
AWS Media Insights Engine Implementation Guide



AWS Media Insights Engine: Implementation Guide

Copyright © Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Amazon's trademarks and trade dress may not be used in connection with any product or service that is not Amazon's, in any manner that is likely to cause confusion among customers, or in any manner that disparages or discredits Amazon. All other trademarks not owned by Amazon are the property of their respective owners, who may or may not be affiliated with, connected to, or sponsored by Amazon.

Table of Contents

Welcome	1
Cost	2
Table 1: Cost per month for 100 workflow runs	2
Architecture overview	3
Components	5
Workflow API	5
Control plane	5
Amazon DynamoDB	5
Operators	5
Data plane	7
Data plane API	7
Data plane pipeline	7
Data pipeline consumers	7
Security	8
IAM roles	8
Design considerations	9
Regional deployments	9
Limitations	9
Template	10
Deployment	11
Launch the stack	11
Resources	13
Uninstall the solution	14
Using the AWS Management Console	14
Using AWS Command Line Interface	14
Deleting the Amazon S3 buckets	14
Operational metrics	15
Source code	16
Revisions	17
Contributors	18
Notices	19

Development framework for building applications that process multimedia

Publication date: *April 2021 (last update (p. 17): January 2022)*

Amazon Web Services (AWS) Media Insights Engine (MIE) is a development framework for building applications that process videos, images, audio, and text with machine learning services on AWS. MIE takes care of workflow orchestration and data persistence so that you can focus on workflow development. By addressing the concerns of running workflows, MIE empowers you to build applications faster with the benefit of inheriting a pre-built and robust back end.

MIE has been successfully used in a variety of scenarios, such as:

- Deriving video features for ad placement
- Transforming video content with redaction
- Indexing videos based on visual and audio content
- Translating videos for automated localization

For additional details and sample use cases, refer to [How to rapidly prototype multimedia applications on AWS with the Media Insights Engine](#) on the AWS Media Blog.

This solution does not provide a graphical user interface (GUI); however, the AWS Media Insights [GitHub repository](#) includes a Media Insights Engine reference application that features a GUI. Consider using this for your own implementation.

This implementation guide describes architectural considerations and configuration steps for deploying AWS Media Insights Engine in the 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 developers who have practical experience architecting in the AWS Cloud. If you are not attempting to develop an application or service that processes videos, images, audio, or text then we recommend you implement another [solutions implementation](#) that better meets your needs.

Cost

You are responsible for the cost of the AWS services used while running this solution. As of January 2022, the cost for running this solution with the default settings in the US East (N. Virginia) Region is approximately **\$24 per month without free tiers**, or **\$13 per month with free tiers** for 100 workflow runs. Add approximately \$2.40 for each additional 100 workflow runs. Most MIE use cases are covered by the free tier for all AWS services except Amazon Kinesis and AWS Lambda.

Table 1: Cost per month for 100 workflow runs

This solution uses the following resources that are billed on a monthly basis.

AWS service	Quantity	Cost
Amazon API Gateway	1 million workflows	\$ 3.50
Amazon DynamoDB	1 million workflows	\$ 0.25
AWS Lambda	100 workflows	\$ 4.75
Amazon Kinesis	100 workflows	\$12.56
Amazon SQS	1 million workflows	\$ 0.40
Amazon SNS	1 million workflows	No charge
Amazon S3	100 workflows	\$ 2.30
AWS X-Ray	100 workflows	\$ 0.0005
Total monthly cost:		\$23.76

The costs in Table 1 are for the AWS Media Insights Engine framework only. User-defined workflows might call other services, such as Amazon Rekognition, which are not included in this estimate. Prices are subject to change. For full details, refer to the pricing webpage for each AWS service you will be using in this solution.

Architecture overview

Deploying this solution with the default parameters builds the following environment in the AWS Cloud.

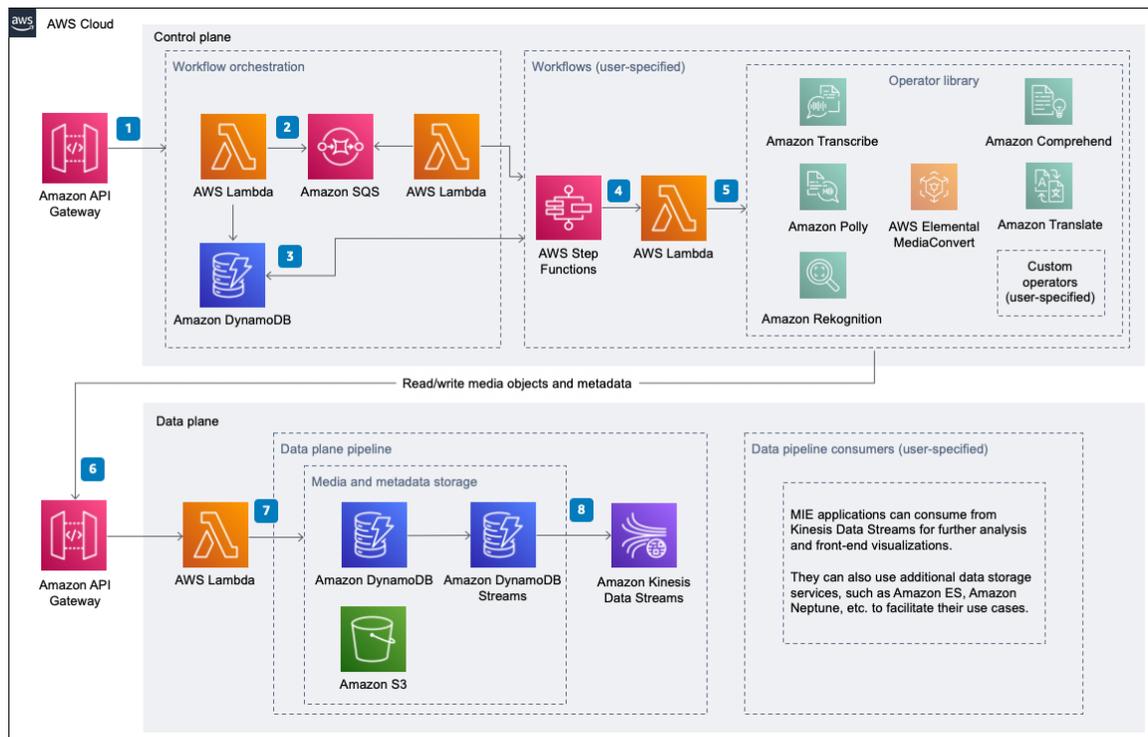


Figure 1: AWS Media Insights Engine architecture on AWS

The AWS CloudFormation template provisions the following resources:

1. An [Amazon API Gateway](#) resource for the control plane REST API. This resource is the entry point where requests to create, read, update, delete (CRUD), or run workflows begin.
2. [AWS Lambda](#) and [Amazon Simple Queue Service \(Amazon SQS\)](#) resources to support workflow orchestration and translating user-defined workflows into [AWS Step Functions](#). A Lambda function updates workflow-related tables in DynamoDB, at which point CRUD workflow requests finish. Requests to run workflows begin when another Lambda function saves the request to an SQS queue. The queue is later read and run by the `WorkflowScheduler` Lambda function that controls how many workflows can run at the same time.
3. [Amazon DynamoDB](#) tables to store workflow-related data, such as state machine definitions for operators, workflow configurations, and workflow run status.
4. A Step Functions resource that is created when a user defines a new workflow using the workflow API. When the `WorkflowScheduler` Lambda function starts a workflow, it starts the Step Functions resource, which then invokes a series of Lambda functions that call external services and/or download results from those services. When all of the Lambda functions in a workflow have fully run, then a Lambda function is called to update the workflow status in Amazon DynamoDB.
5. Lambda functions for using the following commonly used services in MIE workflows: [Amazon Rekognition](#), [Amazon Comprehend](#), [Amazon Translate](#), [Amazon Transcribe](#), [Amazon Polly](#), and [AWS Elemental MediaConvert](#). Operators consist of Lambda functions that call external services and/or download results from those services. They are invoked by a state machine in Step Functions, as

prescribed by the workflow definition. These Lambda functions save results to long-term storage via the data plane REST API.

6. An Amazon API Gateway resource for the data plane REST API. Operators save results to long-term storage by calling this API.
7. [Amazon Simple Storage Service](#) (Amazon S3), DynamoDB, and DynamoDB Streams for media and metadata data storage. The Lambda function behind the data plane API directly accesses Amazon S3 and Amazon DynamoDB to perform incoming CRUD requests. That Lambda function saves files, such as binary media files or JSON metadata files, in Amazon S3. A pointer to those files is saved in an Amazon DynamoDB table. Finally, a time-ordered sequence of modifications to that table are saved in an Amazon DynamoDB Stream and an Amazon Kinesis Data Streams resource.
8. [Amazon Kinesis Data Streams](#) for interfacing with external applications. A Kinesis data stream provides an interface for external applications to access data stored in the MIE data plane. This interface is appropriate for feeding downstream data stores, such as the Amazon OpenSearch Service (OpenSearch Service) or Amazon Neptune, that support specialized data access patterns required by end-user applications. In order to feed a downstream data store, you must implement a consumer (for example, an AWS Lambda function) that consumes records from the data stream and performs the necessary extract, transform, and load (ETL) tasks needed for the external application.

Note

The ETL tasks that feed downstream data stores are entirely use-case dependent and therefore must be user-defined. The [Media Insights Engine Developer Guide](#) includes detailed instructions for implementing ETL functions in MIE.

Solution components

Workflow API

This API creates, updates, deletes, runs, and monitors workflows.

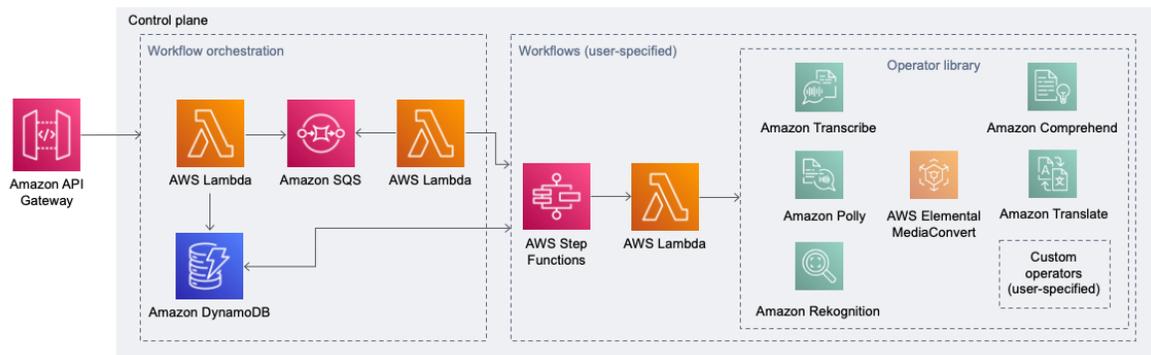


Figure 2: AWS Media Insights Engine control plane

Control plane

This includes the workflow API and state machines for workflows. Workflow state machines are composed of operators from the MIE operator library. When operators within the state machine are run, they interact with the MIE data plane to store and retrieve derived asset and metadata generated from the workflow.

Use the control plane to create, read, update, and delete (CRUD) custom operators and workflows, and to execute those workflows.

Amazon DynamoDB

The following Amazon DynamoDB tables store workflow-related data:

- **Workflow** – This table records user-defined workflows.
- **Workflow Execution** – This table records the details of every workflow run.
- **Operations** – This table records details for each operator in the operator library, such as references to Lambda functions and default runtime parameters.
- **Stage** – This table records the auto-generated AWS Step Functions code needed for each operator.
- **System** – This table records system-wide configurations, such as maximum concurrent workflows.

Operators

Operators are generated state machines that call AWS Lambda functions to perform media analysis or media transformation tasks. Users can define custom operators, but the MIE operator library includes the following pre-built operators:

- **Celebrity Recognition** - An asynchronous operator to identify celebrities in a video using Amazon Rekognition.
- **Content Moderation** - An asynchronous operator to identify unsafe content in videos using Amazon Rekognition.
- **Face Detection** - An asynchronous operator to identify faces in videos using Amazon Rekognition.
- **Face Search** - An asynchronous operator to identify faces from a custom face collection in videos using Amazon Rekognition.
- **Label Detection** - An asynchronous operator to identify objects in a video using Amazon Rekognition.
- **Person Tracking** - An asynchronous operator to identify people in a video using Amazon Rekognition.
- **Shot Detection** - An asynchronous operator to identify camera shots in a video using Amazon Rekognition.
- **Text Detection** - An asynchronous operator to identify text in a video using Amazon Rekognition.
- **Technical Cue Detection** - An asynchronous operator to identify technical cues such as end credits, color bars, and black bars in a video using Amazon Rekognition.
- **Comprehend Key Phrases** - An asynchronous operator to find key phrases in text using Amazon Comprehend.
- **Comprehend Entities** - An asynchronous operator to find references to real-world objects, dates, and quantities in text using Amazon Comprehend.
- **Create SRT Captions** - A synchronous operator to generate SRT formatted caption files from a video transcript generated by Amazon Transcribe.
- **Create VTT Captions** - A synchronous operator to generate VTT formatted caption files from a video transcript generated by Amazon Transcribe.
- **Media Convert** - An asynchronous operator to transcode input video into mpeg4 format using AWS Elemental MediaConvert.
- **Media Info** - A synchronous operator to read technical tag data for video files.
- **Polly** - An asynchronous operator that turns input text into speech using Amazon Polly.
- **Thumbnail** - An asynchronous operator that generates thumbnail images for an input video file using AWS Elemental MediaConvert.
- **Transcribe** - An asynchronous operator to convert input audio to text using Amazon Transcribe.
- **Translate** - An asynchronous operator to translate input text using Amazon Translate.

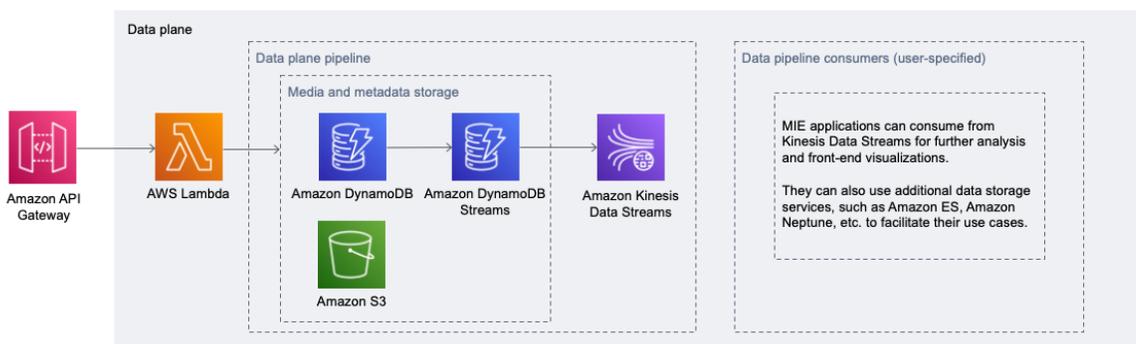


Figure 3: AWS Media Insights Engine data plane

Data plane

This stores the media assets and metadata generated by workflows. Implement a consumer of the Kinesis data stream in the data plane to extract, transform, and load (ETL) data from the master MIE data store to downstream databases that support the data access patterns required by end-user applications.

Data plane API

This API creates, updates, deletes, and retrieves media assets and metadata.

Data plane pipeline

This pipeline stores metadata for an asset that can be retrieved using an object's `AssetId` and `Metadata` type. Writing data to the pipeline initiates a copy of the data to be stored in Kinesis Data Streams. This data stream is the interface that end-user applications can connect to use data stored in the MIE data plane.

Data pipeline consumers

Changes to the data plane DynamoDB table are reflected in a Kinesis data stream. For each record in that data stream, data pipeline consumers perform the necessary extract, transform, and load (ETL) tasks needed to replicate data, such as media metadata, to the data stores used by external applications. These ETL tasks are entirely use-case dependent and therefore must be user-defined. The [Media Insights Engine Developer Guide](#) includes detailed instructions for implementing data pipeline consumers.

Security

When you build systems on AWS infrastructure, security responsibilities are shared between you and AWS. This [shared model](#) reduces your operational burden because AWS operates, manages, and controls the components including the host operating system, the virtualization layer, and the physical security of the facilities in which the services operate. For more information about AWS security, visit [AWS Cloud Security](#).

AWS Media Insights Engine (MIE) uses AWS_IAM to authorize REST API requests. Refer to the documentation for your chosen HTTP client to learn how to use AWS_IAM in your application.

IAM roles

AWS Identity and Access Management (IAM) roles allow customers to assign granular access policies and permissions to services and users on the AWS Cloud. This solution creates IAM roles that grant the solution's AWS Lambda functions access to create Regional resources.

Design considerations

Regional deployments

This solution can use operators for certain AWS services, which is not currently available in all AWS Regions. You must launch this solution in an AWS Region where the AWS services are available. For the most current availability by Region, refer to the [AWS Regional Services List](#).

Limitations

MIE does not inherently limit media attributes such as file size or video duration. Those limitations depend on the services used in user-defined workflows. For example, if a workflow uses Amazon Rekognition, then that workflow will be subject to the [guidelines and quotas](#) listed in the *Amazon Rekognition Developer Guide*.

AWS CloudFormation template

To automate deployment, this solution uses AWS CloudFormation. It includes the following CloudFormation template, which you can download before deployment:

[View
Template](#)

media-insights-stack.template: Use this template to launch the solution and all associated components. The default configuration deploys Amazon API Gateway, AWS Lambda, Amazon Simple Notification Service (Amazon SNS), Amazon Simple Queue Service (Amazon SQS), Amazon DynamoDB, AWS Step Functions, Amazon Simple Storage Service (Amazon S3), Amazon Kinesis Data Streams, and AWS Identity and Access Management (IAM). You can customize the template to meet your specific needs.

Automated deployment

Before you launch the solution, review the architecture, components, and network security considerations discussed in this guide. Follow the step-by-step instructions in this section to configure and deploy the solution into your account.

Time to deploy: Approximately 10 minutes

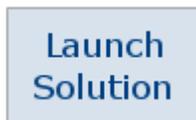
Launch the stack

This automated AWS CloudFormation template deploys the Media Insights Engine solution in the AWS Cloud.

Note

You are responsible for the cost of the AWS services used while running this solution. For more details, visit the [Cost \(p. 2\)](#) section in this guide, and refer to the pricing webpage for each AWS service used in this solution.

1. Sign in to the AWS Management Console and select the button to launch the `media-insights-stack` AWS CloudFormation template.



Alternatively, you can [download the template](#) as a starting point for your own implementation.

2. The template launches 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.
3. On the **Create stack** page, verify that the correct template URL is in the **Amazon S3 URL** text box and choose **Next**.
4. On the **Specify stack details** page, assign a name that is less than 13 characters long to your solution stack. This ensures all the resources in the MIE stack remain under the maximum length allowed by CloudFormation. For information about naming character limitations, refer to [IAM and STS Limits](#) in the *AWS Identity and Access Management User Guide*.
5. Under **Parameters**, review the parameters for this solution template and modify them as necessary. This solution uses the following default values.

Parameter	Default	Description
Max Concurrent Workflows	5	Identifies the maximum number of workflows to run concurrently. When the maximum is reached, additional workflows are added to a wait queue. Recommended range is 2 to 5.

Parameter	Default	Description
Deploy Analytics Pipeline	true	Determines whether to deploy a metadata streaming pipeline that can be consumed by downstream analytics platforms. By default, this capability is activated when the solution is deployed. Set to <code>false</code> to deactivate this capability.
Deploy Test Resources	false	Determines whether to deploy test resources that contain Lambda functions required for integration and end-to-end testing. By default, this capability is deactivated when the solution is deployed. Set to <code>true</code> to activate this capability.
Turn on X-Ray Trace	false	Determines whether to activate Active Xray tracing on all entry points to the stack. By default, this capability is deactivated when the solution is deployed. Set to <code>true</code> to activate this capability.
External Bucket ARN	(none)	Specify the Amazon S3 ARN for media files that MIE needs permission to read as inputs to workflows. Leave blank if the MIE data plane bucket is the only source for media files.
Send Anonymous Data	true	Send anonymous data about MIE performance to AWS to help improve the quality of this solution.
Solution ID	S00163	AWS Solution ID used to track performance in anonymous data.

6. Choose **Next**.
7. On the **Configure stack options** page, choose **Next**.
8. On the **Review** page, review and confirm the settings. Check the box acknowledging that the template will create AWS Identity and Access Management (IAM) resources.
9. Choose **Create stack** to deploy the stack.

You can view the status of the stack in the AWS CloudFormation Console in the **Status** column. You should receive a `CREATE_COMPLETE` status in approximately ten minutes.

Additional resources

AWS services	
<ul style="list-style-type: none">• AWS CloudFormation	<ul style="list-style-type: none">• Amazon Transcribe
<ul style="list-style-type: none">• Amazon API Gateway	<ul style="list-style-type: none">• Amazon Polly
<ul style="list-style-type: none">• AWS Lambda	<ul style="list-style-type: none">• AWS Elemental MediaConvert
<ul style="list-style-type: none">• Amazon DynamoDB	<ul style="list-style-type: none">• AWS Step Functions
<ul style="list-style-type: none">• Amazon Simple Notification Service	<ul style="list-style-type: none">• Amazon Kinesis Data Streams
<ul style="list-style-type: none">• Amazon Simple Queue Service	<ul style="list-style-type: none">• Amazon Simple Storage Service
<ul style="list-style-type: none">• Amazon Rekognition	<ul style="list-style-type: none">• AWS Identity and Access Management (IAM)
<ul style="list-style-type: none">• Amazon Comprehend	<ul style="list-style-type: none">• AWS X-Ray
<ul style="list-style-type: none">• Amazon Translate	

Uninstall the solution

To uninstall the AWS Media Insights Engine solution, delete the AWS CloudFormation stack. This will delete all the resources created by the template except the `Dataplane` S3 bucket and the `DataplaneLogs` bucket. These two buckets will be retained when the solution stack is deleted in order to help prevent accidental data loss. You can use either the AWS Management Console or the AWS Command Line Interface (AWS CLI) to empty, then delete those S3 buckets after deleting the CloudFormation stack.

Using the AWS Management Console

1. Sign in to the [AWS CloudFormation](#) console.
2. Select this solution's installation stack.
3. Choose **Delete**.

Using AWS Command Line Interface

Determine whether the AWS Command Line Interface (AWS CLI) is available in your environment. For installation instructions, refer to [What Is the AWS Command Line Interface](#) in the *AWS CLI User Guide*. After confirming that the AWS CLI is available, run the following command.

```
$ aws cloudformation delete-stack --stack-name <installation-stack-name> --region <aws-region>
```

Deleting the Amazon S3 buckets

Media Insights Engine creates two S3 buckets that are not automatically deleted. To delete these buckets, use the steps below.

1. Sign in to the [Amazon S3 console](#).
2. Select the `Dataplane` bucket.
3. Choose **Empty**.
4. Choose **Delete**.
5. Select the `DataplaneLogs` bucket.
6. Choose **Empty**.
7. Choose **Delete**.

To delete the S3 bucket using AWS CLI, run the following command:

```
$ aws s3 rb s3://<bucket-name> --force
```

Collection of operational metrics

This solution includes an option to send anonymous operational metrics to AWS. We use this data to help improve the quality of the solution. When invoked, the following information is collected and sent to AWS:

- **Solution ID:** The AWS solution identifier
- **Unique ID (UUID):** Randomly generated, unique identifier for each AWS Media Insights Engine deployment
- **Timestamp:** Data-collection timestamp
- **Instance Data:** The version of the solution that was deployed

Example data:

```
{
  "Solution": "S00163",
  "UUID": "d84a0bd5-7483-494e-8ab1-fdfaa7e97687",
  "TimeStamp": "2021-03-01T20:03:05.798545",
  "Data": {
    "Version": "2.0.4",
    "CFTemplate": "Created"
  }
}
```

AWS owns the data gathered through this survey. Data collection is subject to the [AWS Privacy Policy](#).

To opt out of this feature, set the `SendAnonymousData` parameter in the solution's `media-insights-stack` AWS CloudFormation template to a value of `false`. For details, refer to [Launch the stack \(p. 11\)](#).

Source code

Visit our solution [GitHub repository](#) to download the source files for this solution and to share your customizations with others. Refer to the [README.md file](#) for additional information.

Revisions

Date	Change
April 2021	Initial release (v2.0.5)
July 2021	Release v3.0.1: Updated AWS CloudFormation template parameters for capturing operational metrics. For more information, refer to the CHANGELOG.md file in the GitHub repository.
August 2021	Release v3.0.2: bug fixes. For more information, refer to the CHANGELOG.md file in the GitHub repository.
November 2021	Release v3.0.4: bug fixes. For more information, refer to the CHANGELOG.md file in the GitHub repository.
January 2022	Release v4.0.0: Added support for custom language models in Amazon Transcribe and source language auto-detection in Amazon Translate. For more information, refer to the CHANGELOG.md file in the GitHub repository.

Contributors

- Alex Burkleaux
- Brandon Dold
- Ian Downard
- Jill Adams

Notices

Customers are responsible for making their own independent assessment of the information in this document. This document: (a) is for informational purposes only, (b) represents AWS current product offerings and practices, which are subject to change without notice, and (c) does not create any commitments or assurances from AWS and its affiliates, suppliers or licensors. AWS products or services are provided “as is” without warranties, representations, or conditions of any kind, whether express or implied. AWS responsibilities and liabilities 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.

AWS Media Insights Engine is licensed under the terms of the of the Apache License Version 2.0 available at the [Apache Software Foundation](#).