED you can call either `GetLendingAnalysis` or `GetLendingAnalysisSummary` and provide the `JobId` to obtain the results of the analysis.

If using `OutputConfig` to specify an Amazon S3 bucket, the output will be contained within the specified prefix in a directory labeled with the job-id. In the directory there are 3 sub-directories:

- detailedResponse (contains the GetLendingAnalysis response)
- summaryResponse (for the GetLendingAnalysisSummary response)
- splitDocuments (documents split across logical boundaries)

Request Syntax

```
{
   "ClientRequestToken": "string",
   "DocumentLocation": {
      "S3Object": {
         "Bucket": "string",
         "Name": "string",
         "Version": "string"
      }
   },
   "JobTag": "string",
   "KMSKeyId": "string",
```

```
    "NotificationChannel": {
       "RoleArn": "string",
       "SNSTopicArn": "string"
    },
    "OutputConfig": {
       "S3Bucket": "string",
       "S3Prefix": "string"
    }
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters.

The request accepts the following data in JSON format.

ClientRequestToken

The idempotent token that you use to identify the start request. If you use the same token with multiple `StartLendingAnalysis` requests, the same `JobId` is returned. Use `ClientRequestToken` to prevent the same job from being accidentally started more than once. For more information, see Calling Amazon Textract Asynchronous Operations.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: `^[a-zA-Z0-9-_]+$`

Required: No

DocumentLocation

The Amazon S3 bucket that contains the document to be processed. It's used by asynchronous operations.

The input document can be an image file in JPEG or PNG format. It can also be a file in PDF format.

Type: DocumentLocation object

Required: Yes

JobTag

An identifier that you specify to be included in the completion notification published to the Amazon SNS topic. For example, you can use `JobTag` to identify the type of document that the completion notification corresponds to (such as a tax form or a receipt).

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: `[a-zA-Z0-9_.\-:]+`

Required: No

KMSKeyId

The KMS key used to encrypt the inference results. This can be in either Key ID or Key Alias format. When a KMS key is provided, the KMS key will be used for server-side encryption of the objects in the customer bucket. When this parameter is not enabled, the result will be encrypted server side, using SSE-S3.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2048.

Pattern: `^[A-Za-z0-9][A-Za-z0-9:_/+=,@.-]{0,2048}$`

Required: No

NotificationChannel

The Amazon Simple Notification Service (Amazon SNS) topic to which Amazon Textract publishes the completion status of an asynchronous document operation.

Type: NotificationChannel object

Required: No

OutputConfig

Sets whether or not your output will go to a user created bucket. Used to set the name of the bucket, and the prefix on the output file.

`OutputConfig` is an optional parameter which lets you adjust where your output will be placed. By default, Amazon Textract will store the results internally and can only be accessed

by the Get API operations. With `OutputConfig` enabled, you can set the name of the bucket the output will be sent to the file prefix of the results where you can download your results. Additionally, you can set the `KMSKeyID` parameter to a customer master key (CMK) to encrypt your output. Without this parameter set Amazon Textract will encrypt server-side using the AWS managed CMK for Amazon S3.

Decryption of Customer Content is necessary for processing of the documents by Amazon Textract. If your account is opted out under an AI services opt out policy then all unencrypted Customer Content is immediately and permanently deleted after the Customer Content has been processed by the service. No copy of of the output is retained by Amazon Textract. For information about how to opt out, see Managing AI services opt-out policy.

For more information on data privacy, see the Data Privacy FAQ.

Type: OutputConfig object

Required: No

Response Syntax

```
{
   "JobId": "string"
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

JobId

A unique identifier for the lending or text-detection job. The `JobId` is returned from `StartLendingAnalysis`. A `JobId` value is only valid for 7 days.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: `^[a-zA-Z0-9-_]+$`

Errors

For information about the errors that are common to all actions, see Common Error Types.

AccessDeniedException

You aren't authorized to perform the action. Use the Amazon Resource Name (ARN) of an authorized user or IAM role to perform the operation.

HTTP Status Code: 400

BadDocumentException

Amazon Textract isn't able to read the document. For more information on the document limits in Amazon Textract, see Hard limits.

HTTP Status Code: 400

DocumentTooLargeException

The document can't be processed because it's too large. The maximum document size for synchronous operations 10 MB. The maximum document size for asynchronous operations is 500 MB for PDF files.

HTTP Status Code: 400

IdempotentParameterMismatchException

A `ClientRequestToken` input parameter was reused with an operation, but at least one of the other input parameters is different from the previous call to the operation.

HTTP Status Code: 400

InternalServerError

Amazon Textract experienced a service issue. Try your call again.

HTTP Status Code: 500

InvalidKMSKeyException

Indicates you do not have decrypt permissions with the KMS key entered, or the KMS key was entered incorrectly.

HTTP Status Code: 400

InvalidParameterException

An input parameter violated a constraint. For example, in synchronous operations, an `InvalidParameterException` exception occurs when neither of the `S3Object` or `Bytes` values are supplied in the `Document` request parameter. Validate your parameter before calling the API operation again.

HTTP Status Code: 400

InvalidS3ObjectException

Amazon Textract is unable to access the S3 object that's specified in the request. for more information, Configure Access to Amazon S3 For troubleshooting information, see Troubleshooting Amazon S3

HTTP Status Code: 400

LimitExceededException

An Amazon Textract service limit was exceeded. For example, if you start too many asynchronous jobs concurrently, calls to start operations (`StartDocumentTextDetection`, for example) raise a LimitExceededException exception (HTTP status code: 400) until the number of concurrently running jobs is below the Amazon Textract service limit.

HTTP Status Code: 400

ProvisionedThroughputExceededException

The number of requests exceeded your throughput limit. If you want to increase this limit, contact Amazon Textract.

HTTP Status Code: 400

ThrottlingException

Amazon Textract is temporarily unable to process the request. Try your call again.

HTTP Status Code: 500

UnsupportedDocumentException

The format of the input document isn't supported. Documents for operations can be in PNG, JPEG, PDF, or TIFF format.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface V2
- AWS SDK for .NET V4
- AWS SDK for C++
- AWS SDK for Go v2
- AWS SDK for Java V2
- AWS SDK for JavaScript V3
- AWS SDK for Kotlin
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3

TagResource

Adds one or more tags to the specified resource.

Request Syntax

```
{
   "ResourceARN": "string",
   "Tags": {
      "string" : "string"
   }
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters.

The request accepts the following data in JSON format.

ResourceARN

The Amazon Resource Name (ARN) that specifies the resource to be tagged.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1011.

Required: Yes

Tags

A set of tags (key-value pairs) that you want to assign to the resource.

Type: String to string map

Map Entries: Minimum number of 0 items. Maximum number of 200 items.

Key Length Constraints: Minimum length of 1. Maximum length of 128.

Key Pattern: `^(?!aws:)[\p{L}\p{Z}\p{N}_.:/=+\-@]*$`

Value Length Constraints: Minimum length of 0. Maximum length of 256.

Value Pattern: `^([\p{L}\p{Z}\p{N}_.:/=+\-@]*)$`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Error Types.

AccessDeniedException

You aren't authorized to perform the action. Use the Amazon Resource Name (ARN) of an authorized user or IAM role to perform the operation.

HTTP Status Code: 400

InternalServerError

Amazon Textract experienced a service issue. Try your call again.

HTTP Status Code: 500

InvalidParameterException

An input parameter violated a constraint. For example, in synchronous operations, an `InvalidParameterException` exception occurs when neither of the `S3Object` or `Bytes` values are supplied in the `Document` request parameter. Validate your parameter before calling the API operation again.

HTTP Status Code: 400

ProvisionedThroughputExceededException

The number of requests exceeded your throughput limit. If you want to increase this limit, contact Amazon Textract.

HTTP Status Code: 400

ResourceNotFoundException

Returned when an operation tried to access a nonexistent resource.

HTTP Status Code: 400

ServiceQuotaExceededException

Returned when a request cannot be completed as it would exceed a maximum service quota.

HTTP Status Code: 400

ThrottlingException

Amazon Textract is temporarily unable to process the request. Try your call again.

HTTP Status Code: 500

ValidationException

Indicates that a request was not valid. Check request for proper formatting.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface V2
- AWS SDK for .NET V4
- AWS SDK for C++
- AWS SDK for Go v2
- AWS SDK for Java V2
- AWS SDK for JavaScript V3
- AWS SDK for Kotlin
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3

UntagResource

Removes any tags with the specified keys from the specified resource.

Request Syntax

```
{
   "ResourceARN": "string",
   "TagKeys": [ "string" ]
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters.

The request accepts the following data in JSON format.

ResourceARN

The Amazon Resource Name (ARN) that specifies the resource to be untagged.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1011.

Required: Yes

TagKeys

Specifies the tags to be removed from the resource specified by the ResourceARN.

Type: Array of strings

Array Members: Minimum number of 0 items. Maximum number of 200 items.

Length Constraints: Minimum length of 1. Maximum length of 128.

Pattern: `^(?!aws:)[\p{L}\p{Z}\p{N}_.:/=+\-@]*$`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Error Types.

AccessDeniedException

You aren't authorized to perform the action. Use the Amazon Resource Name (ARN) of an authorized user or IAM role to perform the operation.

HTTP Status Code: 400

InternalServerError

Amazon Textract experienced a service issue. Try your call again.

HTTP Status Code: 500

InvalidParameterException

An input parameter violated a constraint. For example, in synchronous operations, an `InvalidParameterException` exception occurs when neither of the `S3Object` or `Bytes` values are supplied in the `Document` request parameter. Validate your parameter before calling the API operation again.

HTTP Status Code: 400

ProvisionedThroughputExceededException

The number of requests exceeded your throughput limit. If you want to increase this limit, contact Amazon Textract.

HTTP Status Code: 400

ResourceNotFoundException

Returned when an operation tried to access a nonexistent resource.

HTTP Status Code: 400

ThrottlingException

Amazon Textract is temporarily unable to process the request. Try your call again.

HTTP Status Code: 500

ValidationException

Indicates that a request was not valid. Check request for proper formatting.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface V2
- AWS SDK for .NET V4
- AWS SDK for C++
- AWS SDK for Go v2
- AWS SDK for Java V2
- AWS SDK for JavaScript V3
- AWS SDK for Kotlin
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3

UpdateAdapter

Update the configuration for an adapter. FeatureTypes configurations cannot be updated. At least one new parameter must be specified as an argument.

Request Syntax

```
{
   "AdapterId": "string",
   "AdapterName": "string",
   "AutoUpdate": "string",
   "Description": "string"
}
```

Request Parameters

For information about the parameters that are common to all actions, see Common Parameters.

The request accepts the following data in JSON format.

AdapterId

A string containing a unique ID for the adapter that will be updated.

Type: String

Length Constraints: Minimum length of 12. Maximum length of 1011.

Required: Yes

AdapterName

The new name to be applied to the adapter.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Pattern: `[a-zA-Z0-9-_]+`

Required: No

AutoUpdate

The new auto-update status to be applied to the adapter.

Type: String

Valid Values: `ENABLED | DISABLED`

Required: No

Description

The new description to be applied to the adapter.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: `^[a-zA-Z0-9\s!"\#\$%'&\(\)\*\+\,\-\./:;=\?@\[\\\]\^_`\{\|\}~><]+$`

Required: No

Response Syntax

```
{
   "AdapterId": "string",
   "AdapterName": "string",
   "AutoUpdate": "string",
   "CreationTime": number,
   "Description": "string",
   "FeatureTypes": [ "string" ]
}
```

Response Elements

If the action is successful, the service sends back an HTTP 200 response.

The following data is returned in JSON format by the service.

AdapterId

A string containing a unique ID for the adapter that has been updated.

Type: String

Length Constraints: Minimum length of 12. Maximum length of 1011.

AdapterName

A string containing the name of the adapter that has been updated.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Pattern: `[a-zA-Z0-9-_]+`

AutoUpdate

The auto-update status of the adapter that has been updated.

Type: String

Valid Values: `ENABLED | DISABLED`

CreationTime

An object specifying the creation time of the the adapter that has been updated.

Type: Timestamp

Description

A string containing the description of the adapter that has been updated.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: `^[a-zA-Z0-9\s!"\#\$%'&\(\)\*\+\,\-\./:;=\?@\[\\\]\^_`\{\|\}~><]+$`

FeatureTypes

List of the targeted feature types for the updated adapter.

Type: Array of strings

Valid Values: `TABLES | FORMS | QUERIES | SIGNATURES | LAYOUT`

Errors

For information about the errors that are common to all actions, see Common Error Types.

AccessDeniedException

You aren't authorized to perform the action. Use the Amazon Resource Name (ARN) of an authorized user or IAM role to perform the operation.

HTTP Status Code: 400

ConflictException

Updating or deleting a resource can cause an inconsistent state.

HTTP Status Code: 400

InternalServerError

Amazon Textract experienced a service issue. Try your call again.

HTTP Status Code: 500

InvalidParameterException

An input parameter violated a constraint. For example, in synchronous operations, an `InvalidParameterException` exception occurs when neither of the `S3Object` or `Bytes` values are supplied in the `Document` request parameter. Validate your parameter before calling the API operation again.

HTTP Status Code: 400

ProvisionedThroughputExceededException

The number of requests exceeded your throughput limit. If you want to increase this limit, contact Amazon Textract.

HTTP Status Code: 400

ResourceNotFoundException

Returned when an operation tried to access a nonexistent resource.

HTTP Status Code: 400

ThrottlingException

Amazon Textract is temporarily unable to process the request. Try your call again.

HTTP Status Code: 500

ValidationException

Indicates that a request was not valid. Check request for proper formatting.

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS Command Line Interface V2
- AWS SDK for .NET V4
- AWS SDK for C++
- AWS SDK for Go v2
- AWS SDK for Java V2
- AWS SDK for JavaScript V3
- AWS SDK for Kotlin
- AWS SDK for PHP V3
- AWS SDK for Python
- AWS SDK for Ruby V3

Data Types

The Amazon Textract API contains several data types that various actions use. This section describes each data type in detail.

> **Note**
>
> The order of each element in a data type structure is not guaranteed. Applications should not assume a particular order.

The following data types are supported:

- Adapter
- AdapterOverview
- AdaptersConfig
- AdapterVersionDatasetConfig
- AdapterVersionEvaluationMetric
- AdapterVersionOverview
- AnalyzeIDDetections
- Block
- BoundingBox
- DetectedSignature
- Document
- DocumentGroup
- DocumentLocation
- DocumentMetadata
- EvaluationMetric
- ExpenseCurrency
- ExpenseDetection
- ExpenseDocument
- ExpenseField
- ExpenseGroupProperty

- ExpenseType
- Extraction
- Geometry
- HumanLoopActivationOutput
- HumanLoopConfig
- HumanLoopDataAttributes
- IdentityDocument
- IdentityDocumentField
- LendingDetection
- LendingDocument
- LendingField
- LendingResult
- LendingSummary
- LineItemFields
- LineItemGroup
- NormalizedValue
- NotificationChannel
- OutputConfig
- PageClassification
- Point
- Prediction
- QueriesConfig
- Query
- Relationship
- S3Object
- SignatureDetection
- SplitDocument
- UndetectedSignature
- Warning

Adapter

An adapter selected for use when analyzing documents. Contains an adapter ID and a version number. Contains information on pages selected for analysis when analyzing documents asychronously.

Contents

AdapterId

A unique identifier for the adapter resource.

Type: String

Length Constraints: Minimum length of 12. Maximum length of 1011.

Required: Yes

Version

A string that identifies the version of the adapter.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Required: Yes

Pages

Pages is a parameter that the user inputs to specify which pages to apply an adapter to. The following is a list of rules for using this parameter.

- If a page is not specified, it is set to `["1"]` by default.
- The following characters are allowed in the parameter's string: `0 1 2 3 4 5 6 7 8 9 - *`. No whitespace is allowed.
- When using * to indicate all pages, it must be the only element in the list.
- You can use page intervals, such as `["1-3", "1-1", "4-*"]`. Where * indicates last page of document.
- Specified pages must be greater than 0 and less than or equal to the number of pages in the document.

Type: Array of strings

Array Members: Minimum number of 1 item.

Length Constraints: Minimum length of 1. Maximum length of 9.

Pattern: `^[0-9\*\-]+$`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

AdapterOverview

Contains information on the adapter, including the adapter ID, Name, Creation time, and feature types.

Contents

AdapterId

A unique identifier for the adapter resource.

Type: String

Length Constraints: Minimum length of 12. Maximum length of 1011.

Required: No

AdapterName

A string naming the adapter resource.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Pattern: `[a-zA-Z0-9-_]+`

Required: No

CreationTime

The date and time that the adapter was created.

Type: Timestamp

Required: No

FeatureTypes

The feature types that the adapter is operating on.

Type: Array of strings

Valid Values: `TABLES | FORMS | QUERIES | SIGNATURES | LAYOUT`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

AdaptersConfig

Contains information about adapters used when analyzing a document, with each adapter specified using an AdapterId and version

Contents

Adapters

A list of adapters to be used when analyzing the specified document.

Type: Array of Adapter objects

Array Members: Minimum number of 1 item. Maximum number of 100 items.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

AdapterVersionDatasetConfig

The dataset configuration options for a given version of an adapter. Can include an Amazon S3 bucket if specified.

Contents

ManifestS3Object

The S3 bucket name and file name that identifies the document.

The AWS Region for the S3 bucket that contains the document must match the Region that you use for Amazon Textract operations.

For Amazon Textract to process a file in an S3 bucket, the user must have permission to access the S3 bucket and file.

Type: S3Object object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

AdapterVersionEvaluationMetric

Contains information on the metrics used to evalute the peformance of a given adapter version. Includes data for baseline model performance and individual adapter version perfromance.

Contents

AdapterVersion

The F1 score, precision, and recall metrics for the baseline model.

Type: EvaluationMetric object

Required: No

Baseline

The F1 score, precision, and recall metrics for the baseline model.

Type: EvaluationMetric object

Required: No

FeatureType

Indicates the feature type being analyzed by a given adapter version.

Type: String

Valid Values: `TABLES | FORMS | QUERIES | SIGNATURES | LAYOUT`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

AdapterVersionOverview

Summary info for an adapter version. Contains information on the AdapterId, AdapterVersion, CreationTime, FeatureTypes, and Status.

Contents

AdapterId

A unique identifier for the adapter associated with a given adapter version.

Type: String

Length Constraints: Minimum length of 12. Maximum length of 1011.

Required: No

AdapterVersion

An identified for a given adapter version.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Required: No

CreationTime

The date and time that a given adapter version was created.

Type: Timestamp

Required: No

FeatureTypes

The feature types that the adapter version is operating on.

Type: Array of strings

Valid Values: `TABLES | FORMS | QUERIES | SIGNATURES | LAYOUT`

Required: No

Status

Contains information on the status of a given adapter version.

Type: String

Valid Values: `ACTIVE | AT_RISK | DEPRECATED | CREATION_ERROR | CREATION_IN_PROGRESS`

Required: No

StatusMessage

A message explaining the status of a given adapter vesion.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 256.

Pattern: `^[a-zA-Z0-9\s!"\#\$%'&\(\)\*\+\,\-\./:;=\?@\[\\\]\^_`\{\|\}~><]+$`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

AnalyzeIDDetections

Used to contain the information detected by an AnalyzeID operation.

Contents

Text

Text of either the normalized field or value associated with it.

Type: String

Required: Yes

Confidence

The confidence score of the detected text.

Type: Float

Valid Range: Minimum value of 0. Maximum value of 100.

Required: No

NormalizedValue

Only returned for dates, returns the type of value detected and the date written in a more machine readable way.

Type: NormalizedValue object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

Block

A `Block` represents items that are recognized in a document within a group of pixels close to each other. The information returned in a `Block` object depends on the type of operation. In text detection for documents (for example DetectDocumentText), you get information about the detected words and lines of text. In text analysis (for example AnalyzeDocument), you can also get information about the fields, tables, and selection elements that are detected in the document.

An array of `Block` objects is returned by both synchronous and asynchronous operations. In synchronous operations, such as DetectDocumentText, the array of `Block` objects is the entire set of results. In asynchronous operations, such as GetDocumentAnalysis, the array is returned over one or more responses.

For more information, see How Amazon Textract Works.

Contents

BlockType

The type of text item that's recognized. In operations for text detection, the following types are returned:

- *PAGE* - Contains a list of the LINE `Block` objects that are detected on a document page.
- *WORD* - A word detected on a document page. A word is one or more ISO basic Latin script characters that aren't separated by spaces.
- *LINE* - A string of space-delimited, contiguous words that are detected on a document page.

In text analysis operations, the following types are returned:

- *PAGE* - Contains a list of child `Block` objects that are detected on a document page.
- *KEY_VALUE_SET* - Stores the KEY and VALUE `Block` objects for linked text that's detected on a document page. Use the `EntityType` field to determine if a KEY_VALUE_SET object is a KEY `Block` object or a VALUE `Block` object.
- *WORD* - A word that's detected on a document page. A word is one or more ISO basic Latin script characters that aren't separated by spaces.
- *LINE* - A string of tab-delimited, contiguous words that are detected on a document page.
- *TABLE* - A table that's detected on a document page. A table is grid-based information with two or more rows or columns, with a cell span of one row and one column each.

- *TABLE_TITLE* - The title of a table. A title is typically a line of text above or below a table, or embedded as the first row of a table.
- *TABLE_FOOTER* - The footer associated with a table. A footer is typically a line or lines of text below a table or embedded as the last row of a table.
- *CELL* - A cell within a detected table. The cell is the parent of the block that contains the text in the cell.
- *MERGED_CELL* - A cell in a table whose content spans more than one row or column. The `Relationships` array for this cell contain data from individual cells.
- *SELECTION_ELEMENT* - A selection element such as an option button (radio button) or a check box that's detected on a document page. Use the value of `SelectionStatus` to determine the status of the selection element.
- *SIGNATURE* - The location and confidence score of a signature detected on a document page. Can be returned as part of a Key-Value pair or a detected cell.
- *QUERY* - A question asked during the call of AnalyzeDocument. Contains an alias and an ID that attaches it to its answer.
- *QUERY_RESULT* - A response to a question asked during the call of analyze document. Comes with an alias and ID for ease of locating in a response. Also contains location and confidence score.

The following BlockTypes are only returned for Amazon Textract Layout.

- `LAYOUT_TITLE` - The main title of the document.
- `LAYOUT_HEADER` - Text located in the top margin of the document.
- `LAYOUT_FOOTER` - Text located in the bottom margin of the document.
- `LAYOUT_SECTION_HEADER` - The titles of sections within a document.
- `LAYOUT_PAGE_NUMBER` - The page number of the documents.
- `LAYOUT_LIST` - Any information grouped together in list form.
- `LAYOUT_FIGURE` - Indicates the location of an image in a document.
- `LAYOUT_TABLE` - Indicates the location of a table in the document.
- `LAYOUT_KEY_VALUE` - Indicates the location of form key-values in a document.
- `LAYOUT_TEXT` - Text that is present typically as a part of paragraphs in documents.

Type: String

Valid Values: KEY_VALUE_SET | PAGE | LINE | WORD | TABLE | CELL | SELECTION_ELEMENT | MERGED_CELL | TITLE | QUERY | QUERY_RESULT | SIGNATURE | TABLE_TITLE | TABLE_FOOTER | LAYOUT_TEXT | LAYOUT_TITLE | LAYOUT_HEADER | LAYOUT_FOOTER | LAYOUT_SECTION_HEADER | LAYOUT_PAGE_NUMBER | LAYOUT_LIST | LAYOUT_FIGURE | LAYOUT_TABLE | LAYOUT_KEY_VALUE

Required: No

ColumnIndex

The column in which a table cell appears. The first column position is 1. ColumnIndex isn't returned by DetectDocumentText and GetDocumentTextDetection.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

ColumnSpan

The number of columns that a table cell spans. ColumnSpan isn't returned by DetectDocumentText and GetDocumentTextDetection.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

Confidence

The confidence score that Amazon Textract has in the accuracy of the recognized text and the accuracy of the geometry points around the recognized text.

Type: Float

Valid Range: Minimum value of 0. Maximum value of 100.

Required: No

EntityTypes

The type of entity.

The following entity types can be returned by FORMS analysis:

- *KEY* - An identifier for a field on the document.
- *VALUE* - The field text.

The following entity types can be returned by TABLES analysis:

- *COLUMN_HEADER* - Identifies a cell that is a header of a column.
- *TABLE_TITLE* - Identifies a cell that is a title within the table.
- *TABLE_SECTION_TITLE* - Identifies a cell that is a title of a section within a table. A section title is a cell that typically spans an entire row above a section.
- *TABLE_FOOTER* - Identifies a cell that is a footer of a table.
- *TABLE_SUMMARY* - Identifies a summary cell of a table. A summary cell can be a row of a table or an additional, smaller table that contains summary information for another table.
- *STRUCTURED_TABLE* - Identifies a table with column headers where the content of each row corresponds to the headers.
- *SEMI_STRUCTURED_TABLE* - Identifies a non-structured table.

`EntityTypes` isn't returned by `DetectDocumentText` and `GetDocumentTextDetection`.

Type: Array of strings

Valid Values: `KEY | VALUE | COLUMN_HEADER | TABLE_TITLE | TABLE_FOOTER | TABLE_SECTION_TITLE | TABLE_SUMMARY | STRUCTURED_TABLE | SEMI_STRUCTURED_TABLE`

Required: No

Geometry

The location of the recognized text on the image. It includes an axis-aligned, coarse bounding box that surrounds the text, and a finer-grain polygon for more accurate spatial information.

Type: Geometry object

Required: No

Id

The identifier for the recognized text. The identifier is only unique for a single operation.

Type: String

Pattern: `.*\S.*`

Required: No

Page

The page on which a block was detected. `Page` is returned by synchronous and asynchronous operations. Page values greater than 1 are only returned for multipage documents that are in PDF or TIFF format. A scanned image (JPEG/PNG) provided to an asynchronous operation, even if it contains multiple document pages, is considered a single-page document. This means that for scanned images the value of `Page` is always 1.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

Query

Type: Query object

Required: No

Relationships

A list of relationship objects that describe how blocks are related to each other. For example, a LINE block object contains a CHILD relationship type with the WORD blocks that make up the line of text. There aren't Relationship objects in the list for relationships that don't exist, such as when the current block has no child blocks.

Type: Array of Relationship objects

Required: No

RowIndex

The row in which a table cell is located. The first row position is 1. `RowIndex` isn't returned by `DetectDocumentText` and `GetDocumentTextDetection`.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

RowSpan

The number of rows that a table cell spans. RowSpan isn't returned by `DetectDocumentText` and `GetDocumentTextDetection`.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

SelectionStatus

The selection status of a selection element, such as an option button or check box.

Type: String

Valid Values: `SELECTED | NOT_SELECTED`

Required: No

Text

The word or line of text that's recognized by Amazon Textract.

Type: String

Required: No

TextType

The kind of text that Amazon Textract has detected. Can check for handwritten text and printed text.

Type: String

Valid Values: `HANDWRITING | PRINTED`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

BoundingBox

The bounding box around the detected page, text, key-value pair, table, table cell, or selection element on a document page. The `left` (x-coordinate) and `top` (y-coordinate) are coordinates that represent the top and left sides of the bounding box. Note that the upper-left corner of the image is the origin (0,0).

The `top` and `left` values returned are ratios of the overall document page size. For example, if the input image is 700 x 200 pixels, and the top-left coordinate of the bounding box is 350 x 50 pixels, the API returns a `left` value of 0.5 (350/700) and a `top` value of 0.25 (50/200).

The `width` and `height` values represent the dimensions of the bounding box as a ratio of the overall document page dimension. For example, if the document page size is 700 x 200 pixels, and the bounding box width is 70 pixels, the width returned is 0.1.

Contents

Height

The height of the bounding box as a ratio of the overall document page height.

Type: Float

Required: No

Left

The left coordinate of the bounding box as a ratio of overall document page width.

Type: Float

Required: No

Top

The top coordinate of the bounding box as a ratio of overall document page height.

Type: Float

Required: No

Width

The width of the bounding box as a ratio of the overall document page width.

Type: Float

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

DetectedSignature

A structure that holds information regarding a detected signature on a page.

Contents

Page

The page a detected signature was found on.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

Document

The input document, either as bytes or as an S3 object.

You pass image bytes to an Amazon Textract API operation by using the `Bytes` property. For example, you would use the `Bytes` property to pass a document loaded from a local file system. Image bytes passed by using the `Bytes` property must be base64 encoded. Your code might not need to encode document file bytes if you're using an AWS SDK to call Amazon Textract API operations.

You pass images stored in an S3 bucket to an Amazon Textract API operation by using the `S3Object` property. Documents stored in an S3 bucket don't need to be base64 encoded.

The AWS Region for the S3 bucket that contains the S3 object must match the AWS Region that you use for Amazon Textract operations.

If you use the AWS CLI to call Amazon Textract operations, passing image bytes using the Bytes property isn't supported. You must first upload the document to an Amazon S3 bucket, and then call the operation using the S3Object property.

For Amazon Textract to process an S3 object, the user must have permission to access the S3 object.

Contents

Bytes

A blob of base64-encoded document bytes. The maximum size of a document that's provided in a blob of bytes is 5 MB. The document bytes must be in PNG or JPEG format.

If you're using an AWS SDK to call Amazon Textract, you might not need to base64-encode image bytes passed using the `Bytes` field.

Type: Base64-encoded binary data object

Length Constraints: Minimum length of 1. Maximum length of 10485760.

Required: No

S3Object

Identifies an S3 object as the document source. The maximum size of a document that's stored in an S3 bucket is 5 MB.

Type: S3Object object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

DocumentGroup

Summary information about documents grouped by the same document type.

Contents

DetectedSignatures

A list of the detected signatures found in a document group.

Type: Array of DetectedSignature objects

Required: No

SplitDocuments

An array that contains information about the pages of a document, defined by logical boundary.

Type: Array of SplitDocument objects

Required: No

Type

The type of document that Amazon Textract has detected. See Analyze Lending Response Objects for a list of all types returned by Textract.

Type: String

Pattern: `.*\S.*`

Required: No

UndetectedSignatures

A list of any expected signatures not found in a document group.

Type: Array of UndetectedSignature objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

DocumentLocation

The Amazon S3 bucket that contains the document to be processed. It's used by asynchronous operations.

The input document can be an image file in JPEG or PNG format. It can also be a file in PDF format.

Contents

S3Object

The Amazon S3 bucket that contains the input document.

Type: S3Object object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

DocumentMetadata

Information about the input document.

Contents

Pages

The number of pages that are detected in the document.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

EvaluationMetric

The evaluation metrics (F1 score, Precision, and Recall) for an adapter version.

Contents

F1Score

The F1 score for an adapter version.

Type: Float

Required: No

Precision

The Precision score for an adapter version.

Type: Float

Required: No

Recall

The Recall score for an adapter version.

Type: Float

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

ExpenseCurrency

Returns the kind of currency detected.

Contents

Code

Currency code for detected currency. the current supported codes are:

- USD
- EUR
- GBP
- CAD
- INR
- JPY
- CHF
- AUD
- CNY
- BZR
- SEK
- HKD

Type: String

Required: No

Confidence

Percentage confideence in the detected currency.

Type: Float

Valid Range: Minimum value of 0. Maximum value of 100.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

ExpenseDetection

An object used to store information about the Value or Label detected by Amazon Textract.

Contents

Confidence

The confidence in detection, as a percentage

Type: Float

Valid Range: Minimum value of 0. Maximum value of 100.

Required: No

Geometry

Information about where the following items are located on a document page: detected page, text, key-value pairs, tables, table cells, and selection elements.

Type: Geometry object

Required: No

Text

The word or line of text recognized by Amazon Textract

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

ExpenseDocument

The structure holding all the information returned by AnalyzeExpense

Contents

Blocks

This is a block object, the same as reported when DetectDocumentText is run on a document. It provides word level recognition of text.

Type: Array of Block objects

Required: No

ExpenseIndex

Denotes which invoice or receipt in the document the information is coming from. First document will be 1, the second 2, and so on.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

LineItemGroups

Information detected on each table of a document, seperated into `LineItems`.

Type: Array of LineItemGroup objects

Required: No

SummaryFields

Any information found outside of a table by Amazon Textract.

Type: Array of ExpenseField objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

ExpenseField

Breakdown of detected information, seperated into the catagories Type, LabelDetection, and ValueDetection

Contents

Currency

Shows the kind of currency, both the code and confidence associated with any monatary value detected.

Type: ExpenseCurrency object

Required: No

GroupProperties

Shows which group a response object belongs to, such as whether an address line belongs to the vendor's address or the recipent's address.

Type: Array of ExpenseGroupProperty objects

Required: No

LabelDetection

The explicitly stated label of a detected element.

Type: ExpenseDetection object

Required: No

PageNumber

The page number the value was detected on.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

Type

The implied label of a detected element. Present alongside LabelDetection for explicit elements.

Type: ExpenseType object

Required: No

ValueDetection

The value of a detected element. Present in explicit and implicit elements.

Type: ExpenseDetection object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

ExpenseGroupProperty

Shows the group that a certain key belongs to. This helps differentiate between names and addresses for different organizations, that can be hard to determine via JSON response.

Contents

Id

Provides a group Id number, which will be the same for each in the group.

Type: String

Required: No

Types

Informs you on whether the expense group is a name or an address.

Type: Array of strings

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

ExpenseType

An object used to store information about the Type detected by Amazon Textract.

Contents

Confidence

The confidence of accuracy, as a percentage.

Type: Float

Valid Range: Minimum value of 0. Maximum value of 100.

Required: No

Text

The word or line of text detected by Amazon Textract.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

Extraction

Contains information extracted by an analysis operation after using StartLendingAnalysis.

Contents

ExpenseDocument

The structure holding all the information returned by AnalyzeExpense

Type: ExpenseDocument object

Required: No

IdentityDocument

The structure that lists each document processed in an AnalyzeID operation.

Type: IdentityDocument object

Required: No

LendingDocument

Holds the structured data returned by AnalyzeDocument for lending documents.

Type: LendingDocument object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

Geometry

Information about where the following items are located on a document page: detected page, text, key-value pairs, tables, table cells, and selection elements.

Contents

BoundingBox

An axis-aligned coarse representation of the location of the recognized item on the document page.

Type: BoundingBox object

Required: No

Polygon

Within the bounding box, a fine-grained polygon around the recognized item.

Type: Array of Point objects

Required: No

RotationAngle

Provides a numerical value corresponding to the rotation of the WORD block. Possible values are 0, 90, 180, and 270.

Type: Float

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

HumanLoopActivationOutput

Shows the results of the human in the loop evaluation. If there is no HumanLoopArn, the input did not trigger human review.

Contents

HumanLoopActivationConditionsEvaluationResults

Shows the result of condition evaluations, including those conditions which activated a human review.

Type: String

Length Constraints: Maximum length of 10240.

Required: No

HumanLoopActivationReasons

Shows if and why human review was needed.

Type: Array of strings

Array Members: Minimum number of 1 item.

Required: No

HumanLoopArn

The Amazon Resource Name (ARN) of the HumanLoop created.

Type: String

Length Constraints: Maximum length of 256.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

HumanLoopConfig

Sets up the human review workflow the document will be sent to if one of the conditions is met. You can also set certain attributes of the image before review.

Contents

FlowDefinitionArn

The Amazon Resource Name (ARN) of the flow definition.

Type: String

Length Constraints: Maximum length of 256.

Required: Yes

HumanLoopName

The name of the human workflow used for this image. This should be kept unique within a region.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 63.

Pattern: `^[a-z0-9](-*[a-z0-9])*`

Required: Yes

DataAttributes

Sets attributes of the input data.

Type: HumanLoopDataAttributes object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

HumanLoopDataAttributes

Allows you to set attributes of the image. Currently, you can declare an image as free of personally identifiable information and adult content.

Contents

ContentClassifiers

Sets whether the input image is free of personally identifiable information or adult content.

Type: Array of strings

Array Members: Maximum number of 256 items.

Valid Values: `FreeOfPersonallyIdentifiableInformation | FreeOfAdultContent`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

IdentityDocument

The structure that lists each document processed in an AnalyzeID operation.

Contents

Blocks

Individual word recognition, as returned by document detection.

Type: Array of Block objects

Required: No

DocumentIndex

Denotes the placement of a document in the IdentityDocument list. The first document is marked 1, the second 2 and so on.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

IdentityDocumentFields

The structure used to record information extracted from identity documents. Contains both normalized field and value of the extracted text.

Type: Array of IdentityDocumentField objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

IdentityDocumentField

Structure containing both the normalized type of the extracted information and the text associated with it. These are extracted as Type and Value respectively.

Contents

Type

Used to contain the information detected by an AnalyzeID operation.

Type: AnalyzeIDDetections object

Required: No

ValueDetection

Used to contain the information detected by an AnalyzeID operation.

Type: AnalyzeIDDetections object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

LendingDetection

The results extracted for a lending document.

Contents

Confidence

The confidence level for the text of a detected value in a lending document.

Type: Float

Valid Range: Minimum value of 0. Maximum value of 100.

Required: No

Geometry

Information about where the following items are located on a document page: detected page, text, key-value pairs, tables, table cells, and selection elements.

Type: Geometry object

Required: No

SelectionStatus

The selection status of a selection element, such as an option button or check box.

Type: String

Valid Values: `SELECTED | NOT_SELECTED`

Required: No

Text

The text extracted for a detected value in a lending document.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

LendingDocument

Holds the structured data returned by AnalyzeDocument for lending documents.

Contents

LendingFields

An array of LendingField objects.

Type: Array of LendingField objects

Required: No

SignatureDetections

A list of signatures detected in a lending document.

Type: Array of SignatureDetection objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

LendingField

Holds the normalized key-value pairs returned by AnalyzeDocument, including the document type, detected text, and geometry.

Contents

KeyDetection

The results extracted for a lending document.

Type: LendingDetection object

Required: No

Type

The type of the lending document.

Type: String

Required: No

ValueDetections

An array of LendingDetection objects.

Type: Array of LendingDetection objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

LendingResult

Contains the detections for each page analyzed through the Analyze Lending API.

Contents

Extractions

An array of Extraction to hold structured data. e.g. normalized key value pairs instead of raw OCR detections .

Type: Array of Extraction objects

Required: No

Page

The page number for a page, with regard to whole submission.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

PageClassification

The classifier result for a given page.

Type: PageClassification object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

LendingSummary

Contains information regarding DocumentGroups and UndetectedDocumentTypes.

Contents

DocumentGroups

Contains an array of all DocumentGroup objects.

Type: Array of DocumentGroup objects

Required: No

UndetectedDocumentTypes

UndetectedDocumentTypes.

Type: Array of strings

Pattern: `.*\S.*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

LineItemFields

A structure that holds information about the different lines found in a document's tables.

Contents

LineItemExpenseFields

ExpenseFields used to show information from detected lines on a table.

Type: Array of ExpenseField objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

LineItemGroup

A grouping of tables which contain LineItems, with each table identified by the table's `LineItemGroupIndex`.

Contents

LineItemGroupIndex

The number used to identify a specific table in a document. The first table encountered will have a LineItemGroupIndex of 1, the second 2, etc.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

LineItems

The breakdown of information on a particular line of a table.

Type: Array of LineItemFields objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

NormalizedValue

Contains information relating to dates in a document, including the type of value, and the value.

Contents

Value

The value of the date, written as Year-Month-DayTHour:Minute:Second.

Type: String

Required: No

ValueType

The normalized type of the value detected. In this case, DATE.

Type: String

Valid Values: DATE

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

NotificationChannel

The Amazon Simple Notification Service (Amazon SNS) topic to which Amazon Textract publishes the completion status of an asynchronous document operation.

Contents

RoleArn

The Amazon Resource Name (ARN) of an IAM role that gives Amazon Textract publishing permissions to the Amazon SNS topic.

Type: String

Length Constraints: Minimum length of 20. Maximum length of 2048.

Pattern: `arn:([a-z\d-]+):iam::\d{12}:role/?[a-zA-Z_0-9+=,.@\-_/]+`

Required: Yes

SNSTopicArn

The Amazon SNS topic that Amazon Textract posts the completion status to.

Type: String

Length Constraints: Minimum length of 20. Maximum length of 1024.

Pattern: `(^arn:([a-z\d-]+):sns:[a-zA-Z\d-]{1,20}:\w{12}:.+$)`

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

OutputConfig

Sets whether or not your output will go to a user created bucket. Used to set the name of the bucket, and the prefix on the output file.

`OutputConfig` is an optional parameter which lets you adjust where your output will be placed. By default, Amazon Textract will store the results internally and can only be accessed by the Get API operations. With `OutputConfig` enabled, you can set the name of the bucket the output will be sent to the file prefix of the results where you can download your results. Additionally, you can set the `KMSKeyID` parameter to a customer master key (CMK) to encrypt your output. Without this parameter set Amazon Textract will encrypt server-side using the AWS managed CMK for Amazon S3.

Decryption of Customer Content is necessary for processing of the documents by Amazon Textract. If your account is opted out under an AI services opt out policy then all unencrypted Customer Content is immediately and permanently deleted after the Customer Content has been processed by the service. No copy of of the output is retained by Amazon Textract. For information about how to opt out, see Managing AI services opt-out policy.

For more information on data privacy, see the Data Privacy FAQ.

Contents

S3Bucket

The name of the bucket your output will go to.

Type: String

Length Constraints: Minimum length of 3. Maximum length of 255.

Pattern: `[0-9A-Za-z\.\-_]*`

Required: Yes

S3Prefix

The prefix of the object key that the output will be saved to. When not enabled, the prefix will be “textract_output".

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1024.

Pattern: `.*\S.*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

PageClassification

The class assigned to a Page object detected in an input document. Contains information regarding the predicted type/class of a document's page and the page number that the Page object was detected on.

Contents

PageNumber

The page number the value was detected on, relative to Amazon Textract's starting position.

Type: Array of Prediction objects

Required: Yes

PageType

The class, or document type, assigned to a detected Page object. The class, or document type, assigned to a detected Page object.

Type: Array of Prediction objects

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

Point

The X and Y coordinates of a point on a document page. The X and Y values that are returned are ratios of the overall document page size. For example, if the input document is 700 x 200 and the operation returns X=0.5 and Y=0.25, then the point is at the (350,50) pixel coordinate on the document page.

An array of `Point` objects, `Polygon`, is returned as part of the Geometry object that's returned in a Block object. A `Polygon` object represents a fine-grained polygon around detected text, a key-value pair, a table, a table cell, or a selection element.

Contents

X

The value of the X coordinate for a point on a `Polygon`.

Type: Float

Required: No

Y

The value of the Y coordinate for a point on a `Polygon`.

Type: Float

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

Prediction

Contains information regarding predicted values returned by Amazon Textract operations, including the predicted value and the confidence in the predicted value.

Contents

Confidence

Amazon Textract's confidence in its predicted value.

Type: Float

Valid Range: Minimum value of 0. Maximum value of 100.

Required: No

Value

The predicted value of a detected object.

Type: String

Pattern: `.*\S.*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

QueriesConfig

Contents

Queries

Type: Array of Query objects

Array Members: Minimum number of 1 item.

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

Query

Each query contains the question you want to ask in the Text and the alias you want to associate.

Contents

Text

Question that Amazon Textract will apply to the document. An example would be "What is the customer's SSN?"

Type: String

Length Constraints: Minimum length of 1. Maximum length of 200.

Pattern: `^[a-zA-Z0-9\s!"\#\$%'&\(\)\*\+\,\-\./:;=\?@\[\\\]\^_`\{\|\}~><]+$`

Required: Yes

Alias

Alias attached to the query, for ease of location.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 200.

Pattern: `^[a-zA-Z0-9\s!"\#\$%'&\(\)\*\+\,\-\./:;=\?@\[\\\]\^_`\{\|\}~><]+$`

Required: No

Pages

Pages is a parameter that the user inputs to specify which pages to apply a query to. The following is a list of rules for using this parameter.

- If a page is not specified, it is set to `["1"]` by default.
- The following characters are allowed in the parameter's string: `0 1 2 3 4 5 6 7 8 9 - *`. No whitespace is allowed.
- When using * to indicate all pages, it must be the only element in the list.
- You can use page intervals, such as `[“1-3”, “1-1”, “4-*”]`. Where * indicates last page of document.

- Specified pages must be greater than 0 and less than or equal to the number of pages in the document.

Type: Array of strings

Array Members: Minimum number of 1 item.

Length Constraints: Minimum length of 1. Maximum length of 9.

Pattern: `^[0-9\*\-]+$`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

Relationship

Information about how blocks are related to each other. A `Block` object contains 0 or more `Relation` objects in a list, `Relationships`. For more information, see Block.

The `Type` element provides the type of the relationship for all blocks in the `IDs` array.

Contents

Ids

An array of IDs for related blocks. You can get the type of the relationship from the `Type` element.

Type: Array of strings

Pattern: `.*\S.*`

Required: No

Type

The type of relationship between the blocks in the IDs array and the current block. The following list describes the relationship types that can be returned.

- *VALUE* - A list that contains the ID of the VALUE block that's associated with the KEY of a key-value pair.
- *CHILD* - A list of IDs that identify blocks found within the current block object. For example, WORD blocks have a CHILD relationship to the LINE block type.
- *MERGED_CELL* - A list of IDs that identify each of the MERGED_CELL block types in a table.
- *ANSWER* - A list that contains the ID of the QUERY_RESULT block that’s associated with the corresponding QUERY block.
- *TABLE* - A list of IDs that identify associated TABLE block types.
- *TABLE_TITLE* - A list that contains the ID for the TABLE_TITLE block type in a table.
- *TABLE_FOOTER* - A list of IDs that identify the TABLE_FOOTER block types in a table.

Type: String

Valid Values: `VALUE | CHILD | COMPLEX_FEATURES | MERGED_CELL | TITLE | ANSWER | TABLE | TABLE_TITLE | TABLE_FOOTER`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

S3Object

The S3 bucket name and file name that identifies the document.

The AWS Region for the S3 bucket that contains the document must match the Region that you use for Amazon Textract operations.

For Amazon Textract to process a file in an S3 bucket, the user must have permission to access the S3 bucket and file.

Contents

Bucket

The name of the S3 bucket. Note that the # character is not valid in the file name.

Type: String

Length Constraints: Minimum length of 3. Maximum length of 255.

Pattern: `[0-9A-Za-z\.\-_]*`

Required: No

Name

The file name of the input document. Image files may be in PDF, TIFF, JPEG, or PNG format.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1024.

Pattern: `.*\S.*`

Required: No

Version

If the bucket has versioning enabled, you can specify the object version.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1024.

Pattern: `.*\S.*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

SignatureDetection

Information regarding a detected signature on a page.

Contents

Confidence

The confidence, from 0 to 100, in the predicted values for a detected signature.

Type: Float

Valid Range: Minimum value of 0. Maximum value of 100.

Required: No

Geometry

Information about where the following items are located on a document page: detected page, text, key-value pairs, tables, table cells, and selection elements.

Type: Geometry object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

SplitDocument

Contains information about the pages of a document, defined by logical boundary.

Contents

Index

The index for a given document in a DocumentGroup of a specific Type.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

Pages

An array of page numbers for a for a given document, ordered by logical boundary.

Type: Array of integers

Valid Range: Minimum value of 0.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

UndetectedSignature

A structure containing information about an undetected signature on a page where it was expected but not found.

Contents

Page

The page where a signature was expected but not found.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

Warning

A warning about an issue that occurred during asynchronous text analysis (StartDocumentAnalysis) or asynchronous document text detection (StartDocumentTextDetection).

Contents

ErrorCode

The error code for the warning.

Type: String

Required: No

Pages

A list of the pages that the warning applies to.

Type: Array of integers

Valid Range: Minimum value of 0.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- AWS SDK for C++
- AWS SDK for Java V2
- AWS SDK for Ruby V3

Common Parameters

The following list contains the parameters that all actions use for signing Signature Version 4 requests with a query string. Any action-specific parameters are listed in the topic for that action. For more information about Signature Version 4, see Signing AWS API requests in the *IAM User Guide*.

X-Amz-Algorithm

The hash algorithm that you used to create the request signature.

Condition: Specify this parameter when you include authentication information in a query string instead of in the HTTP authorization header.

Type: string

Valid Values: `AWS4-HMAC-SHA256`

Required: Conditional

X-Amz-Credential

The credential scope value, which is a string that includes your access key, the date, the region you are targeting, the service you are requesting, and a termination string ("aws4_request"). The value is expressed in the following format: *access_key*/*YYYYMMDD*/*region*/*service*/aws4_request.

For more information, see Create a signed AWS API request in the *IAM User Guide*.

Condition: Specify this parameter when you include authentication information in a query string instead of in the HTTP authorization header.

Type: string

Required: Conditional

X-Amz-Date

The date that is used to create the signature. The format must be ISO 8601 basic format (YYYYMMDD'T'HHMMSS'Z'). For example, the following date time is a valid X-Amz-Date value: `20120325T120000Z`.

Condition: X-Amz-Date is optional for all requests; it can be used to override the date used for signing requests. If the Date header is specified in the ISO 8601 basic format, X-Amz-Date is not required. When X-Amz-Date is used, it always overrides the value of the Date header. For more information, see Elements of an AWS API request signature in the *IAM User Guide*.

Type: string

Required: Conditional

X-Amz-Security-Token

The temporary security token that was obtained through a call to AWS Security Token Service (AWS STS). For a list of services that support temporary security credentials from AWS STS, see AWS services that work with IAM in the *IAM User Guide*.

Condition: If you're using temporary security credentials from AWS STS, you must include the security token.

Type: string

Required: Conditional

X-Amz-Signature

Specifies the hex-encoded signature that was calculated from the string to sign and the derived signing key.

Condition: Specify this parameter when you include authentication information in a query string instead of in the HTTP authorization header.

Type: string

Required: Conditional

X-Amz-SignedHeaders

Specifies all the HTTP headers that were included as part of the canonical request. For more information about specifying signed headers, see Create a signed AWS API request in the *IAM User Guide*.

Condition: Specify this parameter when you include authentication information in a query string instead of in the HTTP authorization header.

Type: string

Required: Conditional

Common Error Types

This section lists common error types that this AWS service may return. Not all services return all error types listed here. For errors specific to an API action for this service, see the topic for that API action.

AccessDeniedException

You don't have permission to perform this action. Verify that your IAM policy includes the required permissions.

HTTP Status Code: 403

ExpiredTokenException

The security token included in the request has expired. Request a new security token and try again.

HTTP Status Code: 403

IncompleteSignature

The request signature doesn't conform to AWS standards. Verify that you're using valid AWS credentials and that your request is properly formatted. If you're using an SDK, ensure it's up to date.

HTTP Status Code: 403

InternalFailure

The request can't be processed right now because of an internal server issue. Try again later. If the problem persists, contact AWS Support.

HTTP Status Code: 500

MalformedHttpRequestException

The request body can't be processed. This typically happens when the request body can't be decompressed using the specified content encoding algorithm. Verify that the content encoding header matches the compression format used.

HTTP Status Code: 400

NotAuthorized

You don't have permissions to perform this action. Verify that your IAM policy includes the required permissions.

HTTP Status Code: 401

OptInRequired

Your AWS account needs a subscription for this service. Verify that you've enabled the service in your account.

HTTP Status Code: 403

RequestAbortedException

The request was aborted before a response could be returned. This typically happens when the client closes the connection.

HTTP Status Code: 400

RequestEntityTooLargeException

The request entity is too large. Reduce the size of the request body and try again.

HTTP Status Code: 413

RequestTimeoutException

The request timed out. The server didn't receive the complete request within the expected time frame. Try again.

HTTP Status Code: 408

ServiceUnavailable

The service is temporarily unavailable. Try again later.

HTTP Status Code: 503

ThrottlingException

Your request rate is too high. The AWS SDKs automatically retry requests that receive this exception. Reduce the frequency of requests.

HTTP Status Code: 400

UnknownOperationException

The action or operation isn't recognized. Verify that the action name is spelled correctly and that it's supported by the API version you're using.

HTTP Status Code: 404

UnrecognizedClientException

The X.509 certificate or AWS access key ID you provided doesn't exist in our records. Verify that you're using valid credentials and that they haven't expired.

HTTP Status Code: 403

ValidationError

The input doesn't meet the required format or constraints. Check that all required parameters are included and that values are valid.

HTTP Status Code: 400