

SVC301

# Increase the value of video with machine learning & media services

Bryan Samis  
Solutions Architect  
Amazon Web Services

# Agenda

Why machine learning for media?

Use cases and relevant services

## AWS solutions

- Media analysis solution
- Live captioning and translation solution

# Why machine learning for media?

# Major shift in the Media and Entertainment industry

Content is no longer king

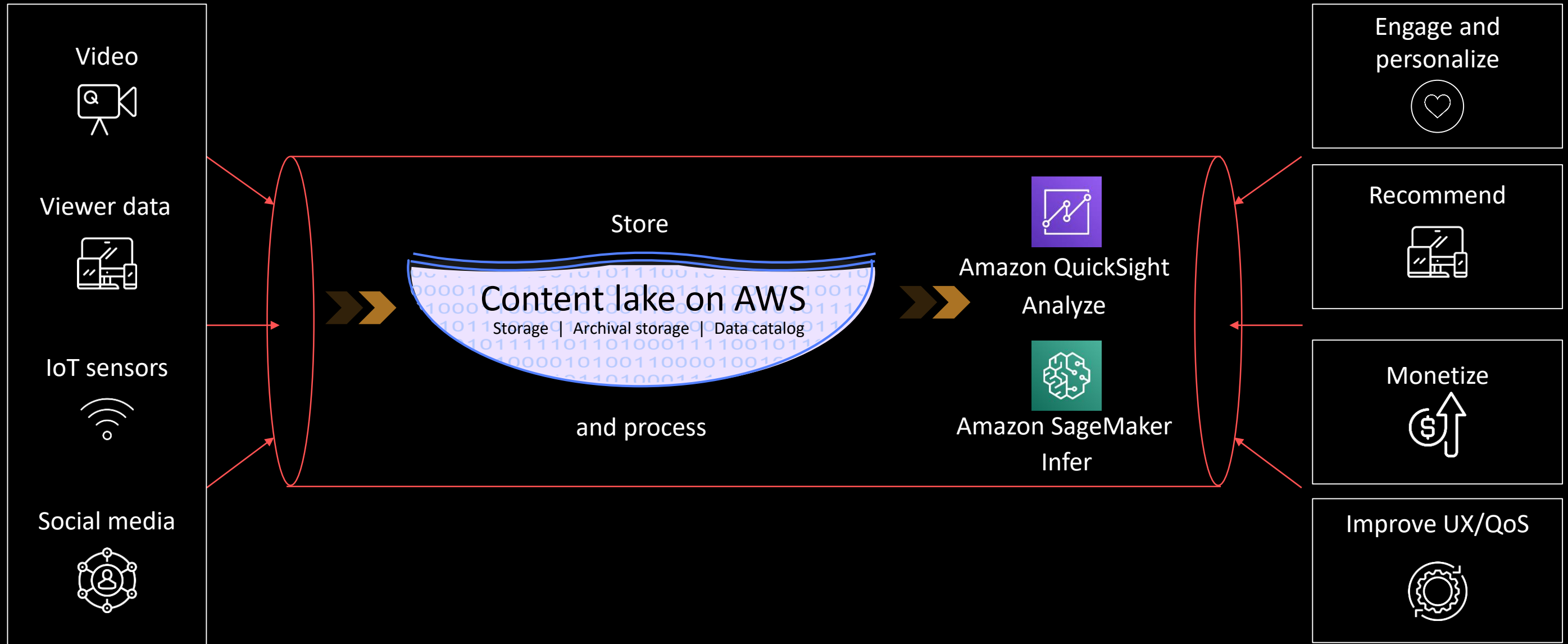
**The viewer is king**

- ✓ Desired shift to a direct-to-consumer business model
- ✓ Success relies on a rich viewer experience
- ✓ Rich viewer experiences rely on data and analytics



# Modern data architecture for media enrichment

## Insights to enhance viewer engagement, personalization, monetization



# Use cases



# Sample ML use cases for media

## Media metadata tagging



Automate the creation of rich metadata index (object, scene, activity, faces, etc.) extracted from audio visual content and integrated into digital and media asset management systems

AWS services to build use case: Amazon Rekognition, Amazon SageMaker, Amazon Transcribe, Amazon Comprehend

## Ad personalization



Identify objects and emotion in content that enable users to integrate personalized ads into subscriber video streams

AWS services to build use case: Apache Spark/Amazon EMR, Amazon SageMaker, AWS Elemental MediaTailor

## Closed captioning



Automated captions, transcription, and translation of audio content

AWS services to build use case: Amazon Transcribe, Amazon Translate, Amazon Comprehend

## Automated compliance marking



Detect potentially inappropriate content to avoid compliance issues in global markets and to increase brand safety for advertisers

AWS services to build use case: Amazon Rekognition, Amazon Transcribe, Amazon SageMaker

# Sample ML use cases for media

## Player/actor identification & activity pathing



Identify and track the path of player activity in the field, or identify the precise timecodes when actors enter and leave a scene, both visually and in spoken dialog (off-screen voice)

AWS services to build use case: Amazon Rekognition, Amazon Transcribe

## Language translation



Translate transcripts and metadata, and improve localization workflows and search experience

AWS services to build use case: Amazon Translate

## Content recommendations



Analyze disparate data and enable the consumer to make personalized content choices and the business to predict behavior

AWS services to build use case: Amazon SageMaker model, Amazon Personalize

## Automated redaction



Detect and pixelate faces captured incidentally to preserve the privacy of non-persons of interest in news feeds and security footage

AWS services to build use case: Amazon Rekognition, Amazon Transcribe, Amazon SageMaker



# Use case: From manual tagging to AI-based tagging



## Background

- **Over 200,000 hours of content**
- Only half the content is manually tagged
- How can they **enrich the metadata** on AWS?
- How can they **unleash the value of content** they already own once on AWS?

## Challenge

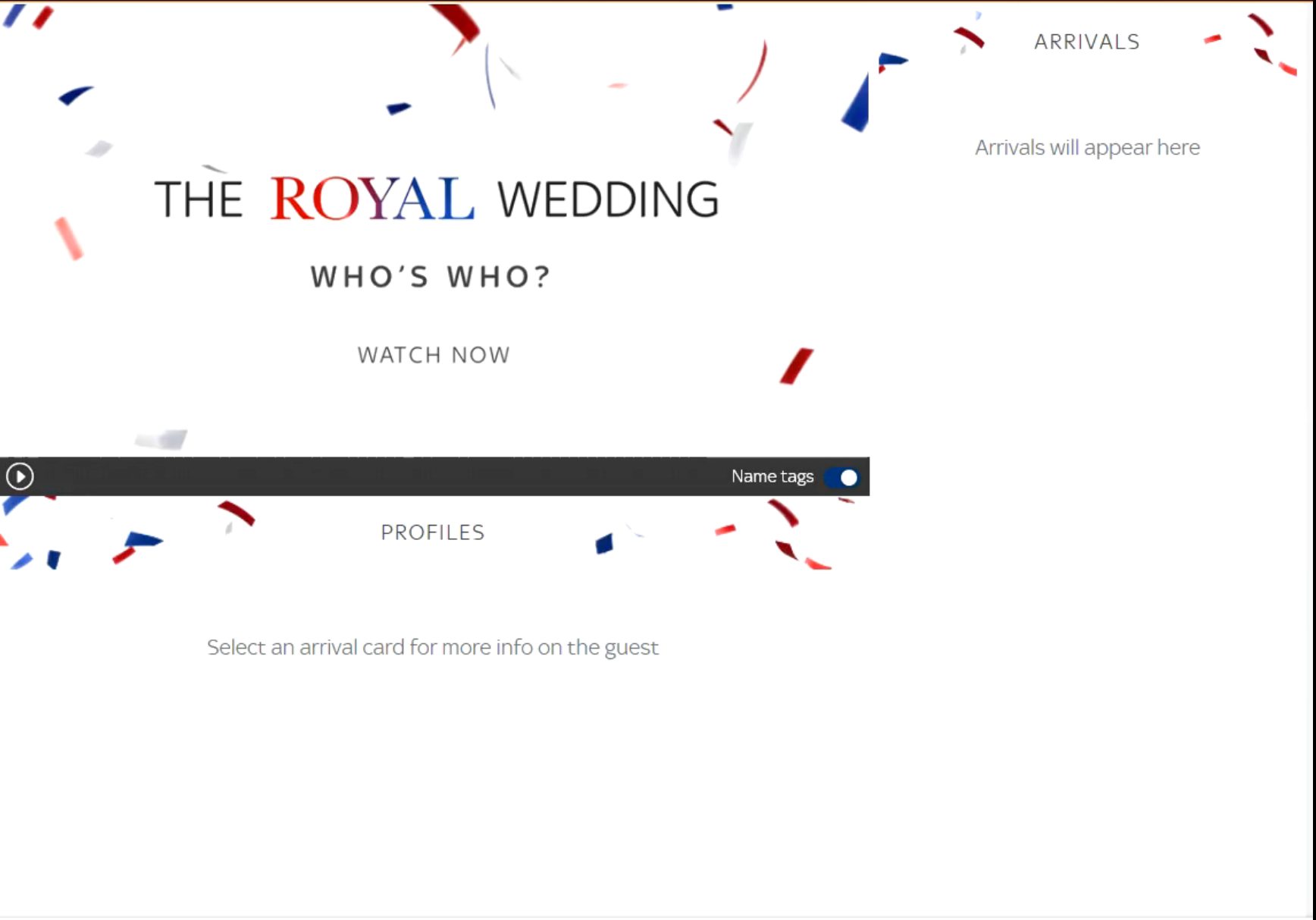
- Large-scale video library
- High accuracy required
- Limited budget
- Ability to extract from video
- Keep up with daily increase in content

## Results

- Solution developed within three weeks
- Live video frame-based analysis
- Established, searchable baseline archive
- All content is now tagged and indexed
- Over 99,000 faces indexed and searchable
- Saved ~9,000 hours a year in manual curation costs



# Use case: Event broadcasting enrichment



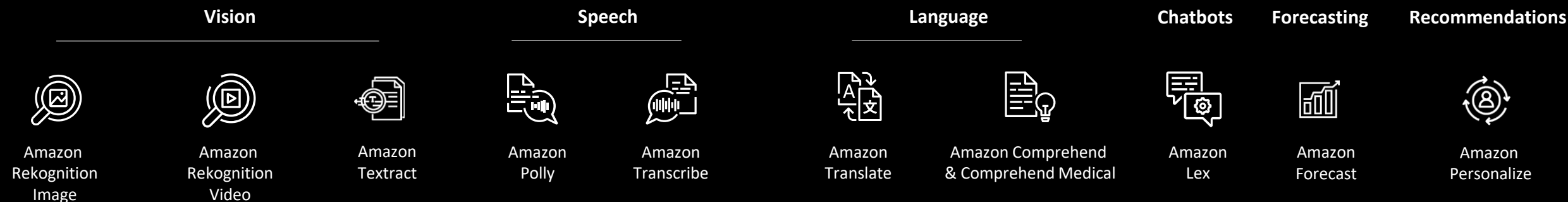
Sky News



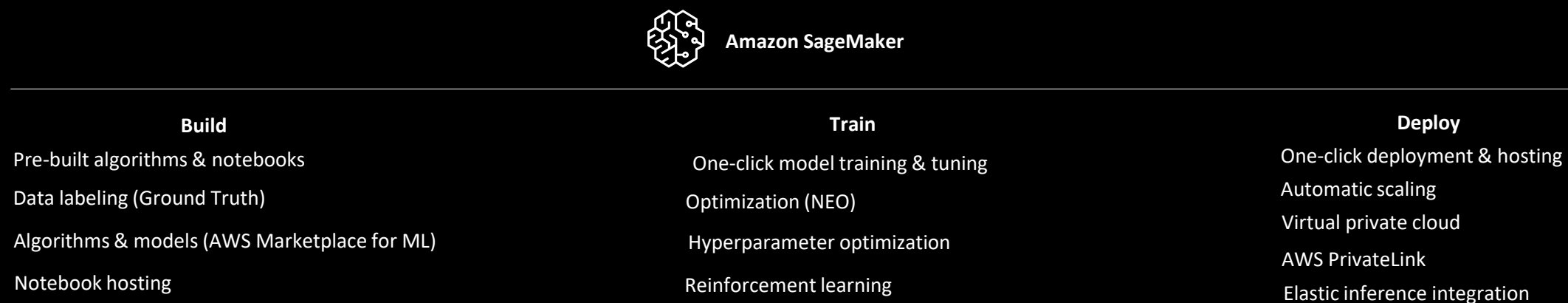
# AWS services

# The Amazon ML stack

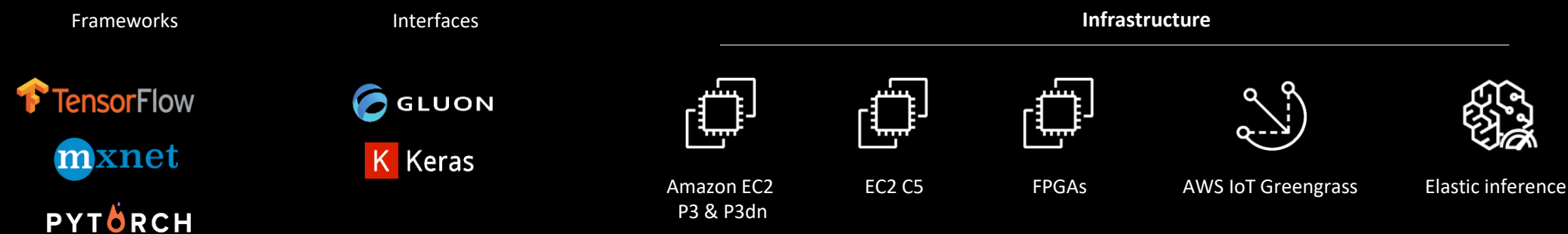
## AI services



## ML services

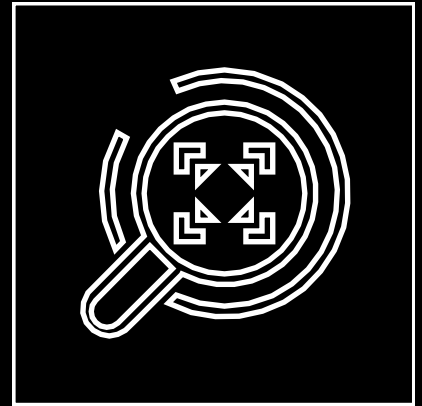


## ML frameworks & infrastructure



# Amazon Rekognition

*Deep learning-based image analysis service*



Object and  
scene detection



Facial  
analysis



Face search  
and match



Content  
moderation



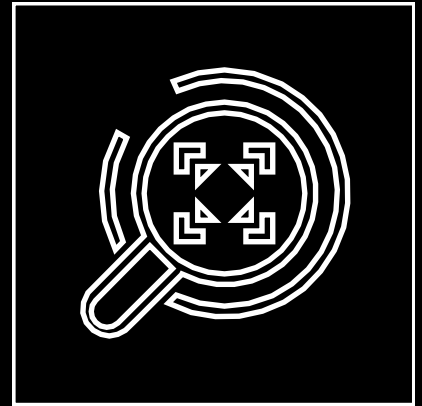
Celebrity  
recognition



Text in  
image

# Amazon Rekognition Video

*Analysis for live or stored video content*



Object and  
activity detection



Face search  
and match



Content  
moderation

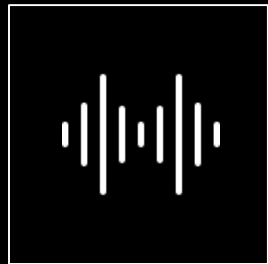
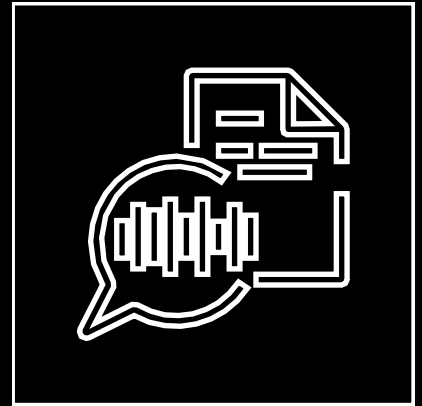


Celebrity  
recognition

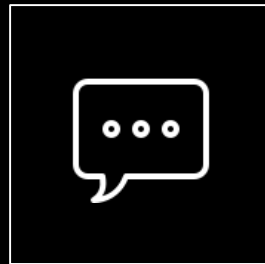


# Amazon Transcribe

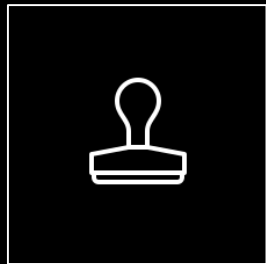
*Automatic speech recognition for stored or real-time audio*



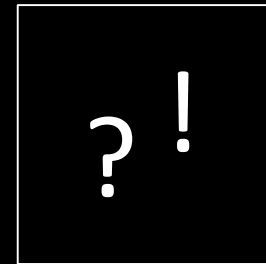
Support for  
both regular and  
telephony audio



Transcription  
of audio from  
16 languages  
and dialects



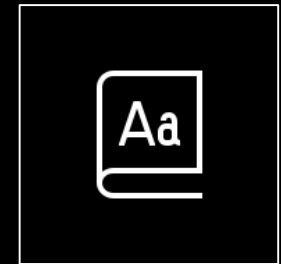
Timestamps



Punctuation  
and formatting



Recognition  
of multiple  
speakers



Custom  
dictionaries

# AWS solutions

# AWS media solutions



## One-click launch

Launch a **fully tested solution starter kit** in your AWS account with **one click**



## Accelerate revenue

Get started with your project **quickly** and start **monetizing your content**



## Well-architected

Rely on the architecture as it reflects **AWS best practices**



## Agility

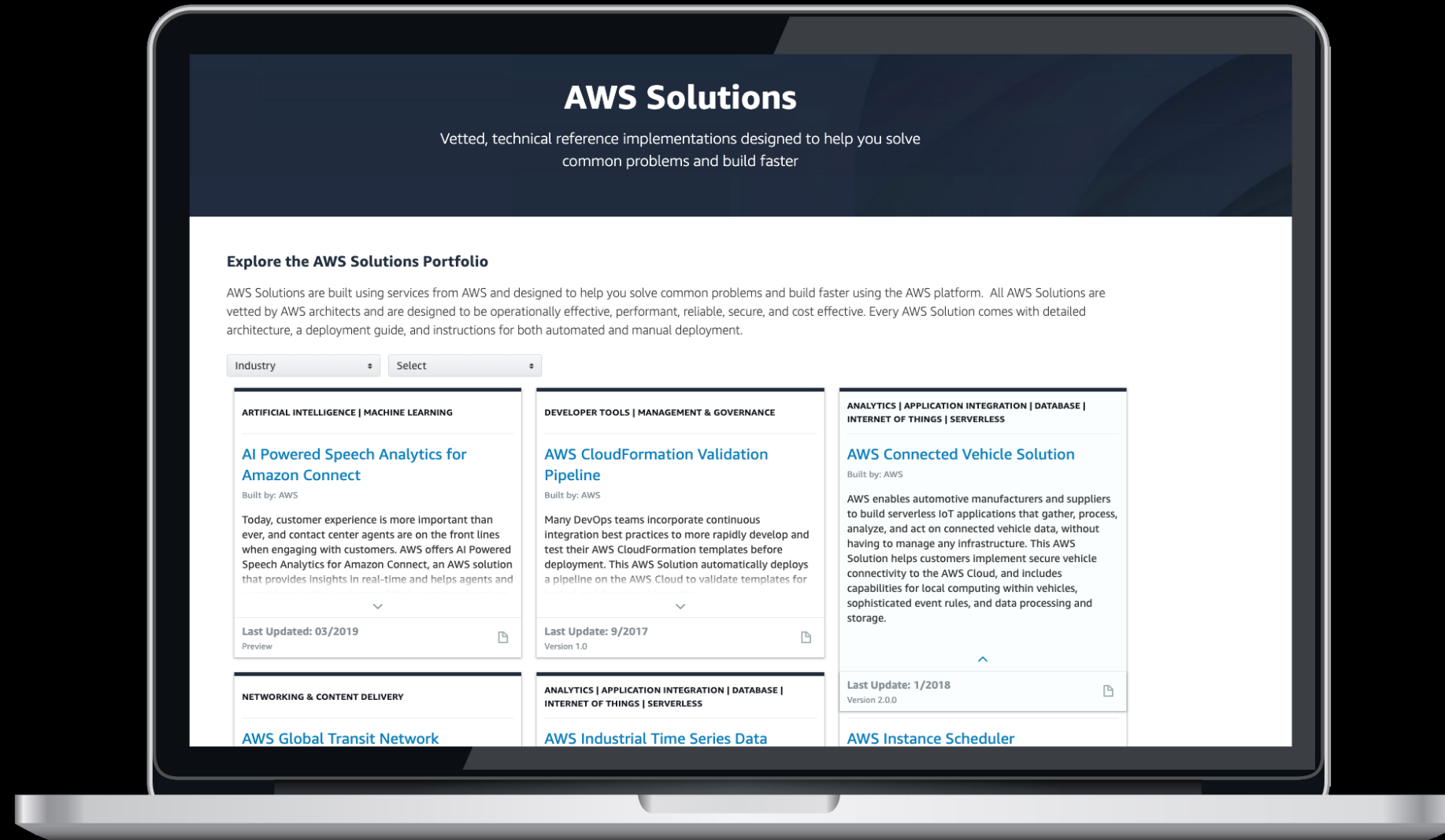
**Save the resources** of developing a solution from the beginning



## Expand and add

Use the framework and available guides to **expand the solution starter kit** based on your needs

# AWS media solutions



Go to: [aws.amazon.com/media/solutions](https://aws.amazon.com/media/solutions)

# Media analysis solution

# Media analysis solution



Amazon  
Rekognition



Amazon Rekognition  
Video



Amazon  
Transcribe



Amazon  
Translate



Amazon  
Comprehend

Automated  
metadata  
generation

Label and face detection  
Celebrity detection  
Face search  
Activity tracking

Transcription

VoD subtitle generation

Context

Key entity and phrase detection

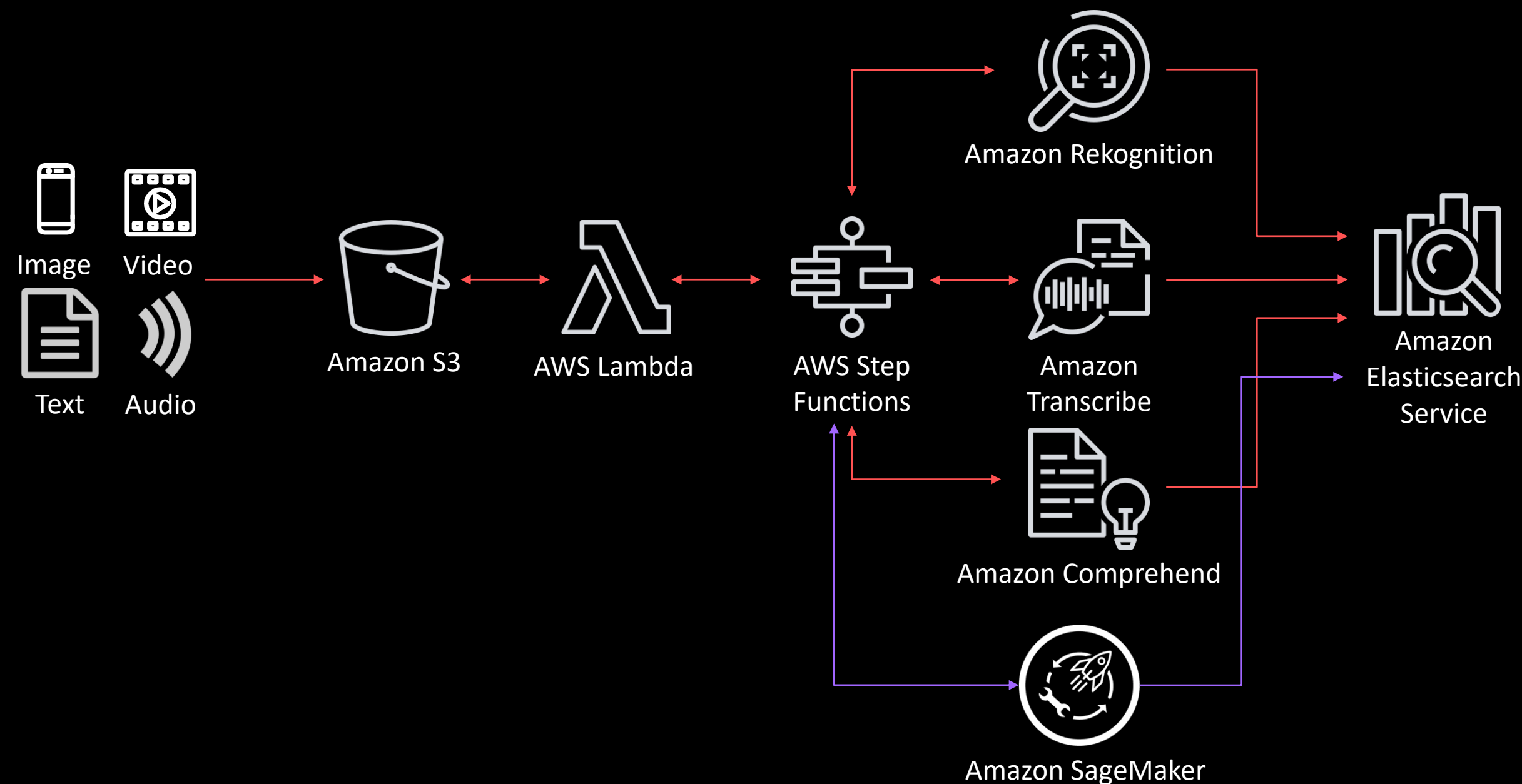


Amazon SageMaker

+ additional solutions for media use cases



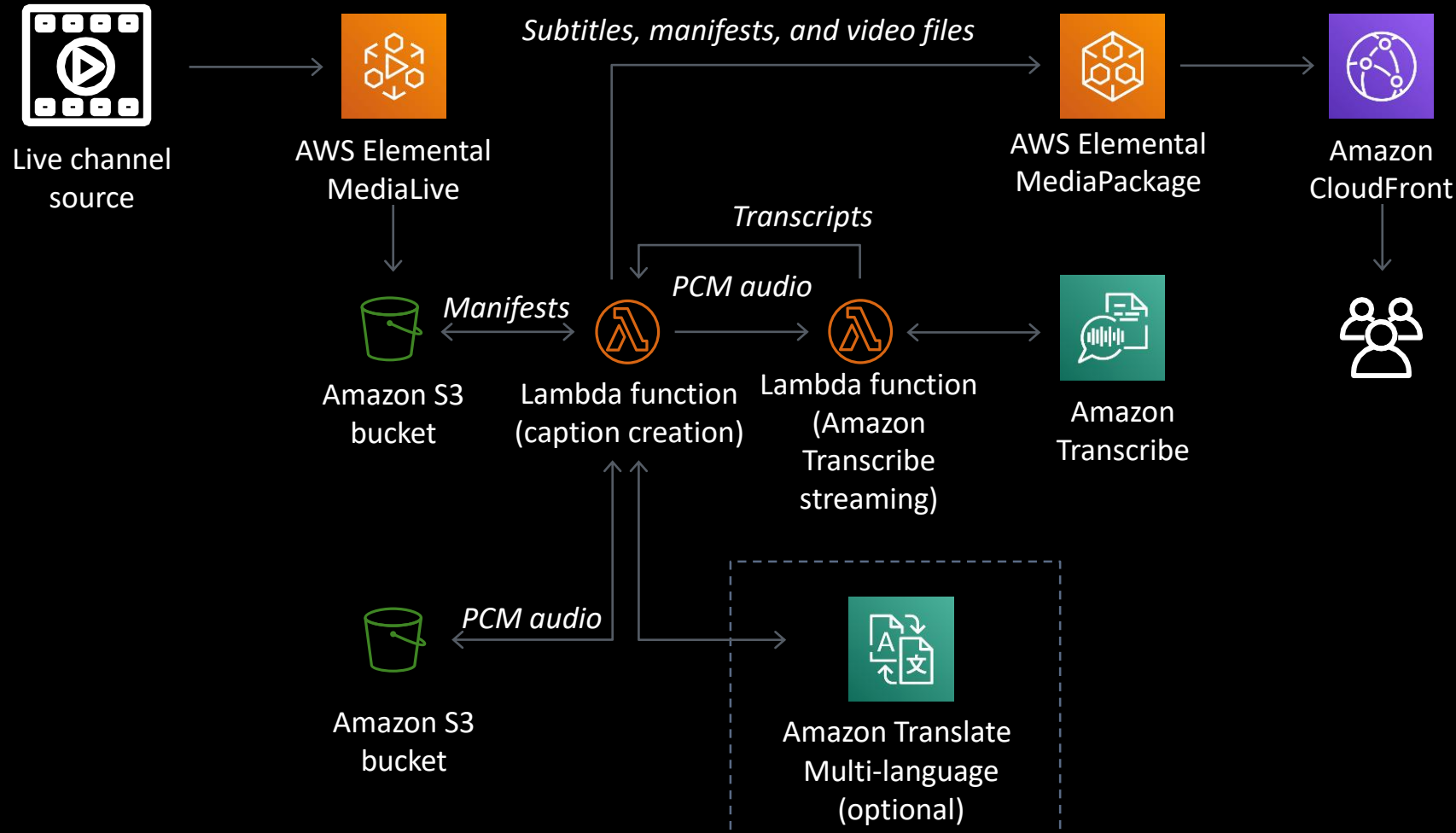
# Media analysis solution



# Media analysis solution walk-through

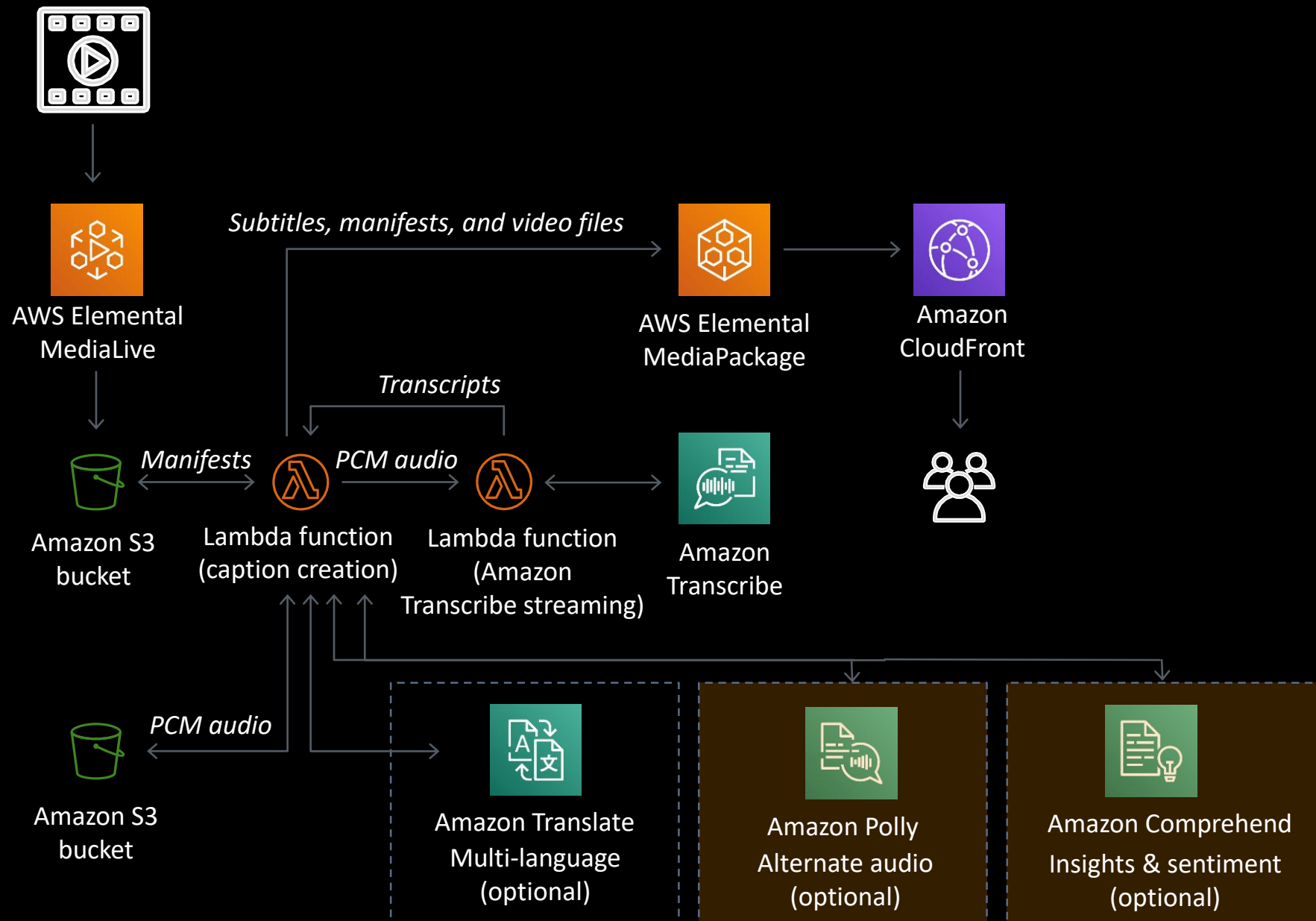
# Live subtitling solution

# Live subtitling solution



- Purpose-built to help customers create live streaming subtitling workflows on the AWS Cloud
- Serverless architecture
- Uses new Amazon Transcribe streaming to extract speech for subtitles, then Amazon Translate to convert into other subtitle languages
- Deploy in minutes with a single click using an AWS CloudFormation template
- Automatically spins up all resources needed for the entire workflow

# Live subtitling solution



- Purpose-built to help customers create live streaming subtitling workflows on the AWS Cloud
- Serverless architecture
- Uses new Amazon Transcribe streaming to extract speech for subtitles, then Amazon Translate to convert into other subtitle languages
- Deploy in minutes with a single click using an AWS CloudFormation template
- Automatically spins up all resources needed for the entire workflow
- Optionally add text to speech with Amazon Polly or use Amazon Comprehend to understand text sentiment

# Live subtitling solution walk-through



# Thank you!

Bryan Samis  
samisb@amazon.com