
AWS
re:Invent
NOV. 28 – DEC. 2, 2022 | LAS VEGAS, NV

GAM204-R

Protect gamer experiences with AI content moderation

Lana Zhang

Sr. Solutions Architect
AWS

Raviteja Yelamanchili

Solutions Architect
AWS

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Agenda

1. Damaging business effects of toxicity and customer expectations for content moderation
2. Content moderation in gaming
3. Bringing AWS AI services together
4. Customer success stories
5. Demo
6. How to get started and Q&A

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Let’s get personal

- How many of you here have experienced toxic interactions when playing games on social media or on online platforms?
- How did that interaction change your perception of the product?





© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Toxicity damages player base and game success

72% gamers have witnessed toxicity towards others

66% experienced toxicity in games

43% quit playing entirely when toxicity arises

Toxicity in Multiplayer Games, 2021 Unity Technologies Report



aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

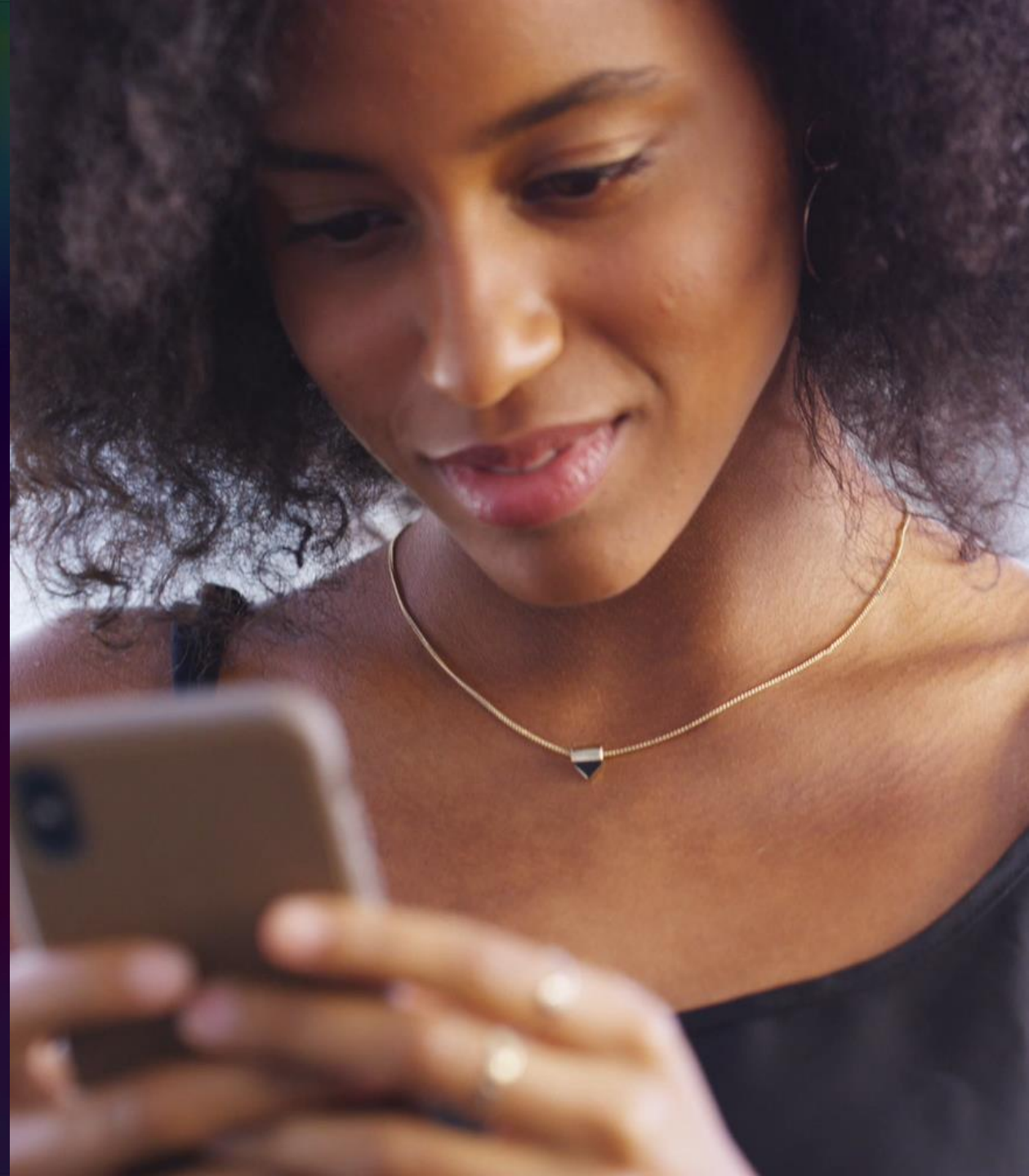
Everyone expects inclusion and safety

50% of people create content at least once daily[1]

40% of consumers disengage after one exposure to toxic UGC[2]

70% of consumers agree that brands are responsible for moderating UGC[2]

[1] Stackla, Customer Survey
[2] BusinessWire, Customer Survey



aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

GAMING



CONTENT TYPE

User comments

In-game voice and text interactions

Shared images and videos in gaming lounges and forums

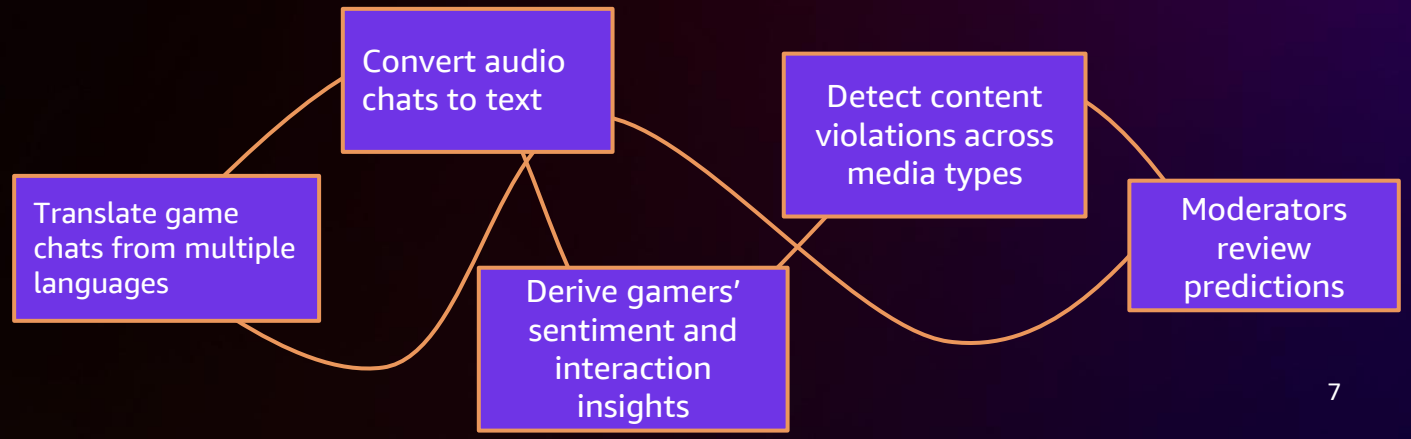
Gamers' live streams and recordings

WHY

Members' inclusiveness and engagement

Prevention of exposure to harmful content (hate, profanity, bullying)

Protect key gaming metrics (K-Factor, daily active users, revenue per user)



Multimodal capabilities with AWS AI and ML



IMAGE

Identify objects, text, logos, scenes, and filter harmful visuals and unwanted associations

Flag and label inappropriate content across hierarchical categories

TEXT

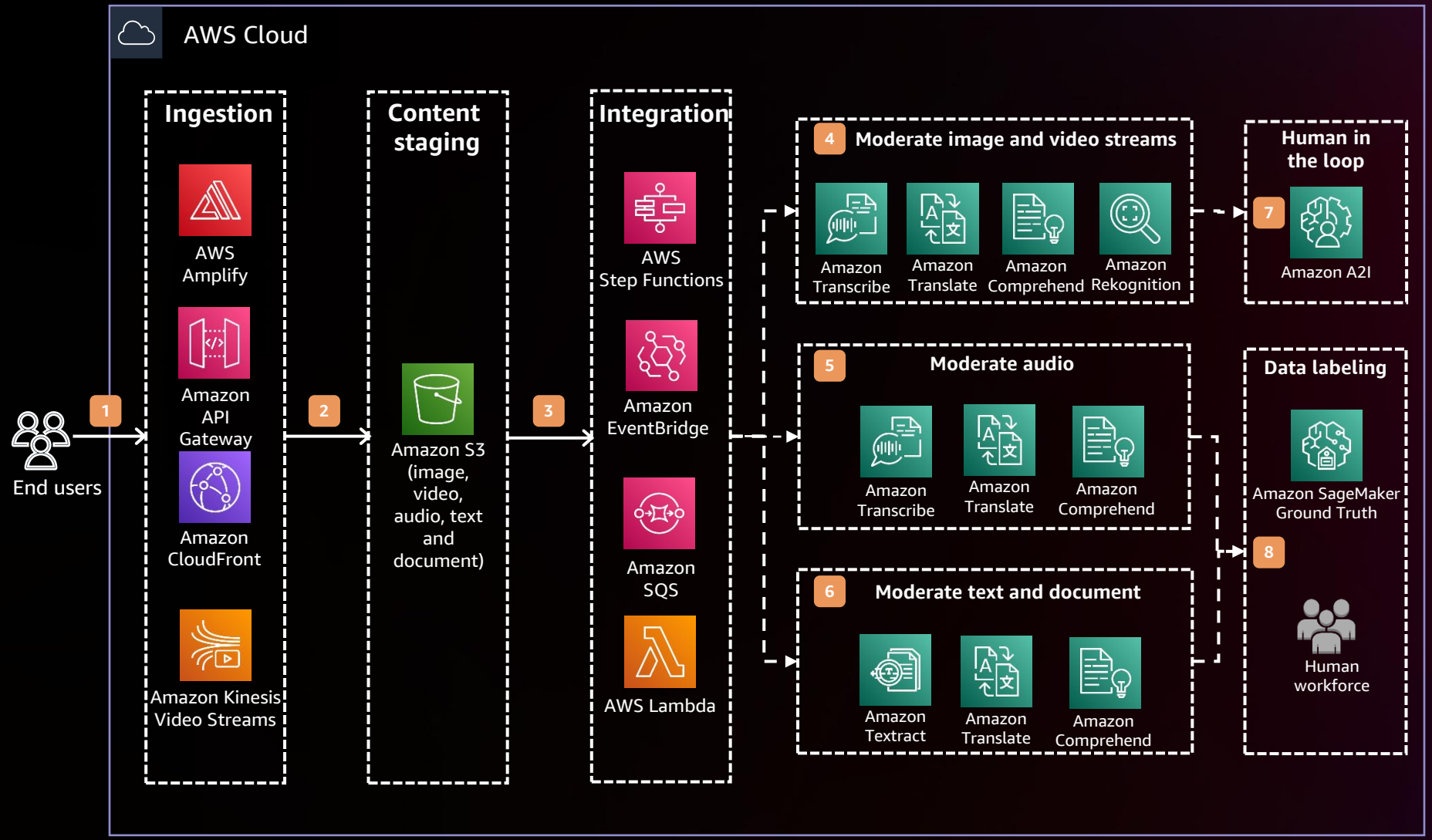
Automatically identify profanity, filter, and redact personal information

AUDIO

Convert speech to text to filter profanity and any unwanted vocabulary

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Bringing all AWS AI services together



1. Customers upload their content into the **AWS Cloud**
2. Content securely persists into an **Amazon Simple Storage Service (Amazon S3)** bucket or another data store
3. Activate workflows, publisher/subscription patterns, and custom code to moderate content
4. Process the audio streams within video streams using **Amazon Transcribe** and **Amazon Rekognition**, and extract content moderation categories using simple APIs
5. Use **Amazon Transcribe** to convert audio into text and leverage natural language processing (NLP) with **Amazon Comprehend**
6. Use **Amazon Textract** to extract and **Amazon Comprehend** natural language processing to moderate content
7. Bring humans into the loop for scenarios that are not fully automatable using **Amazon Augmented AI (Amazon A2I)**
8. Integrate human workforces to customize model vocabularies and image labels using **Amazon SageMaker Ground Truth**

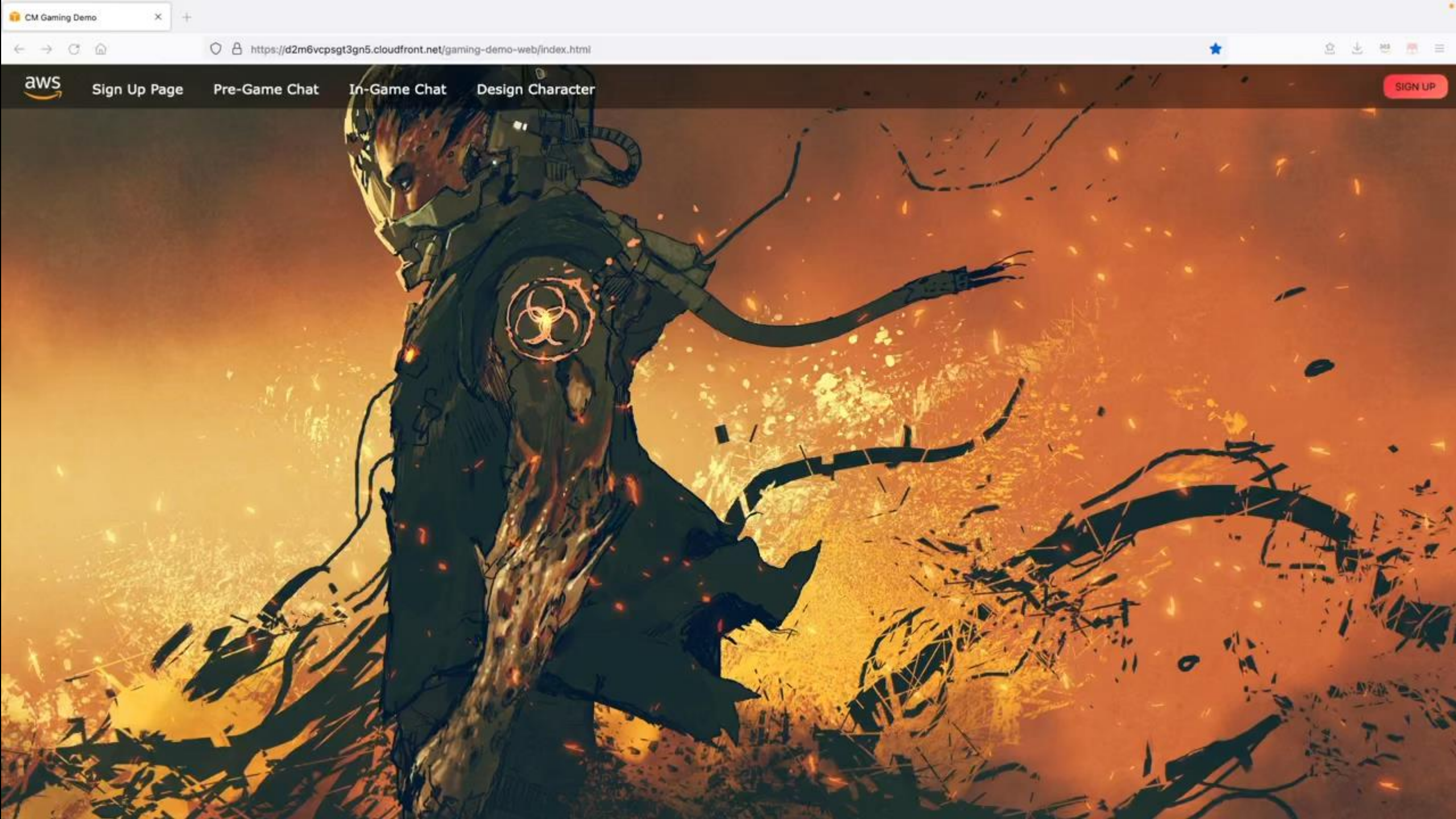
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Customer success stories

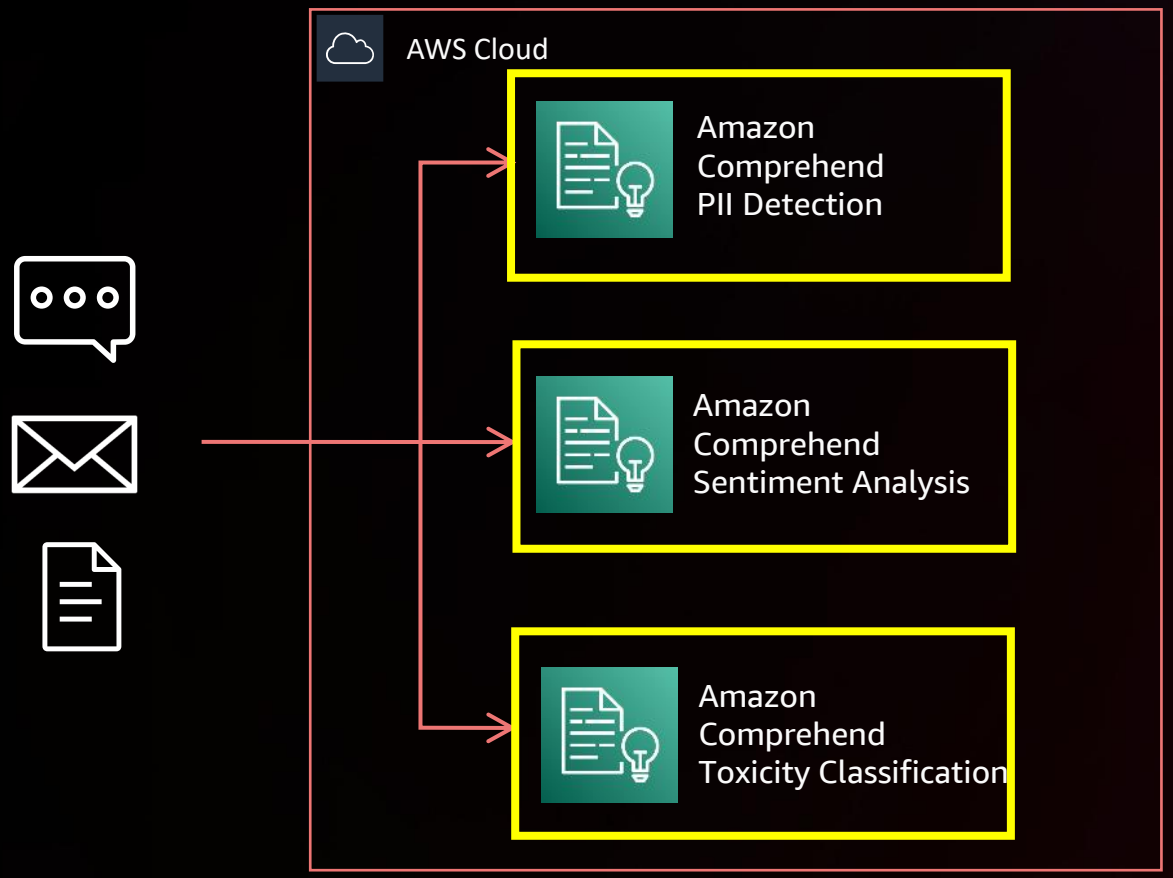
DeNA	DREAM11	MOBISOCIAL
Pococha voice community	Indian fantasy sports community that has grown from 2 to 100+ million members in 4 years	Gaming and streaming platform
Identification and redaction of words not recognized by standard speech-to-text services	Platform that safely automates content sharing and moderation	Flagging of non-compliant images and videos
Higher detection accuracy and faster moderation response times	Thousands of assets efficiently reviewed every day	Manual moderation reduced by 95% with increased accuracy and scalability

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.


CM Gaming Demo
https://d2m6vcpsgt3gn5.cloudfront.net/gaming-demo-web/index.html
aws
Sign Up Page
Pre-Game Chat
In-Game Chat
Design Character
SIGN UP

Uncovering connections in text



Amazon Comprehend
DetectPiiEntity API
Detect and redact PII entities

Amazon Comprehend
DetectSentiment API
Inspect text sentiment

Amazon Comprehend
Custom Classifier
Classify toxicity of sentences

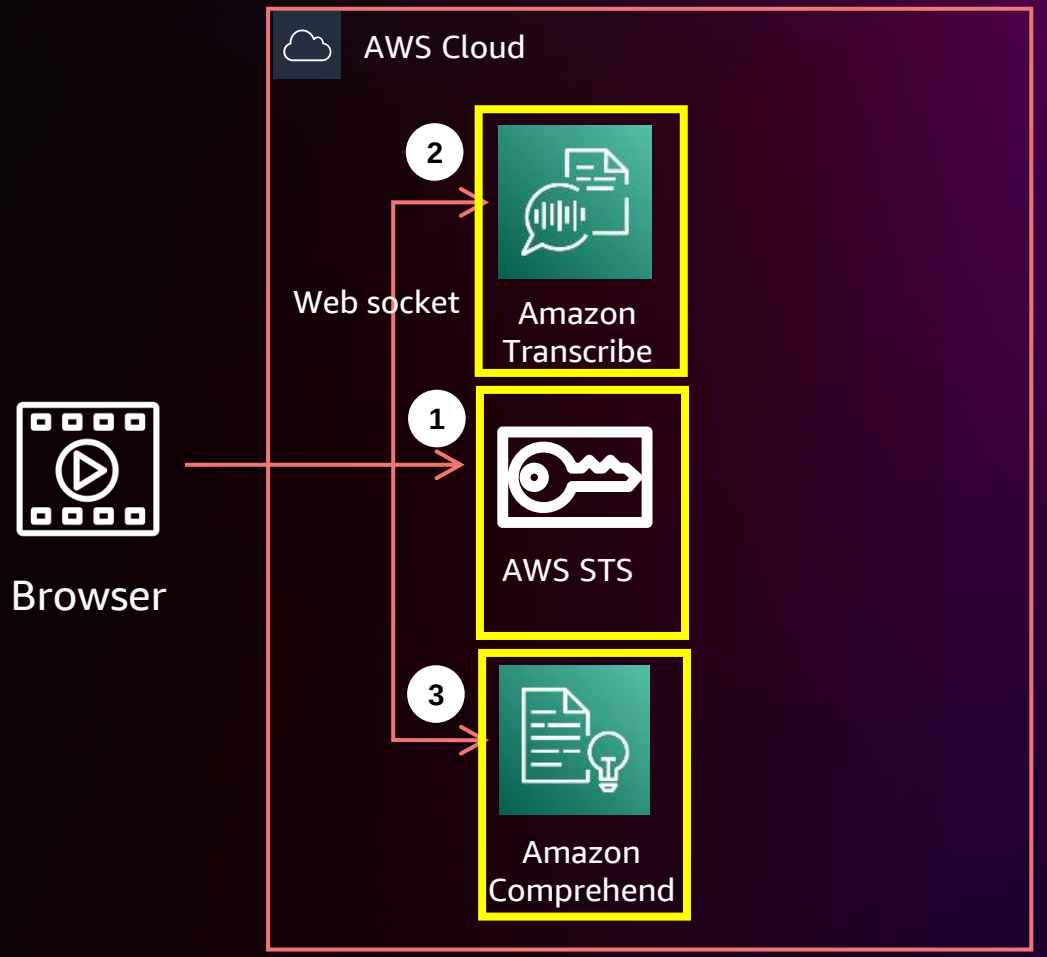
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Live audio moderation

Get STS token

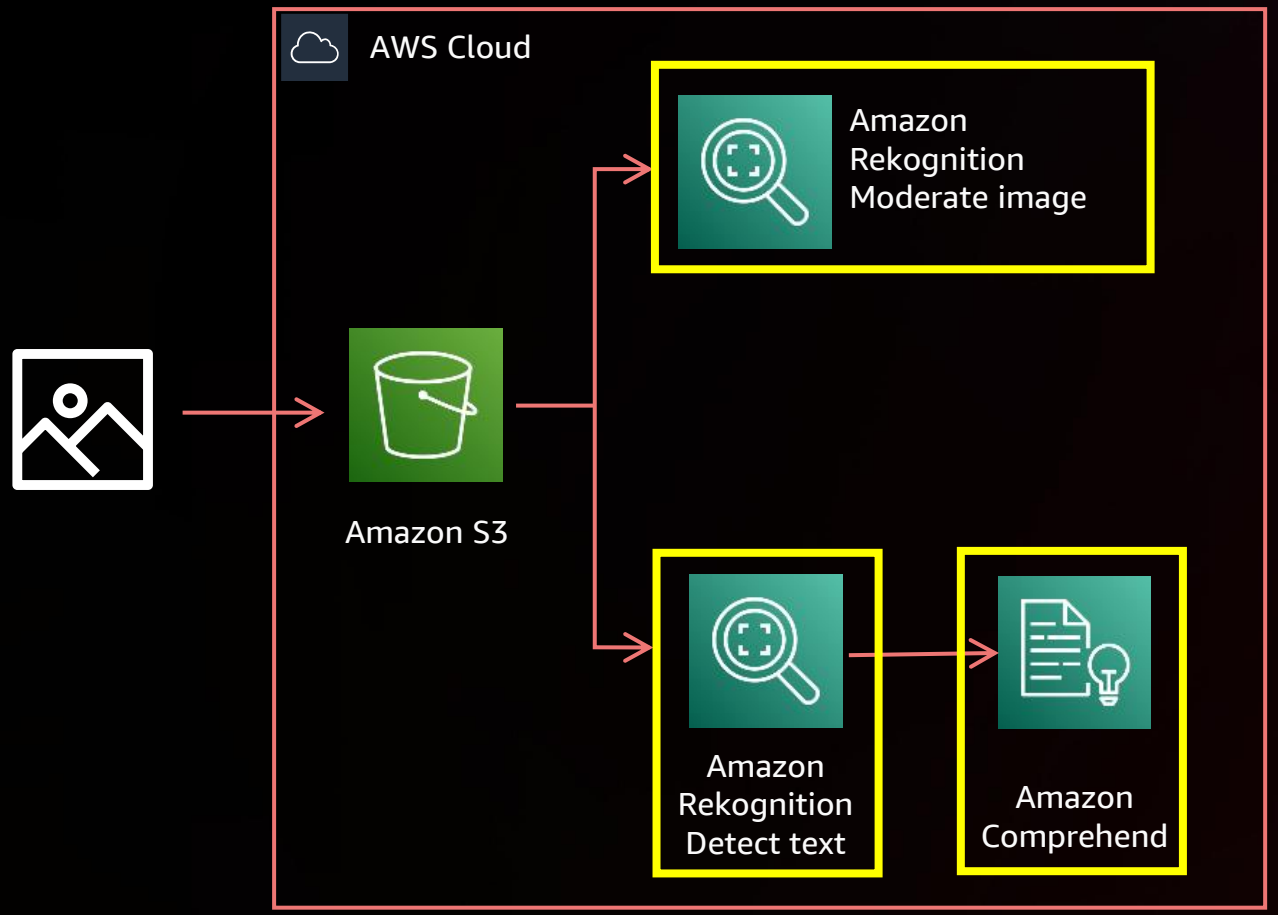
Amazon Transcribe
real-time streaming API
Transcribing audio to text

Amazon Comprehend
Further text moderation



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Image moderation



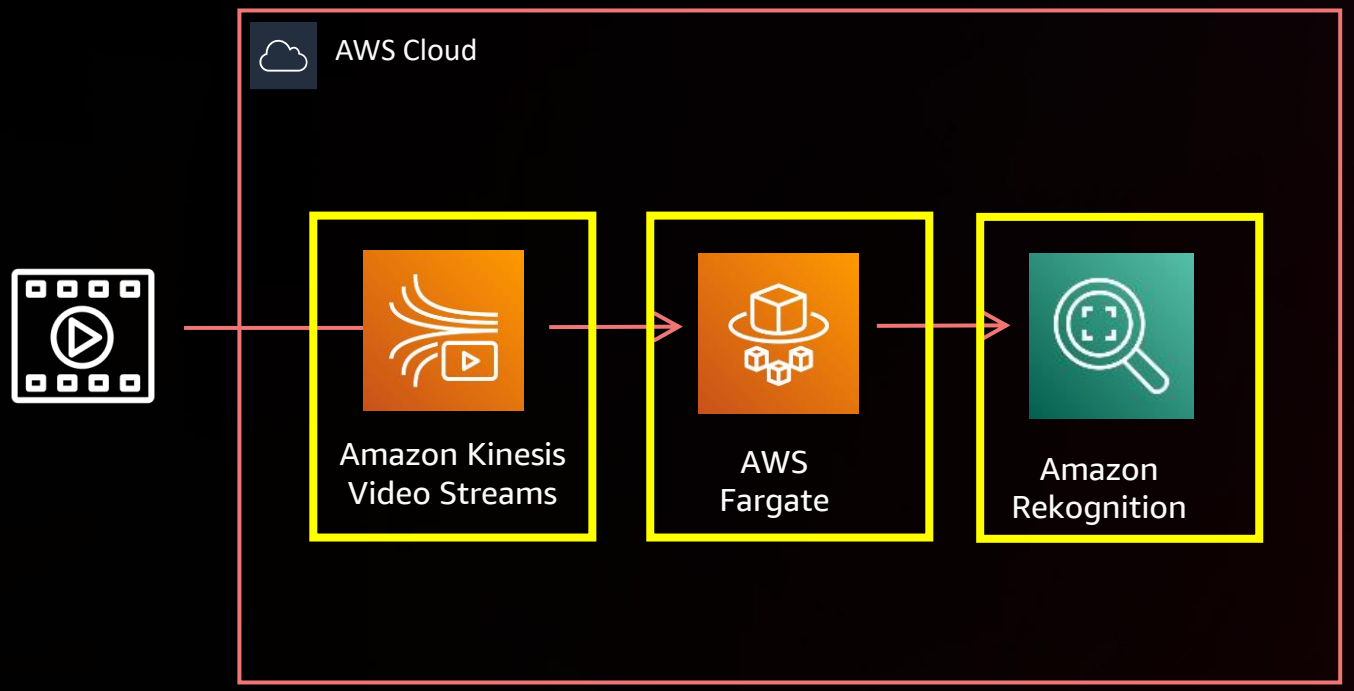
Amazon Rekognition
DetectModerationLabels API
Detecting unsafe content in images

Amazon Rekognition
DetectText API
Detecting text in images

Amazon Comprehend
Further text moderation

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Moderate real-time video stream



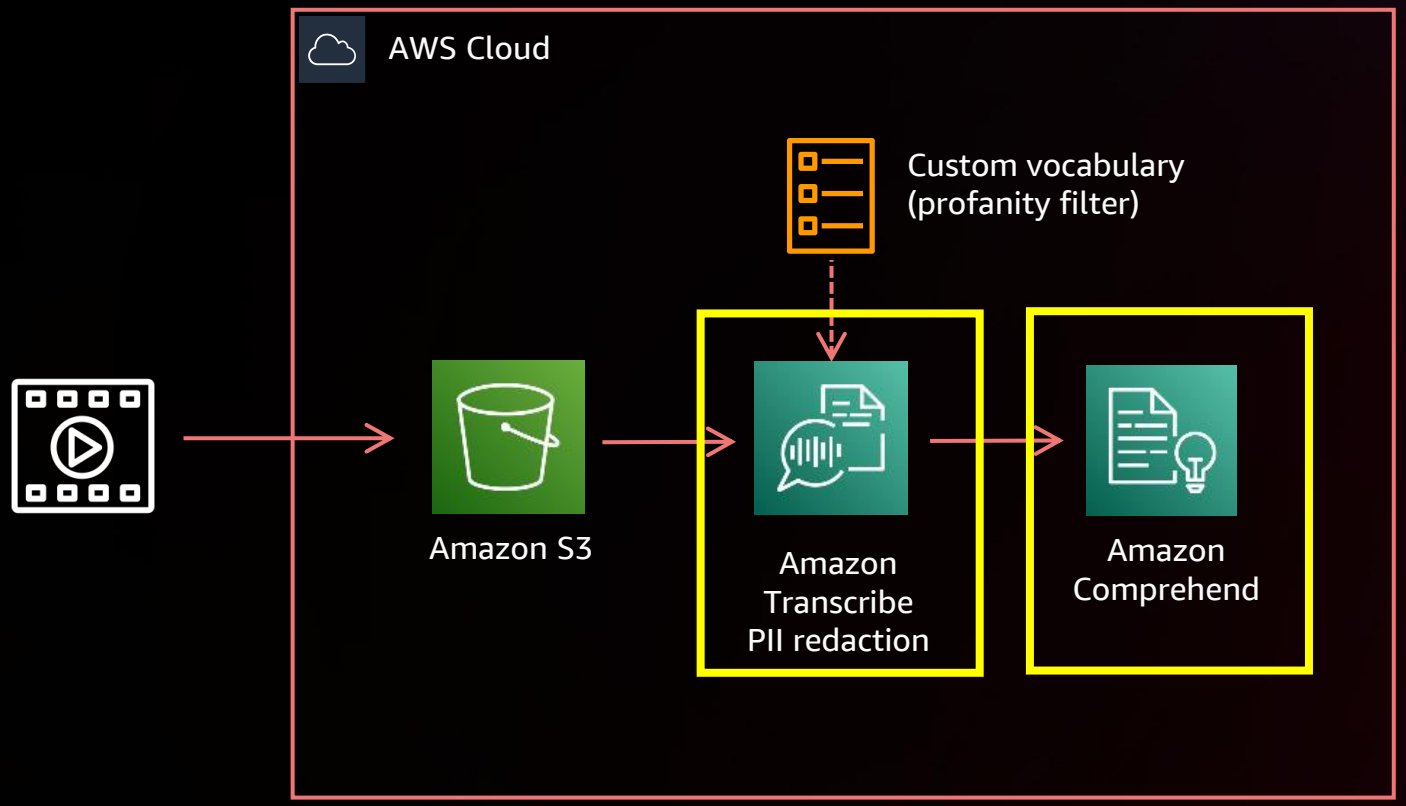
Amazon Kinesis Video Streams
Consume video streams in real-time

AWS Fargate
Send decoded frames to Amazon Rekognition

Amazon Rekognition
DetectModerationLabels API
Detecting unsafe content in video frames

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Moderate audio of the stored media files



Amazon Transcribe
StartTranscriptionJob API
Transcribing audio to text with vocabulary filter and PII redaction

Amazon Comprehend
for text moderation

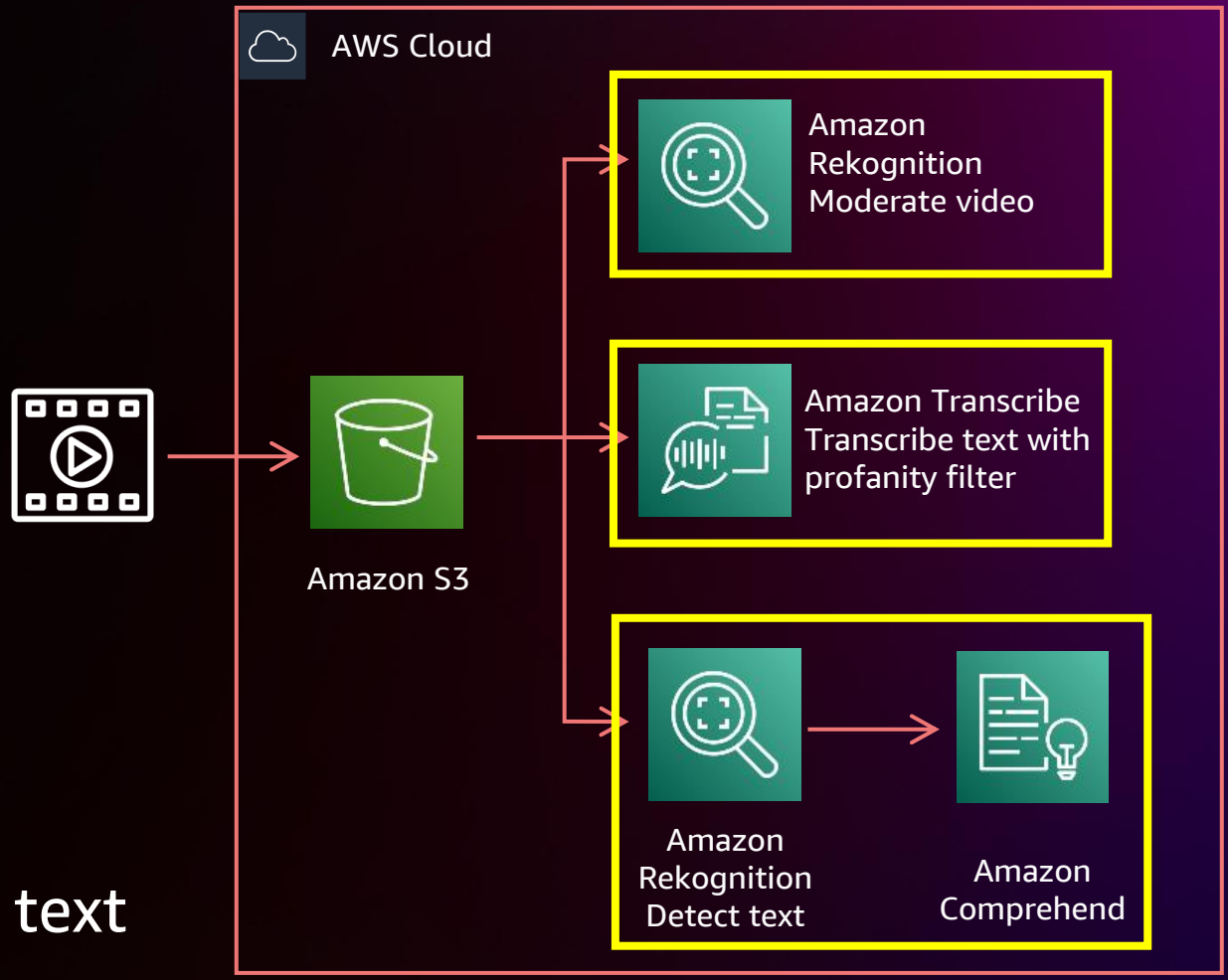
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Moderate stored videos

Amazon Rekognition
DetectModerationLabel API
Detecting unsafe content in video

Amazon Transcribe
StartTranscriptionJob API
Transcribing audio to text with custom vocabulary filter

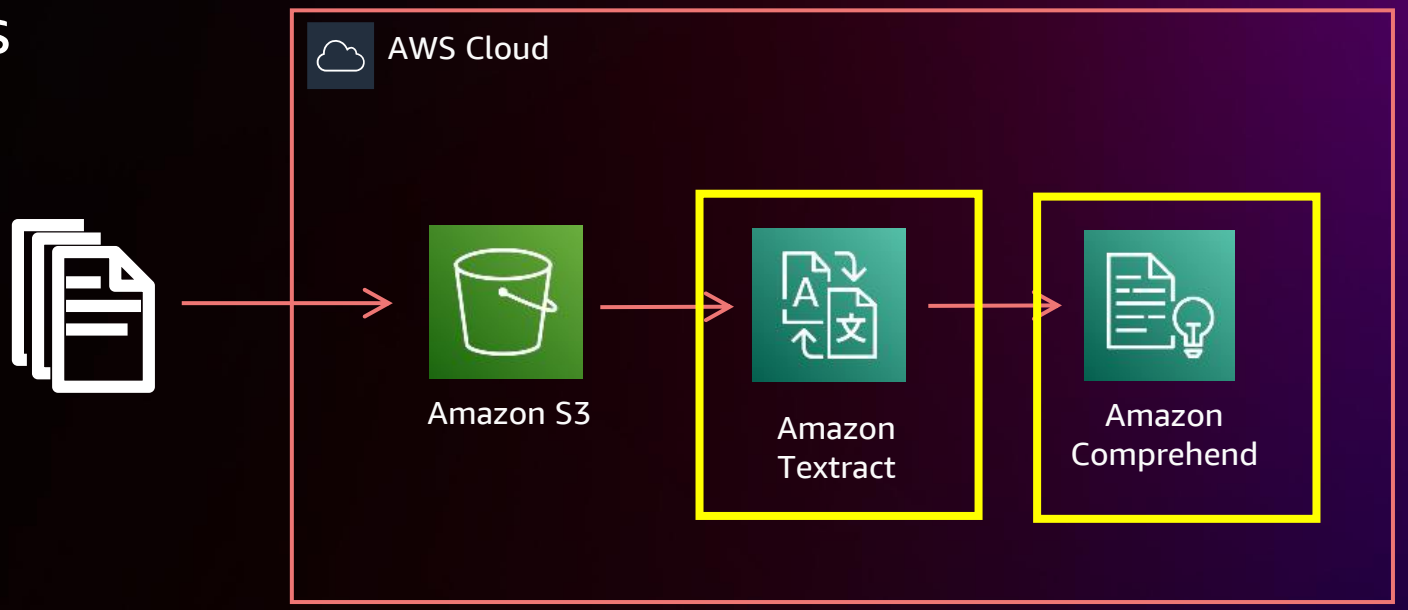
Amazon Rekognition
DetectText API
Detecting text in video and moderate text with Amazon Comprehend



aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Document moderation

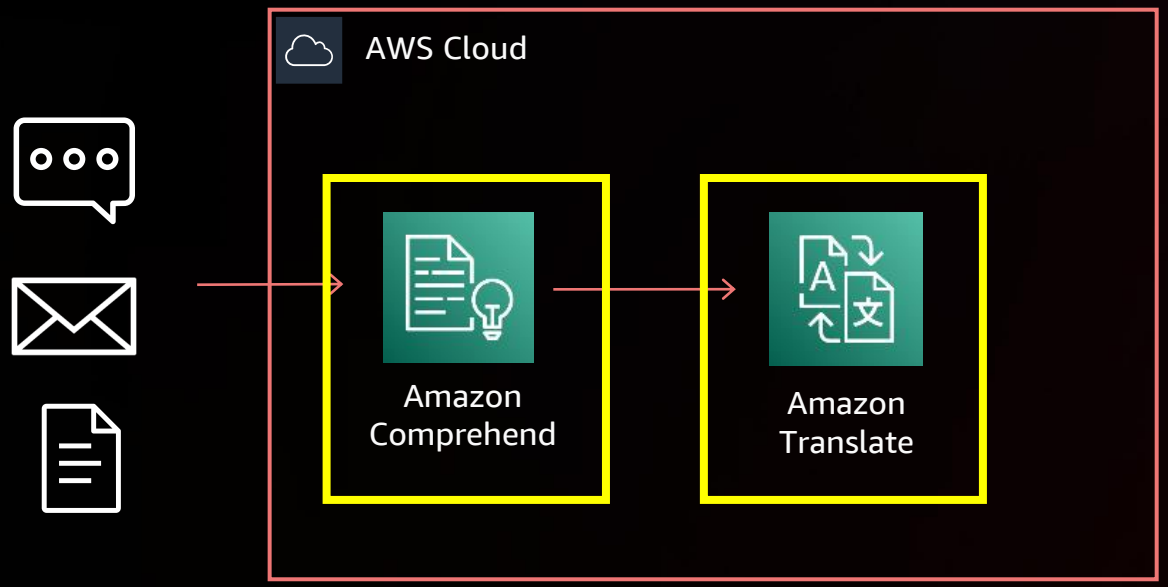
StartDocumentAnalysis API starts the asynchronous analysis of an input document for relationships between detected items



GetDocumentAnalysis API gets the results for an Amazon Textract asynchronous operation that analyzes text in a document

aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Moderation with Amazon Translate



Amazon Comprehend
DetectDomainLanguages API
Detect the language

Amazon Translate
TranslateText API
Translate the text to the target language

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Get started



1 Try and learn
AWS Content Moderation webpage
New AWS Guidance and more

2 Build with us
Contact us

3 Find a partner
AWS Partner Network

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Thank you!

Lana Zhang
LinkedIn: in/lanazhang



Raviteja Yelamanchili
LinkedIn: in/ravitejay





Please complete the session survey in the **mobile app**

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.