

# AWS re:Invent

NOV. 28 – DEC. 2, 2022 | LAS VEGAS, NV

# Improve citizen services using intelligent document processing

Srinath Godavarthi

Principal Solutions Architect  
World Wide Public Sector  
AWS

Ben Snively

Sr. Principal Solutions Architect  
World Wide Public Sector  
AWS

# Agenda

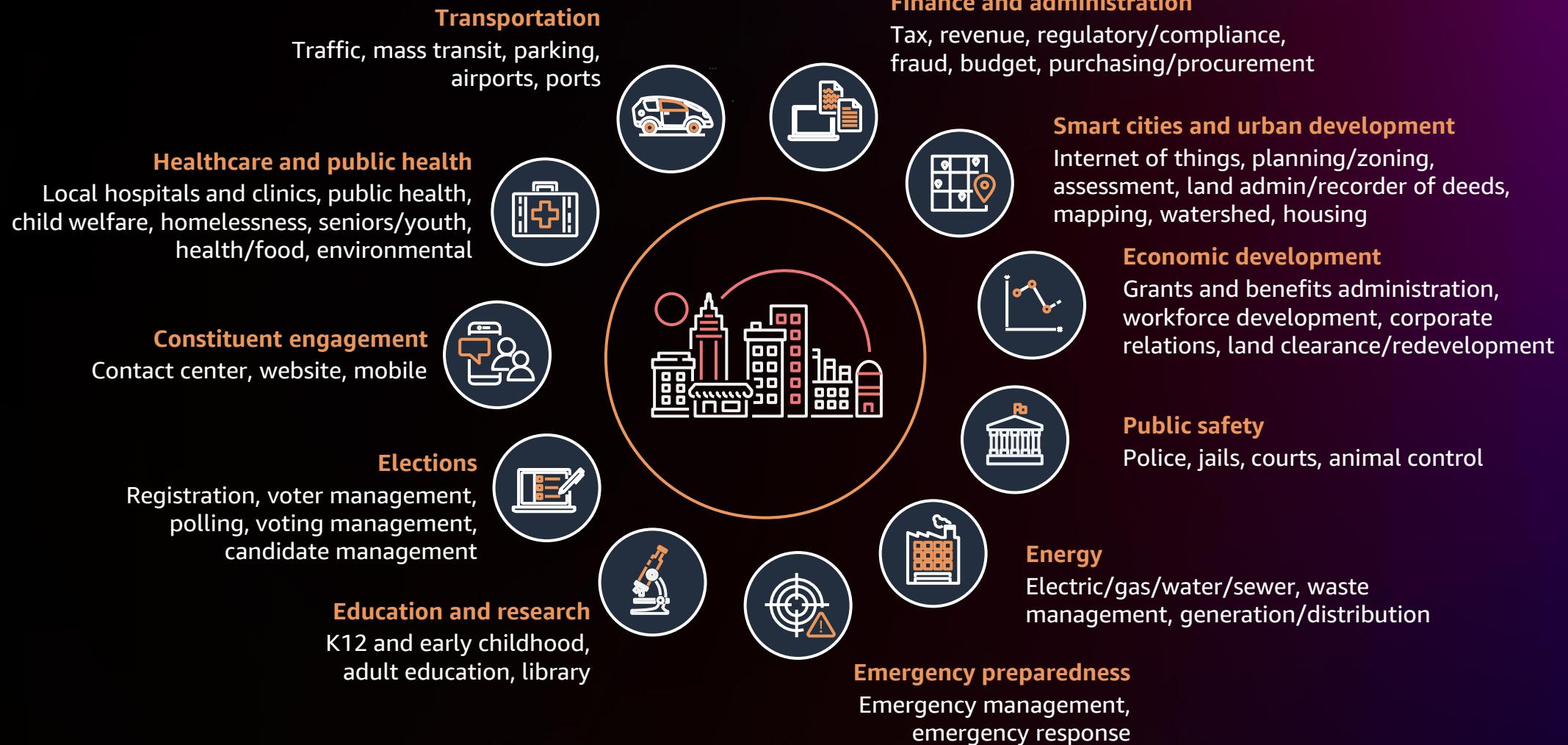
Public sector challenges with document processing

Need for intelligent document processing (IDP)

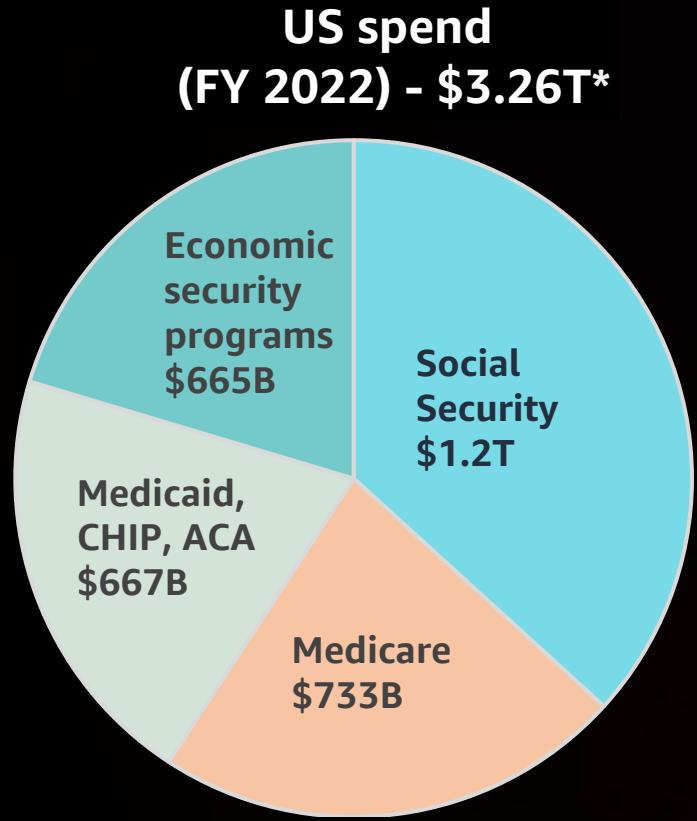
IDP workshop overview/scenarios

Call to action/next steps

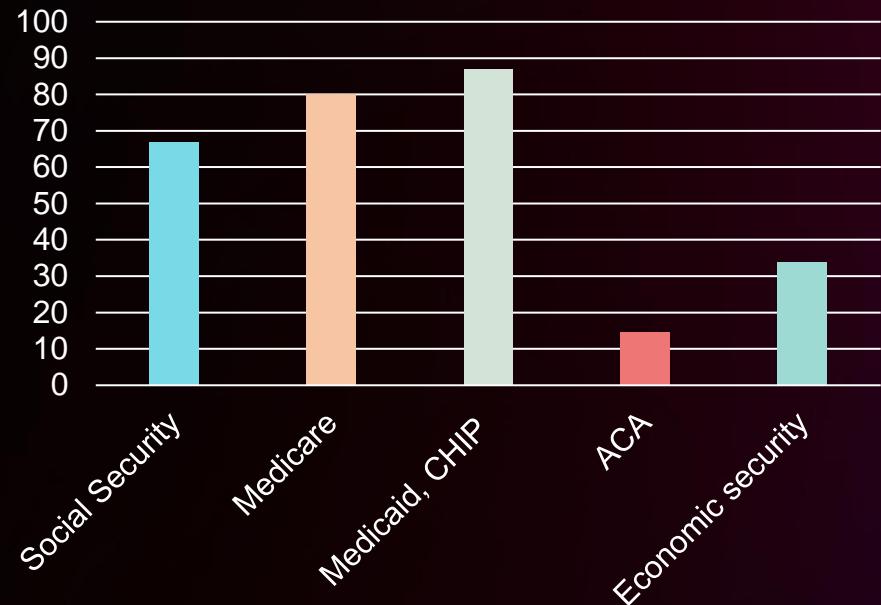
# Document processing is critical for public sector mission



# US spend – social/healthcare programs



**Beneficiaries – FY 2022  
(Millions)**



- Social security
- Medicare
- Medicaid
- Affordable Care Act (ACA)
- Unemployment insurance
- SNAP (food stamps)
- School meals
- Low income housing assistance (LIHEAP)
- Childcare assistance

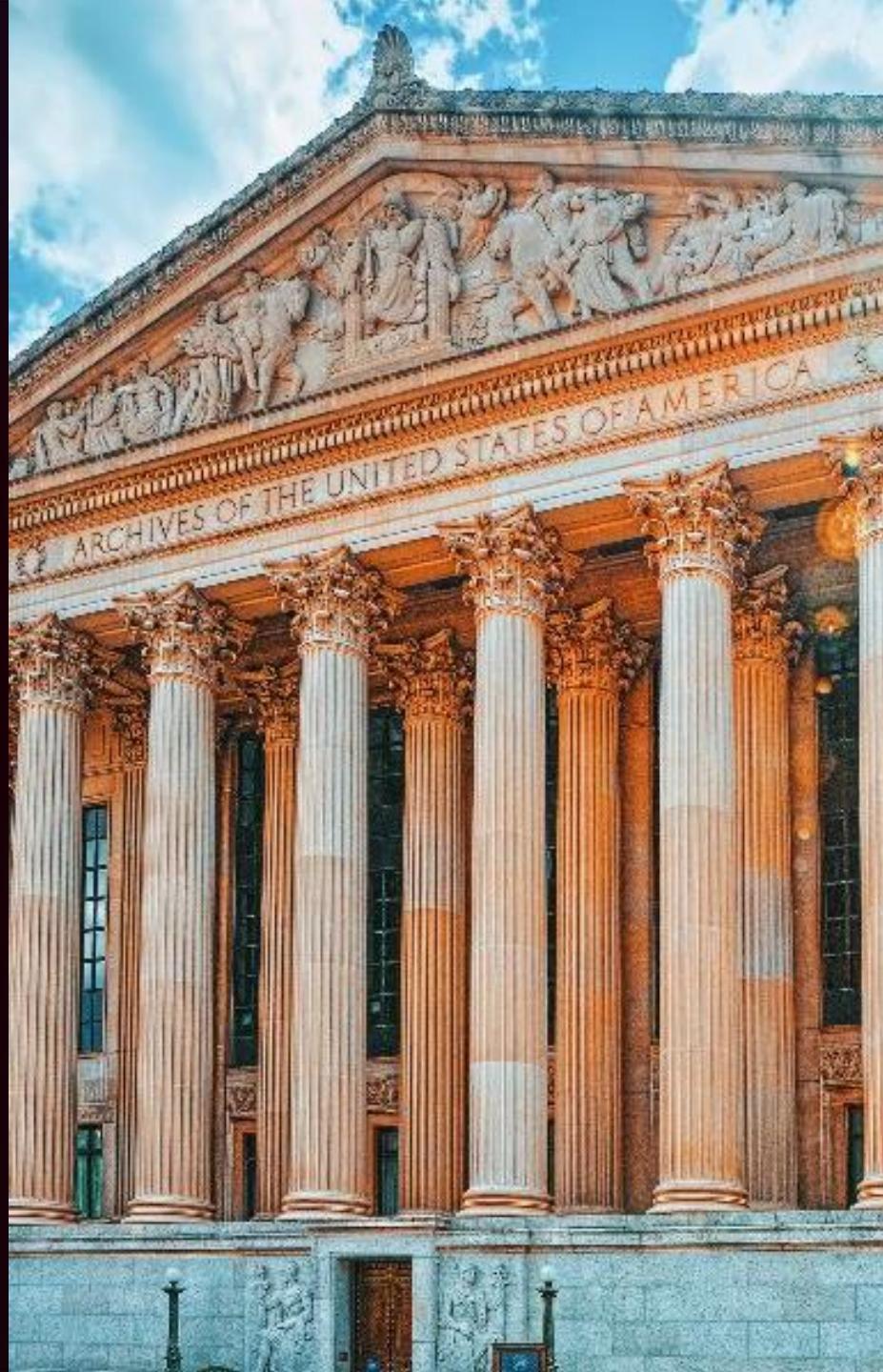
\***Estimated:** 56% of total (\$5.8T) US spending in 2022

**Source:** Center on Budget and Policy Priorities – <https://www.cbpp.org/research/federal-budget/where-do-our-federal-tax-dollars-go>

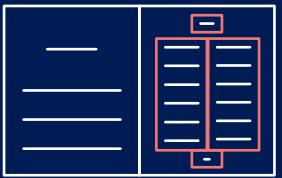
# The need for intelligent document processing

The US federal government has more than 23,000 different forms that lead to more than 11.4B hours of paperwork annually\*

\*Source: American Action Forum – <https://www.americanactionforum.org/press-release/aaf-analysis-finds-federal-government-imposes-23000-forms/>



# Legacy document processes do not meet today's needs



Legacy optical character recognition (OCR) and manual process are time-consuming, error-prone, and expensive

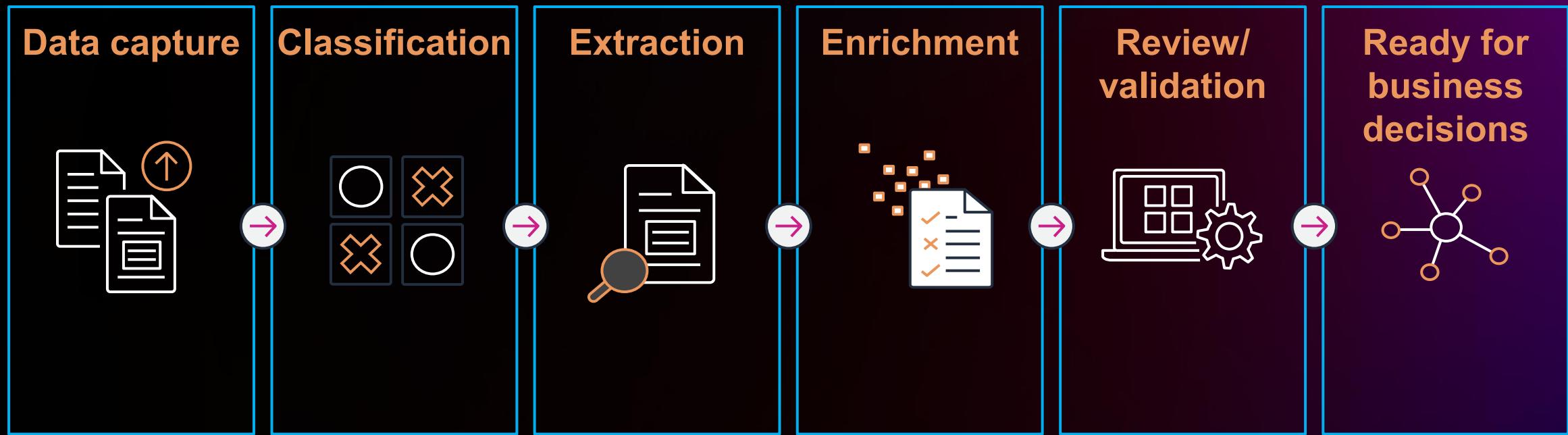


Manual processes do not scale easily with document volume



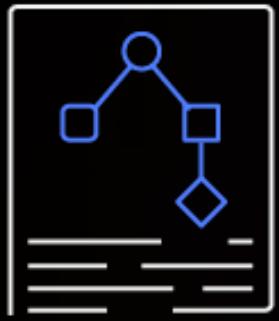
Challenging to find useful information needed for business decisions

# Automate document processing workflows



**Faster document processing shortens decision cycles, which enables you to serve more constituents and have agency staff focus on higher value tasks**

# Summary: Primary use cases



**Process  
automation**



**Search and  
discovery**



**Compliance  
and control**

# Customer case studies

# Improving access to public records



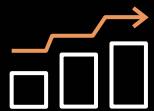
## Problem

The US National Archives and Records Administration (NARA) released the 1950 census data to the general public. Census data is released 72 years after a census is conducted, and it has been 10 years since the last census data for the 1940 Census was publicly released.



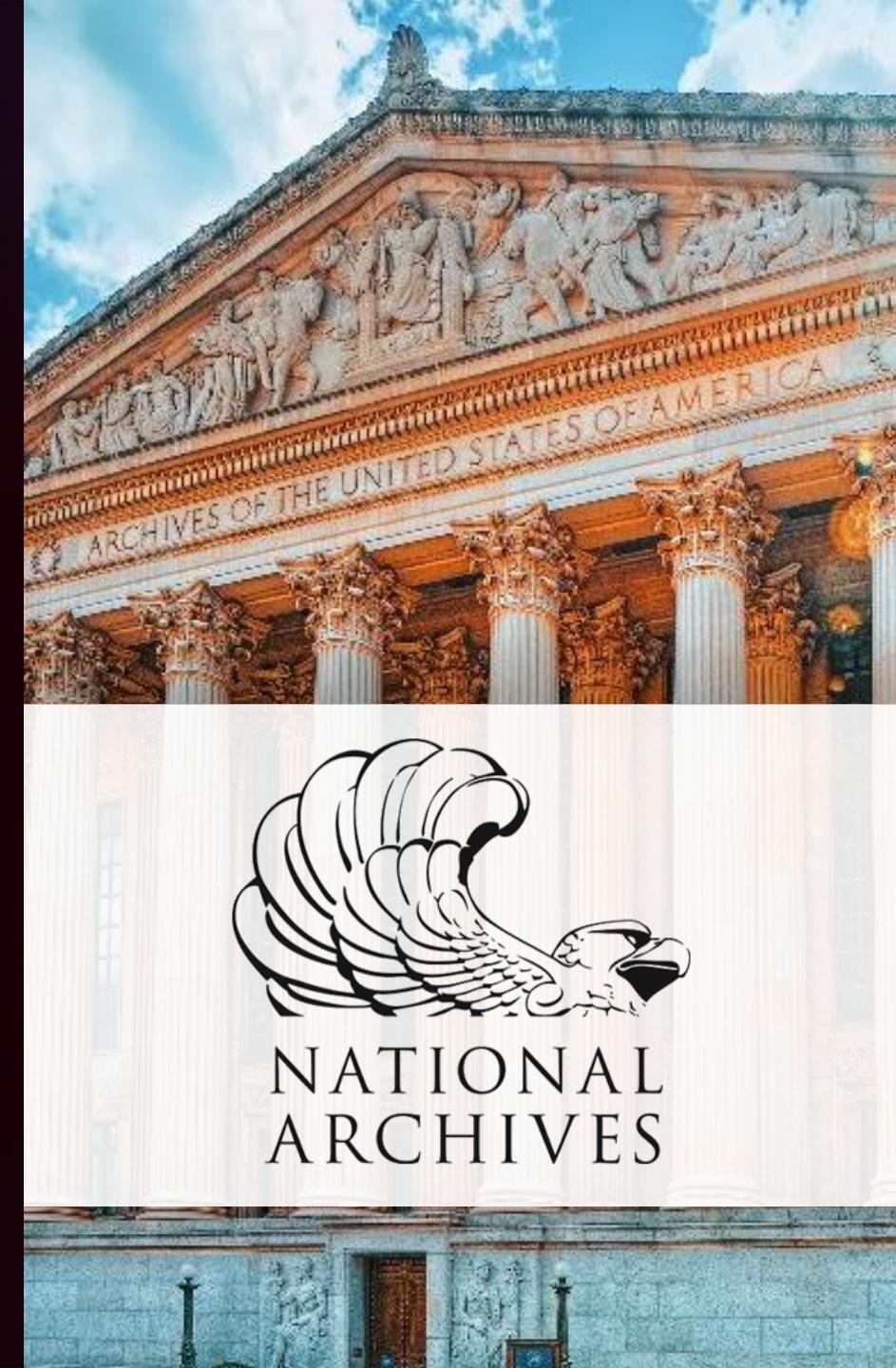
## Solution

NARA is also leveraging machine learning (ML) to improve access to the records. Using Amazon Textract, NARA was able to extract and ingest approximately 130 million handwritten names from the 1950 census population schedules, enabling 1950 Census website visitors to search the schedules by name.



## Impact

Cloud services, including AI and ML, are helping customers across the US government improve the ways they deliver their missions. AWS is proud to support NARA's mission as the Archives makes the 1950 census data available to the public.





## PROBLEM

Since the start of the COVID-19 pandemic, frontline healthcare workers are facing the impossible task of **cutting through a “global tidal wave” of information that might contain clues for the best possible treatments**

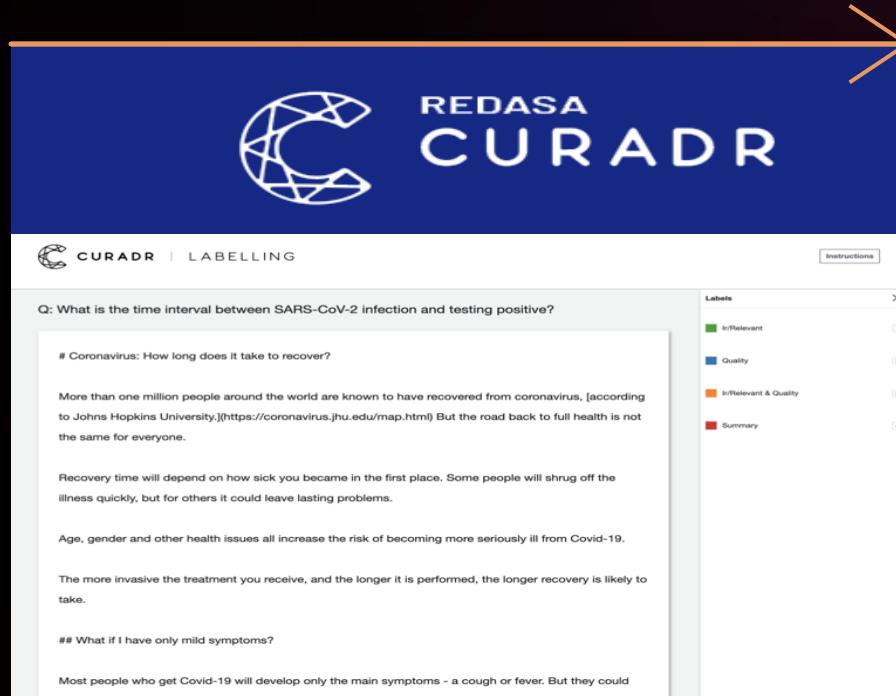
## SOLUTION & IMPACT

Using a combination of Amazon ML services, including natural language search and data labeling built on [Amazon Kendra](#) and [Amazon SageMaker](#) Ground Truth, as well as human review, PanSurg REDASA (REaltime DAta Synthesis and Analysis) can analyze vast amounts of COVID-19 information in real time and quickly extract the most important insights

- ✓ 10,000s of hours of manual research saved

# Implementing intelligent search

PANSURG REDASA (REAL-TIME DATA ANALYSIS AND SYNTHESIS)



Document characterization

Document summary

Document relevance

Topic-based curation package

Document quality

High impact COVID-19 natural language queries

# Workshop overview and instructions

# Workshop overview/structure

You will play the role of three users/personas



---

**End user/constituent –  
applying for benefits**



---

**Case worker –  
leveraging AI/ML to evaluate  
submissions**



---

**Program leadership –  
evaluating program  
effectiveness**

# Workshop overview/structure

- Case study/scenario-based
- Prerequisites + four scenarios
- Focused on personas/mission outcomes



## Improve beneficiary experience

- Reduce overall time to benefits disbursement
- Improve user experience with digital transformation



## Enhance staff productivity

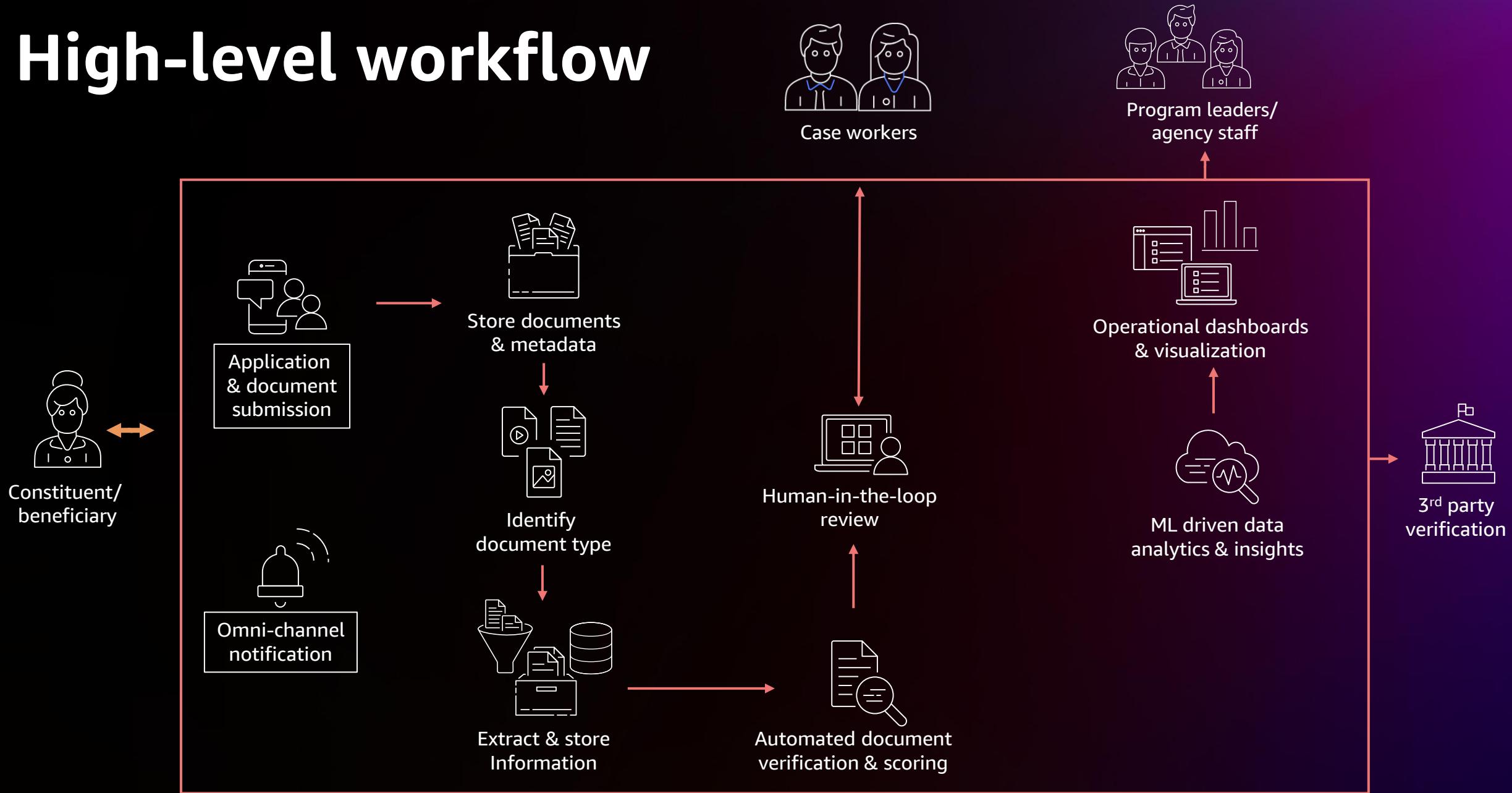
- Reduce manual work/automate processes
- Enhance case review/adjudication processes



## Enable program leadership

- Enable data-driven decision-making
- Enable accurate forecasting on enrollment, staffing, budgets, etc.

# High-level workflow



# The AWS ML stack

BROADEST AND MOST COMPREHENSIVE SET OF MACHINE LEARNING CAPABILITIES

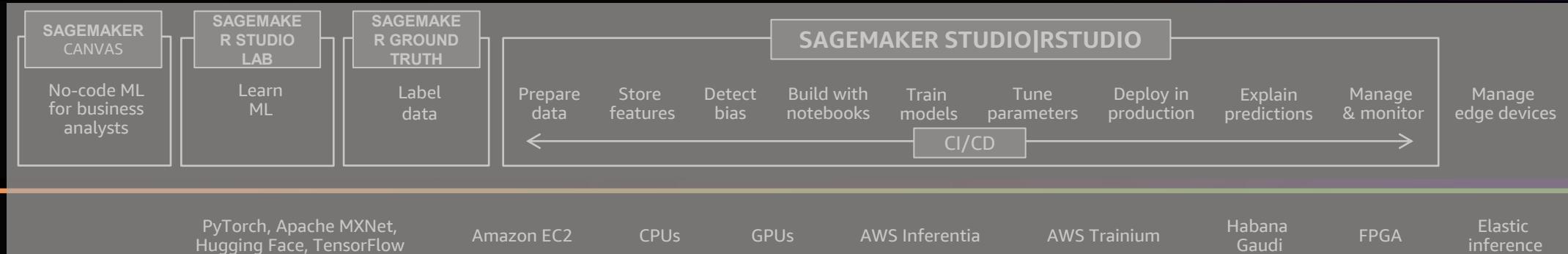
## AI SERVICES



## ML SERVICES



## ML FRAMEWORKS & INFRASTRUCTURE



# Workshop scenarios

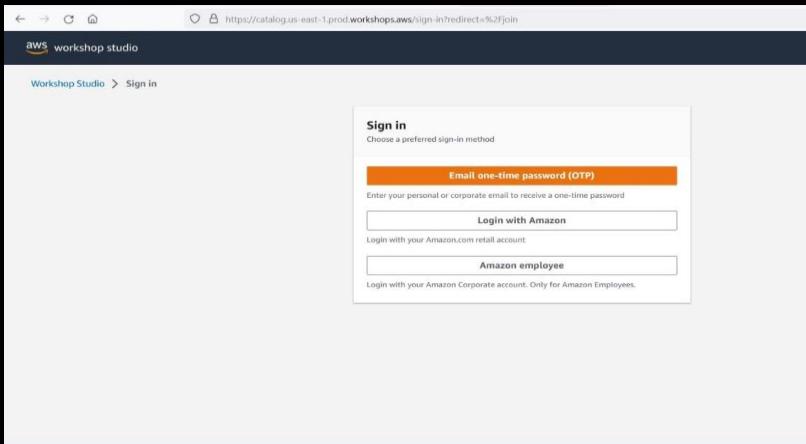
Scenario description	Outcomes
<b>Prerequisites</b>	<ul style="list-style-type: none"><li>Sample documents and website URL for workshop</li><li>Familiarity with AWS AI services</li></ul>
<b>Scenario 1: Standard application processing</b> <ul style="list-style-type: none"><li>User uploads eligibility documents (SSN, DL, passport, utility bill)</li><li>System automatically classifies the document</li><li>Text automatically extracted and stored in a database</li><li>Each document gets a score based on sample business rules</li><li>Auto-adjudication based on a specific application score threshold</li></ul>	<ul style="list-style-type: none"><li>Application successfully processed</li><li>Case worker approves the case – adjudication can be automated as well, if needed</li></ul>
<b>Scenario 2: Incomplete application submission</b> <ul style="list-style-type: none"><li>Document submission with missing fields</li><li>Document submission with face mismatch between DL and passport</li></ul>	<ul style="list-style-type: none"><li>Low application score</li><li>Further review by case worker using human-in-the-loop process</li><li>Case worker denies the application due to potential failure to verify identity of applicant</li></ul>
<b>Technical deep dive</b> <ul style="list-style-type: none"><li>Beneficiary information upload</li><li>Document upload</li><li>Document processing workflow</li></ul>	<ul style="list-style-type: none"><li>Understanding of architecture for each flow</li><li>Understanding of AWS services used in the scenarios</li></ul>
<b>Scenario 3: Automate interview transcription process</b>	<ul style="list-style-type: none"><li>Case worker is relieved of manual note taking tasks</li><li>Text transcript that is further analyzed for sentiment</li><li>Redacted version of the transcript as well to protect sensitive information</li></ul>
<b>Scenario 4: Program leader/benefits administrator dashboard review</b>	<ul style="list-style-type: none"><li>Ability to track program-level metrics using natural language queries</li><li>Data-driven decision-making – forecast enrollment, budgets, and staffing</li></ul>

# Schedule

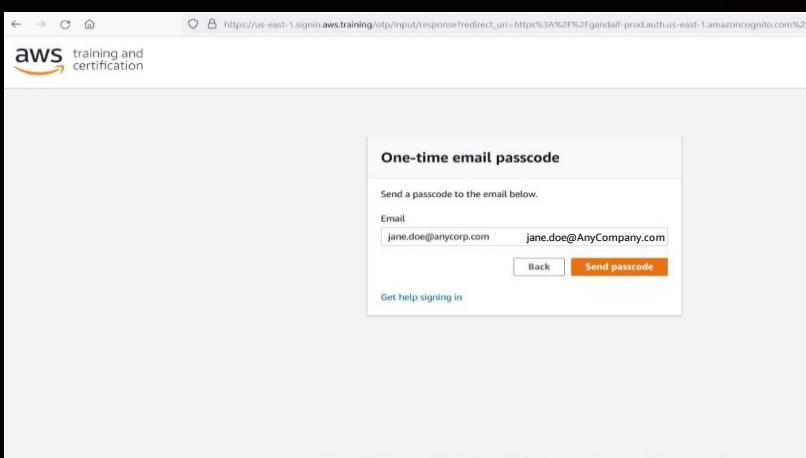
Agenda item	Time
<b>Introduction/overview</b>	<b>20 minutes</b>
<b>Prerequisites</b>	<b>15 minutes</b>
<b>Scenario 1: Standard application processing</b>	<b>20 minutes</b>
<b>Scenario 2: Incomplete application submission</b>	<b>15 minutes</b>
<b>Technical deep dive</b>	<b>15 minutes</b>
<b>Scenario 3: Automate interview transcription process</b>	<b>10 minutes</b>
<b>Scenario 4: Program leader/benefits administrator dashboard review</b>	<b>15 minutes</b>
<b>Wrap up and next steps</b>	<b>10 minutes</b>

# Getting started – Step 1

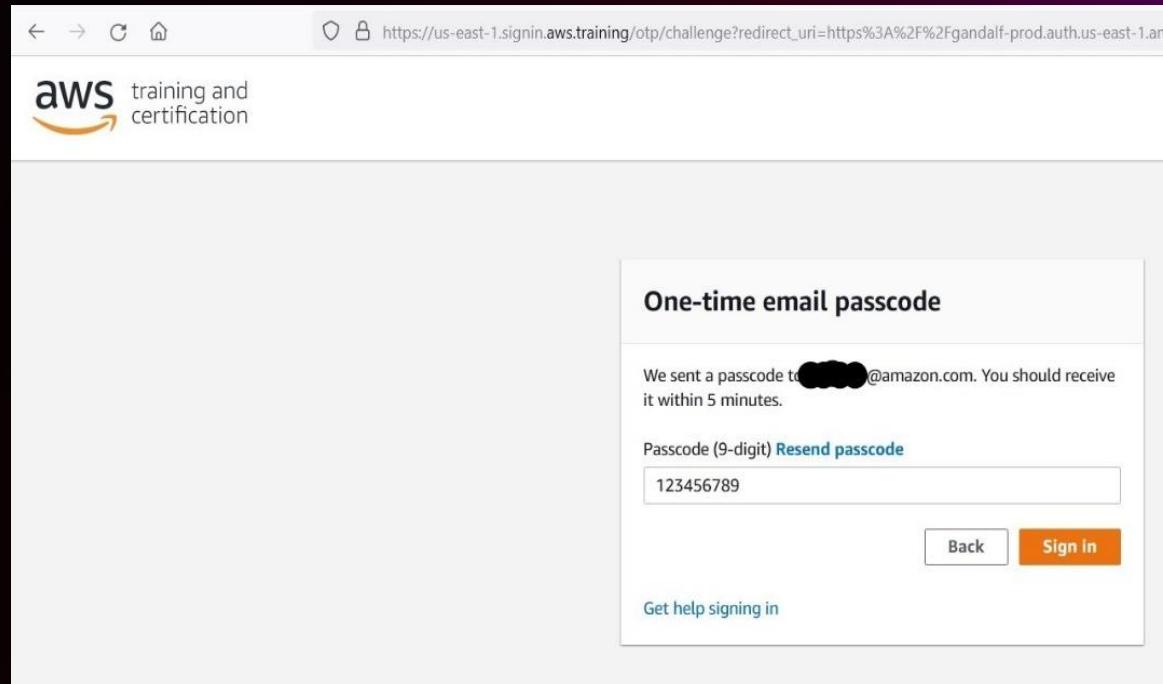
- Go to <https://bit.ly/3u8Btxq>
- Use your email to sign in with one-time Password (OTP)



The screenshot shows the AWS Workshop Studio sign-in page. The URL in the address bar is <https://catalog.us-east-1.prod.workshops.aws/sign-in?redirect=%2Fjoin>. The page title is "aws workshop studio". The main content is a "Sign in" form with the sub-instruction "Choose a preferred sign-in method". It features a prominent orange "Email one-time password (OTP)" button. Below it are three other options: "Login with Amazon", "Login with your Amazon.com retail account", and "Amazon employee". A note at the bottom states: "Login with your Amazon Corporate account. Only for Amazon Employees."



The screenshot shows the "One-time email passcode" page from AWS Training and Certification. The URL in the address bar is [https://us-east-1.siginn.aws.training/otp/input/response?redirect\\_uri=https%3A%2F%2Fgandalf-prod.auth.us-east-1.amazoncognito.com%2Foidc%2Fconnect%2Fcallback](https://us-east-1.siginn.aws.training/otp/input/response?redirect_uri=https%3A%2F%2Fgandalf-prod.auth.us-east-1.amazoncognito.com%2Foidc%2Fconnect%2Fcallback). The page title is "aws training and certification". It displays a message: "Send a passcode to the email below." Below this is a text input field for "Email" containing "jane.doe@anycorp.com" and "jane.doe@AnyCompany.com". At the bottom are "Back" and "Send passcode" buttons, and a "Get help signing in" link.



The screenshot shows the "One-time email passcode" confirmation page. The URL in the address bar is [https://us-east-1.siginn.aws.training/otp/challenge?redirect\\_uri=https%3A%2F%2Fgandalf-prod.auth.us-east-1.amazonaws.com%](https://us-east-1.siginn.aws.training/otp/challenge?redirect_uri=https%3A%2F%2Fgandalf-prod.auth.us-east-1.amazonaws.com%2F). The page title is "aws training and certification". It displays a message: "We sent a passcode to [REDACTED]@amazon.com. You should receive it within 5 minutes." Below this is a text input field for "Passcode (9-digit)" containing "123456789". At the bottom are "Back", "Sign in", and "Get help signing in" buttons.

# Getting started – Step 2

Use the event access code provided to you to join the workshop

*Note that this is part of the bit.ly URL*

The screenshot displays the AWS Workshop Studio interface across three main sections: a landing page, a step 1 join event page, and a step 2 review and join page.

**Landing Page:** Shows the "AWS Workshop Studio" logo and tagline "join hands on events and workshops". It features a "Join an event" callout with a "Get started" button and a "Learning Catalog" section with a "Coming Soon!" message.

**Step 1: Enter event access code:** A modal window titled "Enter event access code" shows a text input field containing the event access code "abcd-012345-ef".

**Step 2: Review and join:** This page contains the "Event details" table and the "Terms and Conditions" section. The "Event details" table includes the following information:

Name	Start time	Duration	Level
sg-test-event-10-15-2022	10/16/2022 12:07 PM	12 hours	-

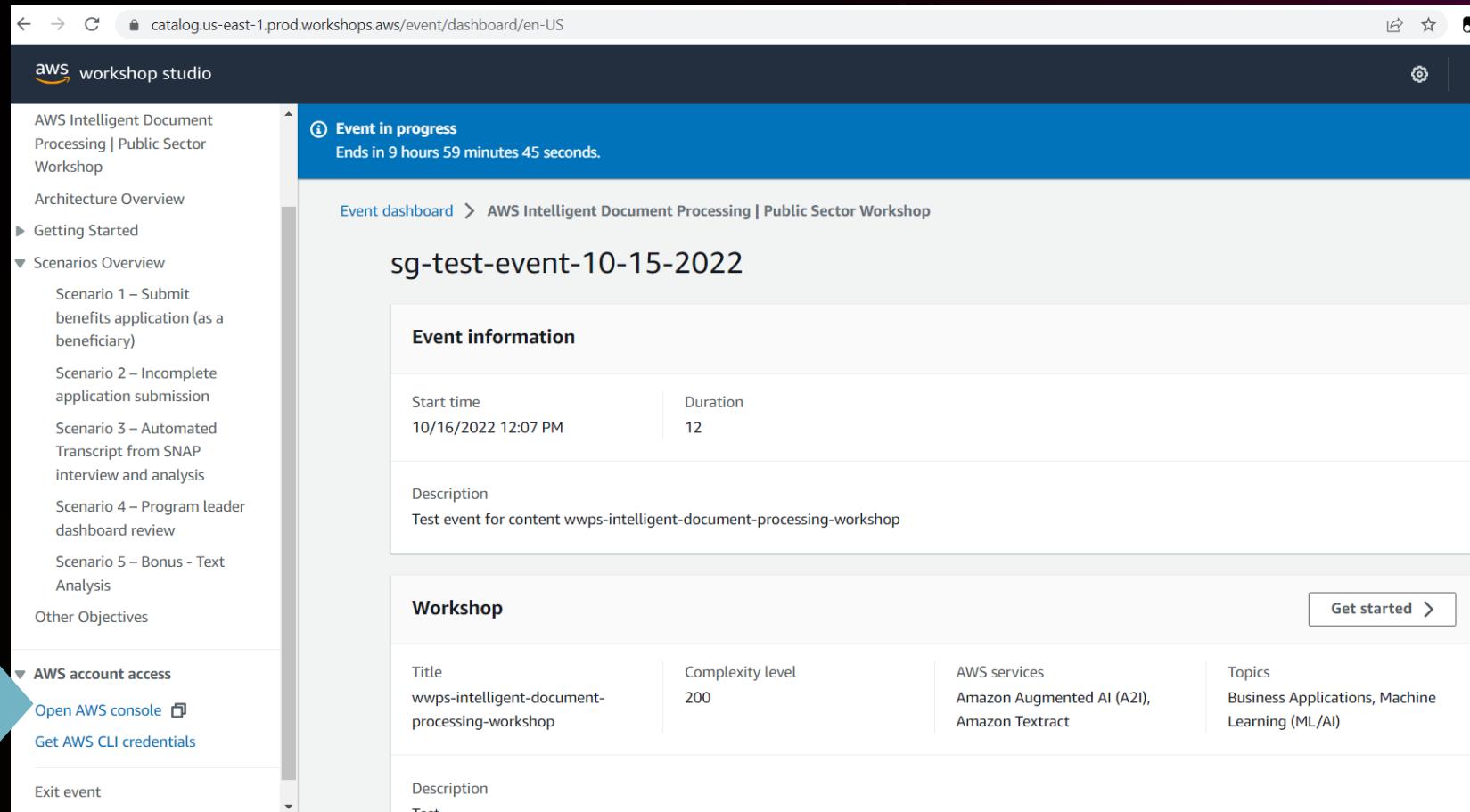
The "Terms and Conditions" section lists four numbered points and a checkbox for agreeing to them. The points are:

1. By using AWS Workshop Studio for the relevant event, you agree to the AWS Event Terms and Conditions and the AWS Acceptable Use Policy. You acknowledge and agree that are using an AWS-owned account that you can only access for the duration of the relevant event. If you find residual resources or materials in the AWS-owned account, you will make us aware and cease use of the account. AWS reserves the right to terminate the account and delete the contents at any time.
2. You will not: (a) process or run any operation on any data other than test data sets or lab-approved materials by AWS, and (b) copy, import, export or otherwise create derivative works of materials provided by AWS, including but not limited to, data sets.
3. AWS is under no obligation to enable the transmission of your materials through Event Engine and may, in its discretion, edit, block, refuse to post, or remove your materials at any time.
4. Your use of AWS Workshop Studio will comply with these terms and all applicable laws, and your access to AWS Workshop Studio will immediately and automatically terminate if you do not comply with any of these terms or conditions.

At the bottom of the page are "Cancel", "Previous", and "Join event" buttons.

# Getting started – Step 3

- Access the AWS Management Console in the left-hand panel



The screenshot shows the AWS Workshop Studio event dashboard for an event titled "sg-test-event-10-15-2022". The left-hand panel is expanded, showing a navigation menu with the following items:

- AWS Intelligent Document Processing | Public Sector Workshop
- Architecture Overview
- ▶ Getting Started
- ▼ Scenarios Overview
  - Scenario 1 – Submit benefits application (as a beneficiary)
  - Scenario 2 – Incomplete application submission
  - Scenario 3 – Automated Transcript from SNAP interview and analysis
  - Scenario 4 – Program leader dashboard review
  - Scenario 5 – Bonus - Text Analysis
- Other Objectives
- ▼ AWS account access
  - Open AWS console
  - Get AWS CLI credentials
- Exit event

A large blue arrow points to the "Open AWS console" link in the AWS account access section of the left-hand panel.

The main content area displays the following information:

**Event in progress**  
Ends in 9 hours 59 minutes 45 seconds.

Event dashboard > AWS Intelligent Document Processing | Public Sector Workshop

**Event information**

Start time	10/16/2022 12:07 PM	Duration	12
Description	Test event for content wwpss-intelligent-document-processing-workshop		

**Workshop**

Title	Complexity level	AWS services	Topics
wwps-intelligent-document-processing-workshop	200	Amazon Augmented AI (A2I), Amazon Textract	Business Applications, Machine Learning (ML/AI)
Description	Test		

**Get started >**

# Scenario description

## Prerequisites

Scenario 1: Standard application processing

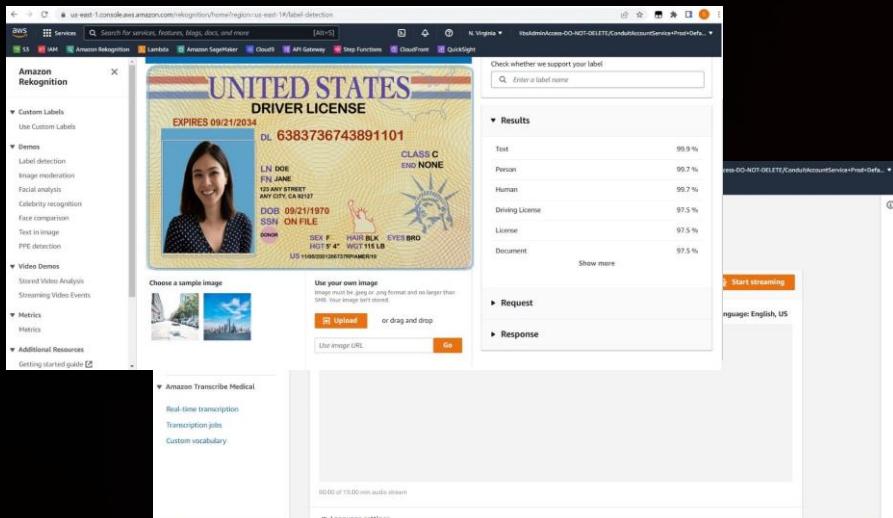
Scenario 2: Incomplete application submission

Scenario 3: Automate interview transcription process

Scenario 4: Program leader dashboard review

Prior to running the workshop scenarios, familiarize yourself with Amazon AI services used in this workshop by following the instructions. Use the AWS Management Console and sample documents.

**IMPORTANT** – Sample documents are already provided to you and are based on **synthetic data** with fictitious names, addresses, etc.; please don't use your own personal info on lab accounts



<https://bit.ly/3u8Btxq>



## Scenario description

### Prerequisites

### Scenario 1: Standard application processing

### Scenario 2: Incomplete application submission

### Scenario 3: Automate interview transcription process

### Scenario 4: Program leader dashboard review

## Objective

You are going to submit a benefits application (SNAP, Medicaid, or others) as a constituent. This application is processed by the state department of human services.

## Core services used

- Amazon Textract
- Amazon Rekognition

## OUTCOMES

- High application score (out of 100)
- Case worker approves the application/potential automation of this approval without any human intervention based on the program and agency policies.
- Applicant gets a notification of approval



## Scenario description

### Prerequisites

Scenario 1: Standard application processing

**Scenario 2: Incomplete application submission**

Scenario 3: Automate interview transcription process

Scenario 4: Program leader dashboard review

## Objective

Submit an application with incomplete documentation/mismatch between documents – say, different photos in identity verification documents

## Core services used

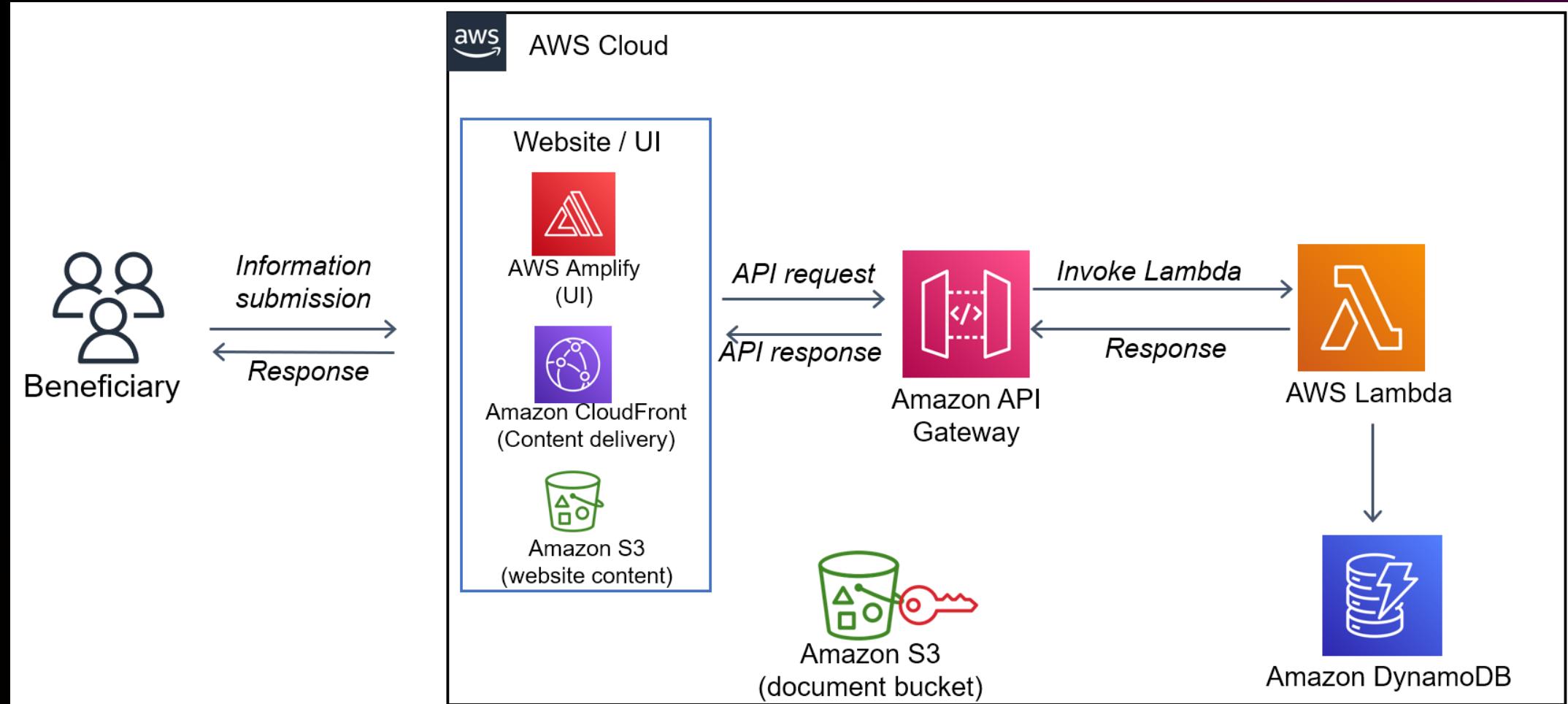
- Amazon Textract
- Amazon Rekognition
- Amazon Augmented AI (A2I)

## OUTCOMES

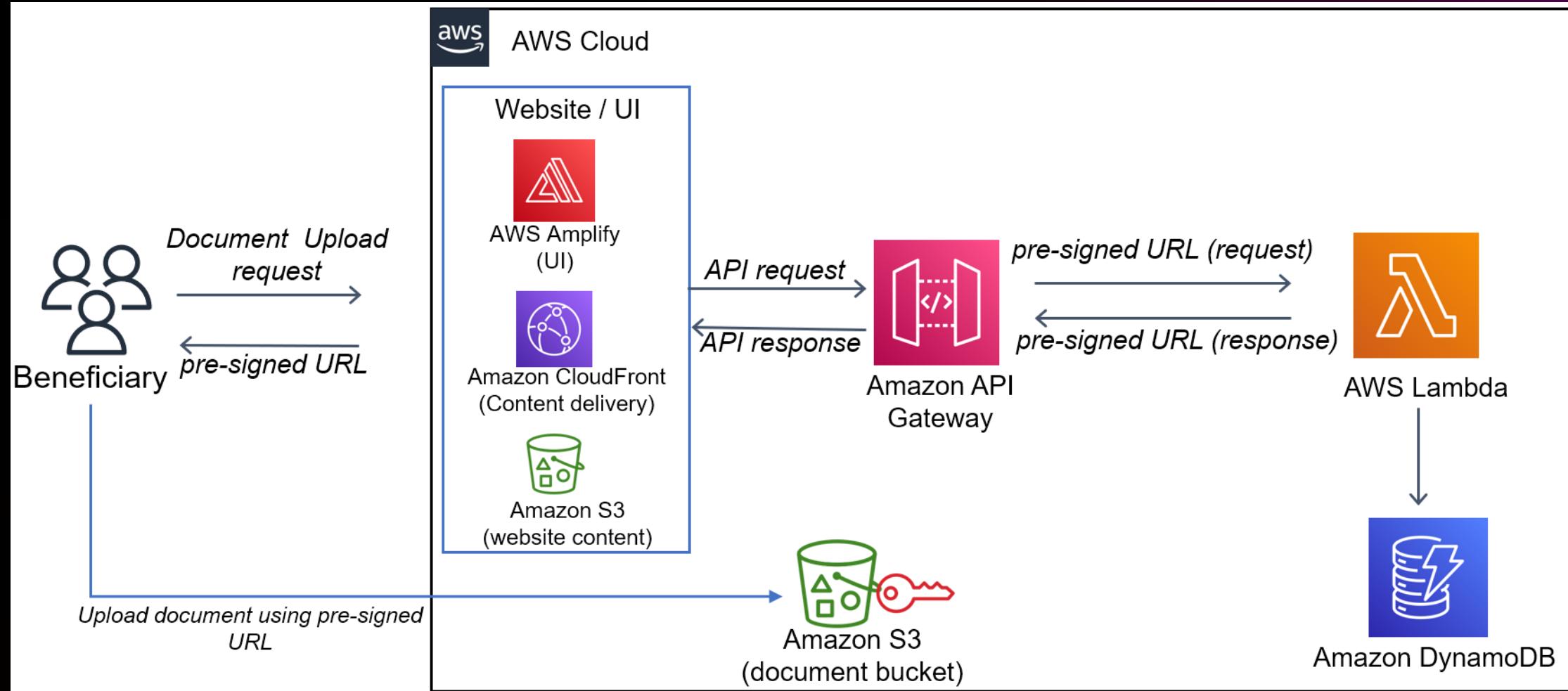
- Low application score
- Further review and potential corrections by case worker using human-in-the-loop (Amazon A2I process)
- Case worker denies the application due to potential failure to verify identity of applicant



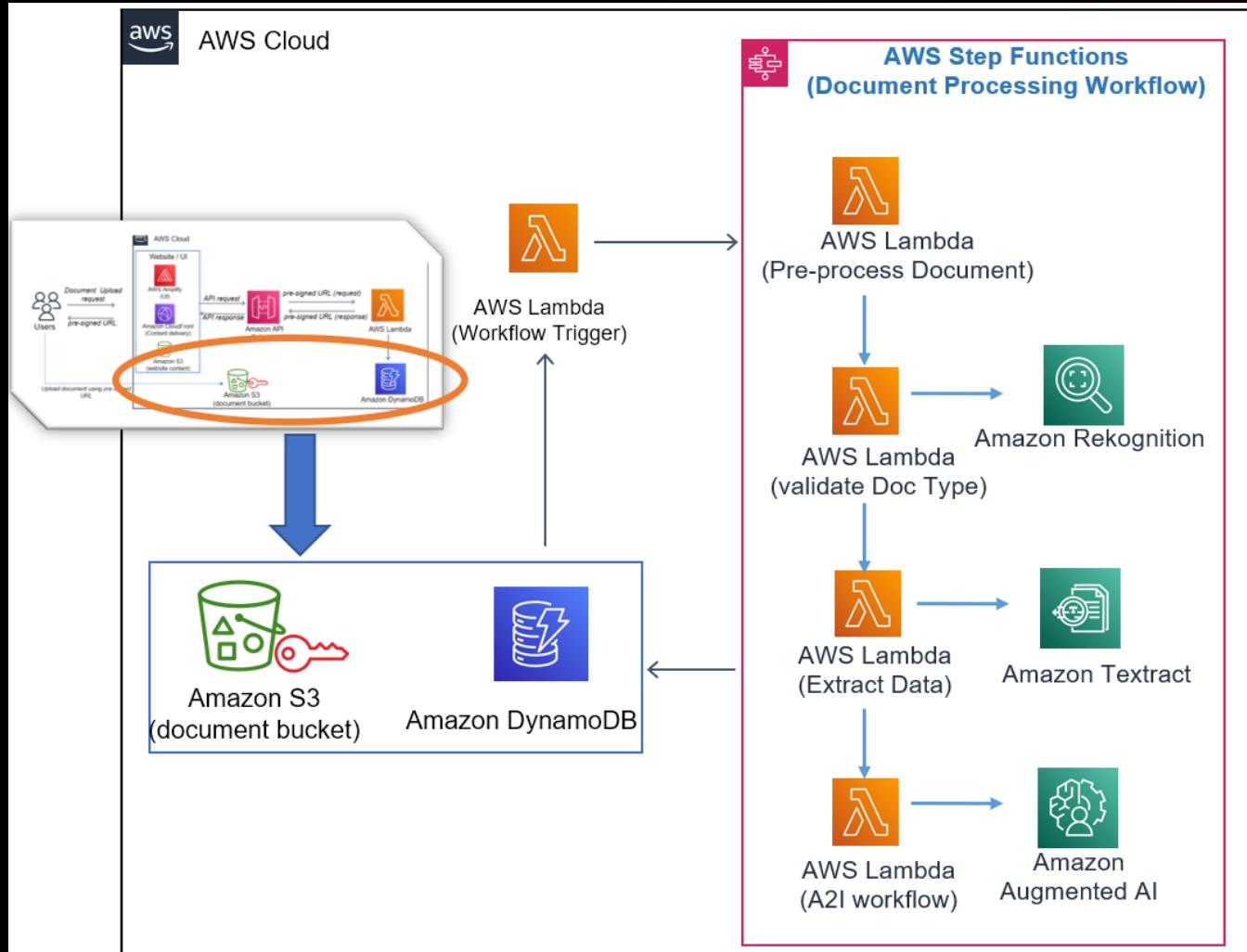
# Technical Deep Dive – Beneficiary information upload



# Technical Deep Dive – Document upload



# Technical Deep Dive – Document processing workflow



## Scenario description

### Prerequisites

Scenario 1: Standard application processing

Scenario 2: Incomplete application submission

**Scenario 3: Automate interview transcription process**

Scenario 4: Program leader dashboard review

## Objective

Automatically transcribe a short interview between the case worker and the constituent to confirm eligibility before benefits are approved

## Core services used

- Amazon Textract
- Amazon Transcribe

## OUTCOMES

- Interview transcript process is automated and the case worker is relieved of manual note taking tasks
- You get a text transcript that can be further analyzed for sentiment
- You get a redacted version of the transcript as well to protect sensitive information, such as SSN, address, etc. – only authorized people have access to the original transcript



# Scenario description

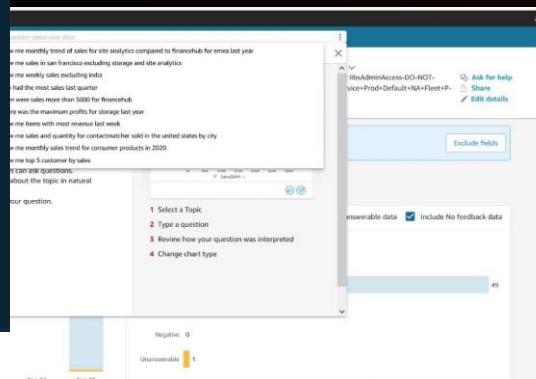
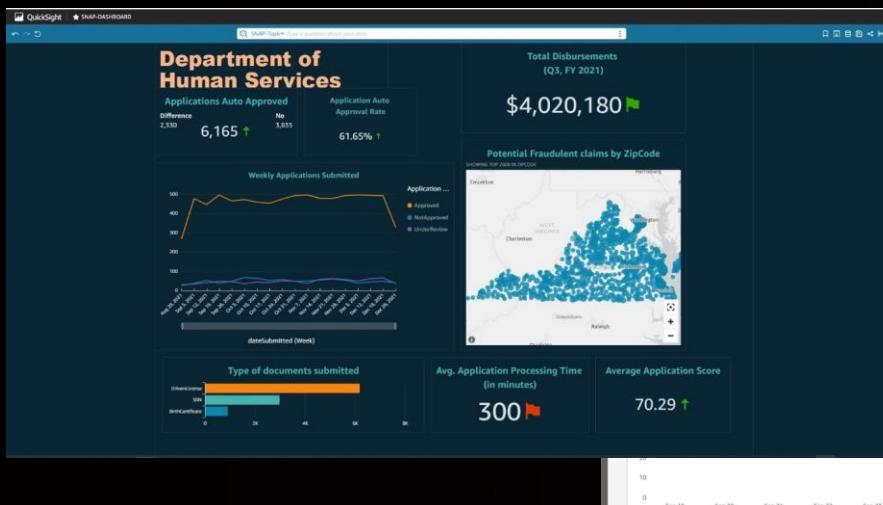
## Prerequisites

### Scenario 1: Standard application processing

### Scenario 2: Incomplete application submission

### Scenario 3: Automate interview transcription process

### Scenario 4: Program leader dashboard review



# Objective

Enable the program leaders, benefits administrators and directors to track program level metrics; helps take program/policy decisions based on deep insights into data.

As an example, these decisions can include staff forecasting, budget or benefits enrollment forecasting

## Core services used

- Amazon QuickSight/QuickSight Q



# Summary and next steps

# We're here to help you with your mission

## AMAZON CUSTOMER RESOURCES

### AWS Professional Services

Skills and experience from AWS experts to supplement your team



### AWS Migration Acceleration Program (MAP)

Use AWS methodology and financial incentives\* to accelerate contact center migrations to the cloud

### AWS support

Go beyond reactive break-fix with proactive support programs



### AWS Training and Certification

Advance your team's skills and build cloud fluency by learning from AWS experts

### AWS Partners

A global network of companies with deep AWS expertise



### Learn from Amazon customer service

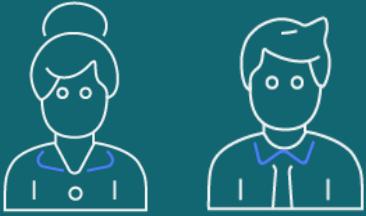
Hear how Amazon's culture of innovation drives our customer experience

# Call to action/next steps



## DISCOVERY WORKSHOP to PROOF OF CONCEPT

- Identify relevant use cases for your mission
- Start a proof of concept



## WHO CAN HELP ?

Work with AWS experts

- AWS Partners
- AWS ProServe
- ML Solution Labs

# Thank you!

Ben Snively

[snivelyb@amazon.com](mailto:snivelyb@amazon.com)

[linkedin.com/in/ben-snively-a10839b/](https://linkedin.com/in/ben-snively-a10839b/)

Srinath Godavarthi

[godavas@amazon.com](mailto:godavas@amazon.com)

[linkedin.com/in/srinathg1/](https://linkedin.com/in/srinathg1/)



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