Contents

Overview ........................................................................................................................................... 4
Cost ................................................................................................................................................ 4
Architecture Overview ..................................................................................................................... 5
Solution Components ....................................................................................................................... 6
Image Analysis .................................................................................................................................. 6
Video Analysis .................................................................................................................................... 6
  Media Conversion ............................................................................................................................ 6
Audio Analysis .................................................................................................................................... 6
Face Indexing ..................................................................................................................................... 7
Web Interface ...................................................................................................................................... 7
  Customization .................................................................................................................................. 7
Considerations ..................................................................................................................................... 8
Limits ................................................................................................................................................ 8
  Custom Sizing .................................................................................................................................. 8
  Regional Deployment ....................................................................................................................... 8
AWS CloudFormation Template .......................................................................................................... 9
Automated Deployment ..................................................................................................................... 9
  What We’ll Cover .............................................................................................................................. 9
  Step 1. Launch the Stack .................................................................................................................. 10
  Step 2. Upload a File for Processing ............................................................................................... 11
  Step 3. View Extracted Metadata .................................................................................................. 11
Security ............................................................................................................................................. 11
  Amazon S3 Server-Side Encryption (SSE) ..................................................................................... 12
  Amazon CloudFront ......................................................................................................................... 12
Additional Resources ....................................................................................................................... 12
Appendix A: Media Analysis Solution API ....................................................................................... 13
  GET /details/{object_id} ................................................................................................................ 13
<table>
<thead>
<tr>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>GET /lookup/{object_id}/{lookup_type}?page={ page}</td>
<td>13</td>
</tr>
<tr>
<td>Description</td>
<td>13</td>
</tr>
<tr>
<td>Request Parameter</td>
<td>13</td>
</tr>
<tr>
<td>Response</td>
<td>13</td>
</tr>
<tr>
<td>GET /search?searchterm={search_term}&amp;page={page}</td>
<td>14</td>
</tr>
<tr>
<td>Description</td>
<td>14</td>
</tr>
<tr>
<td>Request Parameter</td>
<td>14</td>
</tr>
<tr>
<td>Response</td>
<td>14</td>
</tr>
<tr>
<td>GET /status/{object_id}</td>
<td>15</td>
</tr>
<tr>
<td>Description</td>
<td>15</td>
</tr>
<tr>
<td>Request Parameter</td>
<td>15</td>
</tr>
<tr>
<td>Response</td>
<td>15</td>
</tr>
</tbody>
</table>

**Appendix B: Media Analysis State Machine** ........................................... 16

<table>
<thead>
<tr>
<th>Images</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>17</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Videos</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>17</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Audio</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>18</td>
</tr>
</tbody>
</table>

**Appendix C: Customizing the Solution** ............................................ 19

<table>
<thead>
<tr>
<th>Step 1. Update the State Machine</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>20</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Step 2. Add Logic to AWS Lambda Code</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>20</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Step 3. Update the AWS IAM Policy</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>22</td>
</tr>
</tbody>
</table>

**Appendix D: Collection of Anonymous Data** ..................................... 23

<table>
<thead>
<tr>
<th>Source Code</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>24</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Document Revisions</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>24</td>
</tr>
</tbody>
</table>
About This Guide
This implementation guide discusses architectural considerations and configuration steps for deploying the Media Analysis Solution on the Amazon Web Services (AWS) Cloud. It includes links to an AWS CloudFormation template that launches, configures, and runs the AWS services required to deploy this solution using AWS best practices for security and availability.

The guide is intended for IT infrastructure architects and developers who have practical experience architecting on the AWS Cloud.

Overview
Amazon Web Services (AWS) offers powerful and cost-effective services to help customers process, analyze, and extract meaningful data from their audio, image, and video files. Customers who want to obtain a broader understanding of their media libraries can use these services to develop solutions that analyze and extract valuable metadata from their media files. Customers can also use various machine learning tools and algorithms to develop their own analytics solutions in the AWS Cloud. However, developing these solutions can require extensive knowledge of deep-learning algorithms and artificial intelligence (AI) services.

To help customers more easily analyze and understand their media files, AWS offers the Media Analysis Solution, a reference implementation that uses serverless, AWS-native AI services to automatically extract valuable metadata from media files. This solution combines Amazon Rekognition, which provides highly accurate object, scene, and activity detection; facial analysis and recognition; pathing; and celebrity detection in videos and images, Amazon Transcribe, an automatic speech recognition service, and Amazon Comprehend, to automatically transcribe audio and extract key phrases and entities from transcripts, to quickly and seamlessly extract key details from their media files in their AWS accounts without machine learning expertise.

Cost
You are responsible for the cost of the AWS services used while running the Media Analysis Solution. The total cost for running this solution depends on the size of media and metadata files stored in Amazon Simple Storage Service (Amazon S3), the duration and number of executed AWS Lambda functions, the number of AWS Step Functions state transitions, running the solution’s Amazon Elasticsearch Service cluster, and the size and length of media files analyzed with Amazon Rekognition, Amazon Transcribe, and Amazon Comprehend. You will also incur charges for requests made to the Amazon API Gateway. For full details, see the pricing webpage for each AWS service you will be using in this solution.
Architecture Overview
Deploying the Media Analysis Solution builds the following environment in the AWS Cloud.

The AWS CloudFormation template deploys an Amazon API Gateway RESTful API, an AWS Lambda function to process API requests, and an Amazon Cognito user pool which enable customers to securely interact with their media files. The solution also creates an AWS Step Functions state machine that orchestrates the media analysis processes, and a second Lambda function that executes the analysis and metadata extraction using managed artificial intelligence (AI) services such as Amazon Rekognition, Amazon Transcribe, and Amazon Comprehend. When an MP4 video file is uploaded, AWS Elemental MediaConvert extracts audio for the analysis by Amazon Transcribe and Amazon Comprehend. Metadata results are stored in an Amazon Simple Storage Service (Amazon S3) bucket and indexed in an Amazon Elasticsearch Service (Amazon ES) cluster.

The solution also creates a web interface and deploys it in an Amazon S3 bucket to allow customers to immediately start analyzing small media files with a simple web interface. Amazon CloudFront is used to restrict access to the solution’s website bucket contents.
Solution Components

Image Analysis
When you upload a PNG, JPG, or JPEG image to the solution’s encrypted Amazon S3 bucket, a Lambda function invokes the Step Functions state machine which enters a parallel state that simultaneously executes three branches, each responsible for orchestrating a different type of image analysis: label detection, celebrity recognition, and face search. The metadata results are stored in the Amazon S3 bucket and returned to the state machine as the task output.

When the analysis is complete, the state machine enters a final state that indexes the results in the Elasticsearch cluster. You can search and retrieve the image metadata using the solution API or web interface.

For more information on the image analysis state machine, see Appendix B.

Video Analysis
When you upload an MOV or MP4 video, a Lambda function starts the state machine which enters a parallel state that executes five branches, each responsible for orchestrating a different type of video analysis: label detection, celebrity recognition, face detection, face search, and pathing. The metadata results are stored in Amazon S3 and returned to the state machine as the task output.

When the analysis is complete, the results are indexed in the Elasticsearch cluster where they can be searched and retrieved using the solution API or web interface.

Media Conversion
When an MP4 video file is uploaded, a Lambda function triggers an AWS Elemental MediaConvert job to create a separate audio file. When the job is complete MediaConvert will upload the audio file into Amazon S3 and returned to the audio analysis state machine.

For more information on the video analysis state machine, see Appendix B.

Audio Analysis
When a FLAC, MP3, WAV, or MP4 audio file is uploaded, a Lambda function starts the state machine which orchestrates the audio transcription process using Amazon Transcribe.

Once the transcription is complete, the state machine leverages the resulting transcript to perform natural language processing using Amazon Comprehend. The state machine enters
a parallel state that simultaneously executes two branches to perform key phrase and key entity detection.

When the analysis is complete, the results are indexed in the Elasticsearch cluster where they can be searched and retrieved using the solution API or web interface.

For more information on the audio analysis state machine, see Appendix B.

**Face Indexing**
The Media Analysis Solution allows you to upload images of faces to be indexed in an Amazon Rekognition collection. These faces are automatically used during face searches on images and videos.

The first time you upload a face image to the solution, a Lambda function will create a collection for you, then it will index the face. When a new face image is uploaded, the face is added to your collection.

**Web Interface**
The Media Analysis Solution features a simple static web interface that makes it easier to upload files, index faces, browse media files, and view detailed search results. This web interface can be used as a reference for building your own media analysis applications. The interface leverages Amazon Cognito for user authentication, an Amazon API Gateway RESTful API for search and metadata retrieval, AWS Amplify for interacting with cloud services, and is powered by web assets hosted in an Amazon S3 bucket.

When authenticated users upload files through the interface, the files are stored in private folders that correspond to their unique Amazon Cognito identifier to ensure fine-grained access control using AWS Identity and Access Management (IAM) policies.

**Customization**
To leverage existing tools and applications to upload files or develop your own, consider the following Amazon S3 naming conventions.

<table>
<thead>
<tr>
<th>File Type</th>
<th>Convention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Media files</td>
<td>private/&lt;Amazon Cognito Identity ID&gt;/media/&lt;UUID v4&gt;/content/filename.ext</td>
</tr>
<tr>
<td>Raw extracted metadata</td>
<td>Private/&lt;Amazon Cognito Identity ID&gt;/media/&lt;UUID v4&gt;/results/ filename.ext</td>
</tr>
<tr>
<td>File Type</td>
<td>Convention</td>
</tr>
<tr>
<td>-----------</td>
<td>------------</td>
</tr>
<tr>
<td>Face images (to be added to your Amazon Rekognition collection)</td>
<td>Private/&lt;Amazon Cognito Identity ID&gt;/collection/&lt;UUID v4&gt;/externalImageId.ext</td>
</tr>
</tbody>
</table>

**Considerations**

**Limits**
Currently, the Media Analysis Solution has the following known limitations:

- **Web interface files are limited to 100MB**
  Files uploaded to the Media Analysis Solution through the demo web interface must be limited to 100MB. Customers can upload larger files to Amazon S3 through their own tools and applications by using an AWS SDK or the AWS CLI.

- **Audio analysis is limited to audio files 1024MB or smaller**
  The audio analysis state machine will not process audio files greater than 1024MB.

**Custom Sizing**
Choose from three preset Amazon ES cluster sizes to support your anticipated metadata results:

Small:
- 3 dedicated master nodes; t2.small.elasticsearch instance type
- 2 data nodes; m4.large.elasticsearch instance type

Medium:
- 3 dedicated master nodes; t2.medium.elasticsearch instance type
- 4 data nodes; m4.large.elasticsearch instance type

Large:
- 3 dedicated master nodes; t2.medium.elasticsearch instance type
- 6 data nodes; m4.large.elasticsearch instance type

**Regional Deployment**
This solution uses the Amazon Rekognition, Amazon Comprehend, and Amazon Transcribe services which are currently available in specific AWS Regions only. Therefore, you must
launch this solution in an AWS Region where the AWS Services are available. For the most current availability by region, see the [AWS service offerings by region](#).

## AWS CloudFormation Template

This solution uses AWS CloudFormation to automate the deployment of the Media Analysis Solution on the AWS Cloud. It includes the following AWS CloudFormation template which you can download before deployment:

**media-analysis-deploy.template**: Use this template to launch the solution and all associated components. The default configuration deploys Amazon Simple Storage Service (Amazon S3) buckets, an Amazon ElasticSearch Service (Amazon ES) cluster, AWS Lambda functions, an AWS Step Functions state machine, an Amazon API Gateway RESTful API, an Amazon CloudFront distribution, and an Amazon Cognito user pool, but you can also customize the template based on your specific needs.

## Automated Deployment

Before you launch the automated deployment, please review the architecture, configuration, and other considerations discussed in this guide. Follow the step-by-step instructions in this section to configure and deploy the Media Analysis Solution into your account.

**Time to deploy**: Approximately 20 minutes

## What We’ll Cover

The procedure for deploying this architecture on AWS consists of the following steps. For detailed instructions, follow the links for each step.

**Step 1. Launch the stack**

- Launch the AWS CloudFormation template into your AWS account.
- Enter values for required parameters: **Email**, **Elasticsearch Cluster Sizing**

**Step 2. Upload a File for Processing**

- Upload a file using the web interface to begin analysis

**Step 3. View Extracted Metadata**

- Review the results extracted from the previously uploaded file
Step 1. Launch the Stack
This automated AWS CloudFormation template deploys the Media Analysis Solution on the AWS Cloud.

Note: You are responsible for the cost of the AWS services used while running this solution. See the Cost section for more details. For full details, see the pricing webpage for each AWS service you will be using in this solution.

1. Sign in to the AWS Management Console and click the button to the right to launch the media-analysis-deploy AWS CloudFormation template. You can also download the template as a starting point for your own implementation.

2. The template is launched in the US East (N. Virginia) Region by default. To launch the solution in a different AWS Region, use the region selector in the console navigation bar.

Note: This solution uses Amazon Rekognition, Amazon Comprehend, and Amazon Transcribe, which are currently available in specific AWS Regions only. Therefore, you must launch this solution in an AWS Region where these services are available.

3. On the Select Template page, keep the default settings for Choose a template and select Next.

4. On the Specify Details page, assign a name to your solution stack.

5. Under Parameters, review the parameters for the template and modify them as necessary. This solution uses the following default values.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Default</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Email</td>
<td>&lt;Requires Input&gt;</td>
<td>Email address of the user that will be created in the Amazon Cognito Identity Pool. After launch, an email will be sent to this address with instructions for logging in to the demo website</td>
</tr>
<tr>
<td>Elasticsearch Cluster Sizing</td>
<td>Small</td>
<td>A drop-down box with three options: small, medium, large cluster</td>
</tr>
<tr>
<td>Send Anonymous Usage Data</td>
<td>Yes</td>
<td>A drop-down box indicating whether you would like to send anonymous usage data to AWS</td>
</tr>
</tbody>
</table>

6. Choose Next.

7. On the Options page, choose Next.

---

1 For the most current AWS services availability by region, see [https://aws.amazon.com/about-aws/global-infrastructure/regional-product-services/](https://aws.amazon.com/about-aws/global-infrastructure/regional-product-services/)
8. On the **Review** page, review and confirm the settings. Be sure to check the box acknowledging that the template will create AWS Identity and Access Management (IAM) resources.

9. Choose **Create** to deploy the stack.

   You can view the status of the stack in the AWS CloudFormation console in the **Status** column. You should see a status of **CREATE_COMPLETE** in approximately 20 minutes.

   **Note:** In addition to the primary AWS Lambda functions, this solution includes the MediaAnalysisHelperFunction Lambda function, which runs only during initial configuration or when resources are updated or deleted.

   When running this solution, the MediaAnalysisHelperFunction function is inactive. However, do not delete this function as it is necessary to manage associated resources.

**Step 2. Upload a File for Processing**

After the solution successfully launches, you can start uploading video, image, or audio files for processing. The solution sends an email with instructions for signing into the solution’s provided web interface.

1. In the email, follow the instructions to sign in to the website.

2. Navigate to the **Upload** tab.

3. In the **Media** section, upload an image, video, or audio file.

4. Select **View in Console**, to track the progress of your media analysis in the AWS Step Functions console.

**Step 3. View Extracted Metadata**

Once the media file metadata is extracted, you can view the results by selecting **View results**, or searching for the file in the **Browse** tab. Depending on the media file uploaded, you will see different types of extracted metadata. You can then view detected labels, facial attributes, known faces from provided images, celebrities, persons, captions, entities, and phrases in your uploaded files.

**Security**

When you build systems on AWS infrastructure, security responsibilities are shared between you and AWS. This shared model can reduce your operational burden as AWS operates, manages, and controls the components from the host operating system and virtualization
layer down to the physical security of the facilities in which the services operate. For more information about security on AWS, visit the [AWS Security Center](https://aws.amazon.com/security/).

**Amazon S3 Server-Side Encryption (SSE)**

AWS highly recommends that customers encrypt sensitive data in transit and at rest. This solution automatically encrypts media files and metadata at rest with [Amazon Simple Storage Service (Amazon S3) SSE](https://aws.amazon.com/s3/encryption/).

**Amazon CloudFront**

This solution deploys a static website hosted in an Amazon S3 bucket. To help reduce latency and improve security, this solution includes an Amazon CloudFront distribution with an origin access identity, which is a special CloudFront user that helps restrict access to the solution’s website bucket contents. For more information, see [Restricting Access to Amazon S3 Content by Using an Origin Access Identity](https://docs.aws.amazon.com/AmazonCloudFront/latest/DeveloperGuide/origin-access-identity.html).

**Additional Resources**

**AWS services**

- [AWS CloudFormation](https://aws.amazon.com/cloudformation/)
- [AWS Lambda](https://aws.amazon.com/lambda/)
- [Amazon Rekognition](https://aws.amazon.com/rekognition/)
- [Amazon Cognito](https://aws.amazon.com/cognito/)
- [Amazon Simple Storage Service](https://aws.amazon.com/s3/)
- [Amazon Transcribe](https://aws.amazon.com/transcribe/)
- [Amazon Comprehend](https://aws.amazon.com/comprehend/)
- [Amazon API Gateway](https://aws.amazon.com/api-gateway/)
- [Amazon Elasticsearch Service](https://aws.amazon.com/elasticsearch-service/)
- [AWS Step Functions](https://aws.amazon.com/stepfunctions/)
- [Amazon CloudFront](https://aws.amazon.com/cloudfront/)
Appendix A: Media Analysis Solution API

The Media Analysis Solution API enables you to access media metadata in a secure manner. The API acts as a “front door” to an AWS Lambda function that applications use to search and retrieve metadata extracted from media files.

The following operations are available in the Media Analysis Solution API:

**GET /details/{object_id}**

**Description**
The GET /details/{object_id} operation enables you to retrieve details about a specific media file that has been uploaded and analyzed by the Media Analysis Solution.

**Request Parameter**

`object_id`

- The ID of the media file
- Type: String
- Required: Yes

**Response**

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>object_id</td>
<td>The unique ID of the media file</td>
</tr>
<tr>
<td>details</td>
<td>A mapping containing details about the media file</td>
</tr>
</tbody>
</table>

**GET /lookup/{object_id}/{lookup_type}?page={page}**

**Description**
The GET /lookup/{object_id}/{lookup_type}?page={page} operation enables you to retrieve metadata that has been extracted by the media file by the solution.

**Request Parameter**

`object_id`

- The ID of the media file
- Type: String
Required: Yes

**lookup_type**

The type of metadata requested

Type: String

Required: Yes

Valid Values: labels | celebs | faces | face_matches | persons | transcript | phrases | entities

**page**

The page number to return

Type: Number

Required: No

### Response

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>S3</td>
<td>Amazon S3 bucket and key of the media file</td>
</tr>
<tr>
<td>&lt;LookupType&gt;</td>
<td>Details for each &lt;lookup_type&gt; detected by the Media Analysis Solution</td>
</tr>
<tr>
<td>MediaType</td>
<td>The format of the media file (image and video only)</td>
</tr>
<tr>
<td>Next</td>
<td>If the response is truncated, the Media Analysis Solution returns this number that you can use in the subsequent request to retrieve the next set of &lt;lookup_type&gt;</td>
</tr>
</tbody>
</table>

**GET /search?searchterm={search_term}&page={page}**

**Description**

The **GET /search?searchterm={search_term}&page={page}** operation enables you to search for media files based on metadata that has been extracted by the solution. Up to 30 results are returned at a time.

**Request Parameter**

**search_term**

Term to search for
Type: String
Required: Yes

Page number to request
Type: Number
Required: Yes
Minimum Value: 1

Response

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>name</td>
<td>The name of the media file</td>
</tr>
<tr>
<td>file_type</td>
<td>The format of the media file</td>
</tr>
<tr>
<td>media_id</td>
<td>The unique ID of the media file</td>
</tr>
<tr>
<td>thumbnail</td>
<td>The location of the media file in Amazon S3</td>
</tr>
<tr>
<td>upload_time</td>
<td>The date and time that the media file was uploaded to the solution</td>
</tr>
</tbody>
</table>

GET /status/{object_id}

Description
The GET /status/{object_id} operation enables you to retrieve the status of the state machine analyzing the media file.

Request Parameter
object_id

The ID of the media file
Type: String
Required: Yes
### Response

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>state_machine_status</td>
<td>The status of the state machine analyzing the media file</td>
</tr>
<tr>
<td>analysis</td>
<td>The status of each individual analysis being performed by the state machine</td>
</tr>
</tbody>
</table>

### Appendix B: Media Analysis State Machine

The Media Analysis Solution coordinates the analysis of media files using an AWS Step Functions state machine that triggers an AWS Lambda function to orchestrate the analysis and extraction of metadata using managed Artificial Intelligence (AI) services. When a new media file is uploaded to Amazon Simple Storage Service (Amazon S3), a Lambda function is invoked that parses the event details and starts the state machine using the following input:

```json
{
    "Records": [
        {
            "eventSource": "media-analysis"
        }
    ],
    "upload_time": "2018-05-16T10:00:00.000Z",
    "key": "private/<Cognito-Identity-Id>/media/<object-id>/content/filename.ext",
    "file_type": "ext",
    "size": 50000,
    "owner_id": "<Cognito-Identity-Id>",
    "object_id": "<object-id>",
    "file_name": "image_name.ext"
}
```

Depending on the format of the file uploaded, the state machine will execute image, video, and/or audio analysis.

The solution leverages Pass states to inform Task states of the type of analysis that should be performed by the Lambda function. The Pass state simply passes its input to its output. Each Pass state updates the $.lambda field in the state machine output with parameters for the Lambda function. When invoked, the Lambda function consumes this output, and depending on the $.lambda.service_name and $.lambda.function_name, it will execute the specified analysis. The code snippet below coordinates label detection in the image analysis state machine:
Images
If a file_type is PNG, JPG, or JPEG, the state machine enters a parallel state that automatically coordinates the analysis of the image using the synchronous Amazon Rekognition API.

Videos
If a file_type is MOV or MP4 the state machine enters a parallel state that automatically coordinates the analysis of the video using the asynchronous Amazon Rekognition Video API.
Figure 3: Video analysis state machine

Each branch of the video analysis state machine starts the analysis job, then uses the job status poller pattern to check on the status of the video analysis job. Once the analysis is complete, the state machine retrieves, processes, and stores the results.

Audio

If a file_type is MP3, MP4, WAV, or FLAC, the state machine automatically coordinates the analysis of the audio using the asynchronous Amazon Transcribe API.
The audio analysis state machine starts the analysis job, then uses the `job status poller` pattern to check on the status of the audio analysis job. Once the analysis is complete, the results are retrieved, processed, and stored and the state machine enters a parallel state that automatically coordinates the detection of key entities and phrases in the resulting transcript using the asynchronous Amazon Comprehend API. The results are immediately processed and stored.

**Appendix C: Customizing the Solution**

The Media Analysis Solution includes a default set of features, but you can extend and customize the solution to include additional features and functionalities for your specific use
case. Follow the step-by-step instructions in this section to add functionality that will automatically detect text in images uploaded to the solution.

**Step 1. Update the State Machine**

To coordinate text detection in concurrence with the existing image analysis, add another branch to the image analysis state machine. Follow the example below:

```json
{
  "StartAt": "Image-Text Params",
  "States": {
    "Image-Text Params": {
      "Type": "Pass",
      "Result": {
        "service_name": "image",
        "function_name": "get_text"
      },
      "ResultPath": "$lambda",
      "Next": "Image-Get Text"
    },
    "Image-Get Text": {
      "Type": "Task",
      "InputPath": "$",
      "ResultPath": "$results.text",
      "End": true
    }
  }
}
```

**Step 2. Add Logic to AWS Lambda Code**

When the MediaAnalysisFunction AWS Lambda function is invoked by the state machine, the function will parse the event, invoke the appropriate Amazon Rekognition API, process and store the response, and return the status of the job to the state machine. Update the MediaAnalysisFunction Lambda function.

In the `source/analysis/lib/image/index.js` file, add the following code.

```javascript
else if (event.lambda.function_name == 'get_text') {
  _image.getText(event, function(err, data) {
    if (err) {
      return cb(err, null);
    } else {
      return cb(null, data);
    }
  });
}
```
In the source/analysis/lib/image/image.js file, add the following code.

```javascript
image.prototype.getText = function(image_info, cb) {
    let rek_params = {
        Image: {
            S3Object: {
                Bucket: s3Bucket,
                Name: image_info.key
            }
        }
    };

    let rekognition = new AWS.Rekognition();
    rekognition.detectText(rek_params, function(err, data) {
        if (err) {
            return cb(err, null);
        } else {

            var text = [];
            for (var t in data.TextDetections) {
                if (data.TextDetections[t].Confidence >= confidence_score) {
                    text.push(data.TextDetections[t].DetectedText);
                }
            }

            let text_key = ['private', image_info.owner_id, 'media', image_info.object_id, 'results', 'text.json'].join('/');

            let s3_params = {
                Bucket: s3Bucket,
                Key: text_key,
                Body: JSON.stringify(data),
                ContentType: 'application/json'
            };

            upload.respond(s3_params, function(err, response) {
                if (err) {
                    return cb(err, null);
                } else {
                    let text_response = {'key': text_key, 'text': text, 'status': 'COMPLETE'};
                    return cb(null, text_response);
                }
            });
        }
    });
};
```
Step 3. Update the AWS IAM Policy

The Media Analysis Solution uses AWS Identity and Access Management (IAM) policies to ensure least privilege access to AWS resources. By default, the solution does not allow the MediaAnalysisFunction Lambda function to access the Amazon Rekognition DetectText API, if you want to access the API you must update the IAM policy to provide access. Add the following permission:

```json
{
    "Action": [
        "rekognition:DetectText"
    ],
    "Resource": "*",
    "Effect": "Allow"
}
```

You can quickly and easily add new functionality to the Media Analysis Solution, enabling you to extract even more valuable data from your media files. In addition to the above changes, you should consider updating the MediaAnalysisApiFunction Lambda function to enable API access to the metadata extracted from your media files.

Customers who want to write their own code using a different language or logic structure can easily coordinate those functions from the Media Analysis Solution state machine. You can also use the solution to coordinate analysis using your own machine learning models by specifying an activity in the state machine’s Task resource.
Appendix D: Collection of Anonymous Data

This solution includes an option to send anonymous usage data to AWS. We use this data to better understand how customers use this solution and related services and products. When enabled, the following information is collected and sent to AWS:

- **Solution ID**: The AWS solution identifier
- **Unique ID (UUID)**: Randomly generated, unique identifier for each deployment
- **Timestamp**: Media file upload timestamp
- **Format**: The format of the uploaded media file
- **Size**: The size of the file the solution processes
- **Duration**: The length of the uploaded video file

Note that AWS will own the data gathered via this survey. Data collection will be subject to the [AWS Privacy Policy](https://aws.amazon.com/about-aws/privacy-legal/). To opt out of this feature, modify the AWS CloudFormation template mapping section as follows:

```
"Send" : {
  "AnonymousUsage" : { "Data" : "Yes" }
},
```

OR

```
"Send" : {
  "AnonymousUsage" : { "Data" : "No" }
},
```

b) After the solution has been launched, find the `media-analysis-solution` function in the Lambda console and set the `SEND_ANONYMOUS_DATA` environment variable to `No`. 
Source Code

You can visit our GitHub repository to download the templates and scripts for this solution, and to share your customizations with others.

Document Revisions

<table>
<thead>
<tr>
<th>Date</th>
<th>Change</th>
<th>In sections</th>
</tr>
</thead>
<tbody>
<tr>
<td>May 2018</td>
<td>Initial Release</td>
<td>-</td>
</tr>
<tr>
<td>September 2018</td>
<td>Added information about media conversion, the increased size limitation for audio files, and custom sizing for the Amazon ES cluster</td>
<td>Solution Components; Considerations</td>
</tr>
<tr>
<td>December 2018</td>
<td>Added information about the Amazon CloudFront distribution for the static website hosted in the Amazon S3 bucket</td>
<td>Security</td>
</tr>
</tbody>
</table>

© 2018, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Notices

This document is provided for informational purposes only. It represents AWS’s current product offerings and practices as of the date of issue of this document, which are subject to change without notice. Customers are responsible for making their own independent assessment of the information in this document and any use of AWS’s products or services, each of which is provided “as is” without warranty of any kind, whether express or implied. This document does not create any warranties, representations, contractual commitments, conditions or assurances from AWS, its affiliates, suppliers or licensors. The responsibilities and liabilities of AWS to its customers are controlled by AWS agreements, and this document is not part of, nor does it modify, any agreement between AWS and its customers.

The Media Analysis Solution on AWS is licensed under the terms of the Amazon Software License available at https://aws.amazon.com/asl/.