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

FSI306-R

Build intelligent document processing with AI/ML on AWS

Daniel Sezanoff
Solutions Architect
Amazon Web Services

Pierre Nkola
Solutions Architect
Amazon Web Services

Radhika Gupta
Solutions Architect
Amazon Web Services

Daniela Vargas
Solutions Architect
Amazon Web Services

Madison Paulik
Solutions Architect
Amazon Web Services

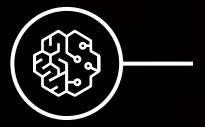
aws

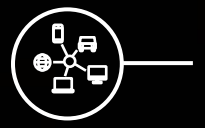
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Agenda

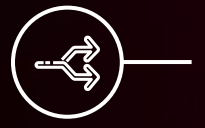

Business need


Benefits of intelligent processing


Services in scope


Pipeline architecture


Workshop


Next steps


References

aws

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

Documents processing across financial services

Customer onboarding & KYC



Streamline customer onboarding document processing

Investment Research



Automate research document processing and creation

Bank loan origination



Automate loan document processing

Credit application



Expedite credit card application processing

Insurance claims processing



Extract data from documents and automate claims processing

Fraud detection



Leverage data from documents to identify anomalous behavior

aws

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

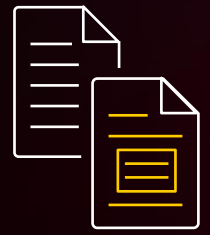
Challenges with manual processing of documents



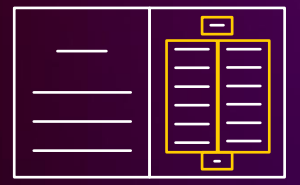
Manual processes do not scale easily with large volumes of documents



Current rules-based systems are not intelligent and break with format changes



Onboarding new documents is effort-intensive with long lead time



Extracting text manually is time-consuming, error prone, and expensive

aws

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

Benefits of intelligent document processing

Go beyond OCR with accurate, versatile information extraction



Analyze documents and send key insights to downstream systems and workflows

Serve end customers faster

Reduce the operational cost of document processing

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

Services in scope

Event-driven services



Amazon S3



AWS Lambda



Amazon DynamoDB

Machine learning services



Amazon Textract



Amazon Comprehend



Amazon Polly

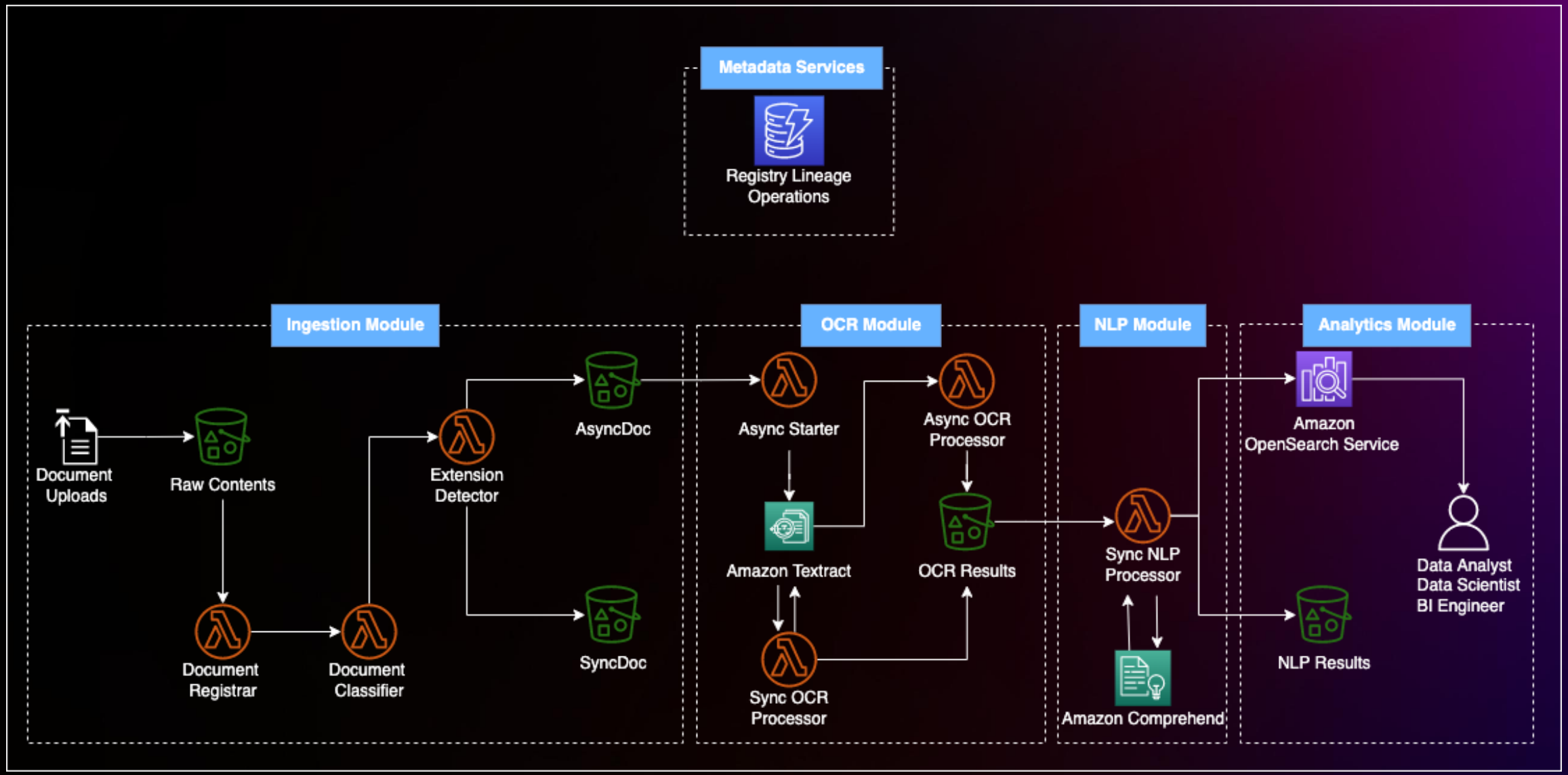
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

What to Expect

Document Upload

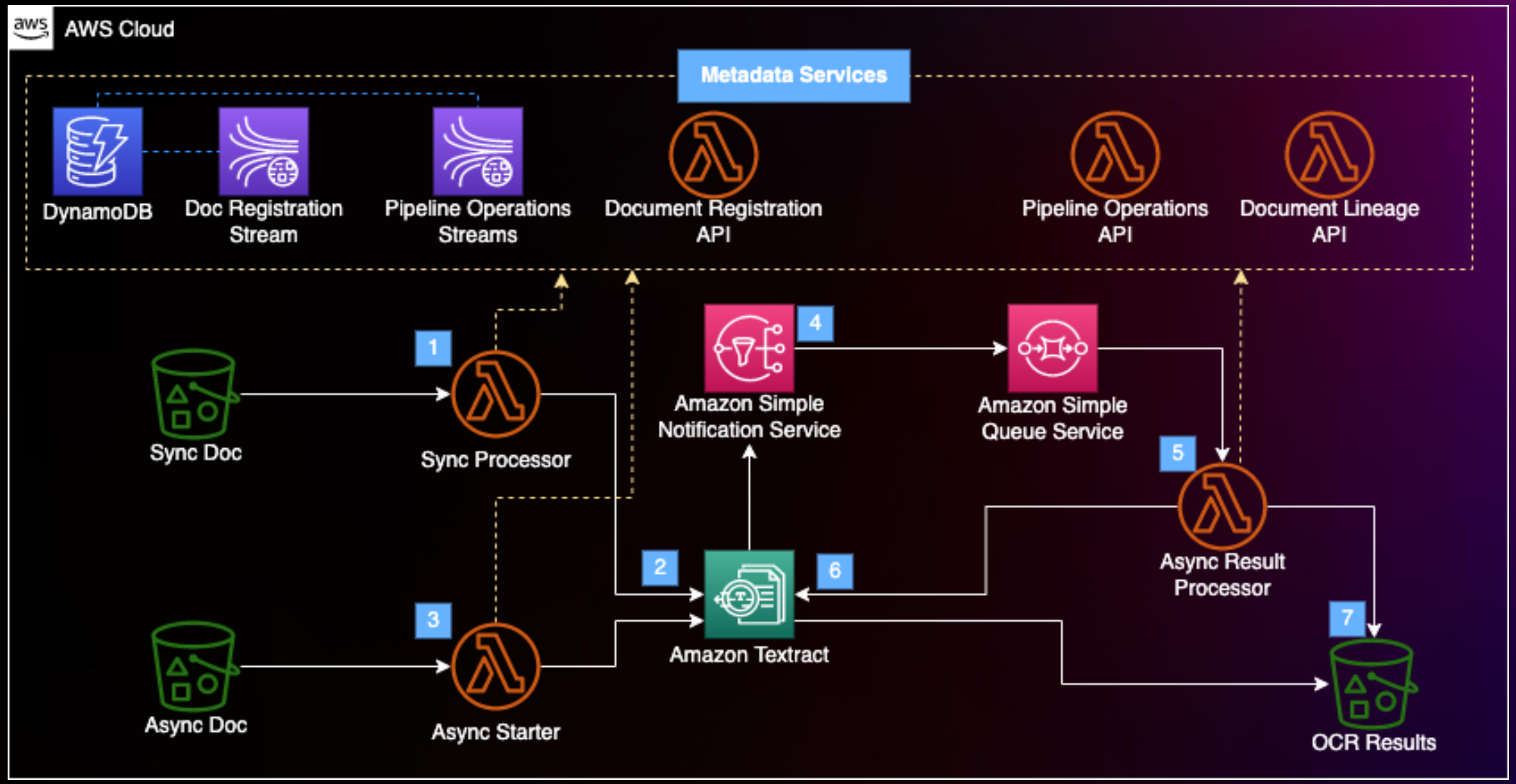

PASSPORT
PASSEPORT
PASSPORTE
UNITED STATES OF AMERICA
Type/Type/Tipo
P
Code/Code/Codigo
USA
Passport No./No.du Passeport / No. de Passporte
918268822
Surname / Nom / Appelidos
Alejandro
Given Names / Prénoms / Nombres
Rosalez
Nationality / Nationalité / Nacionalidad
UNITED STATES OF AMERICA
Date of birth / Date de naissance / Fecha de nacimiento
15 Apr 1990
Place of birth / Lieu de naissance / Lugar de nacimiento
Texas, U.S.A.
Sex / Sexe / Sexo
M
Date of issue / Date de délivrance / Fecha de expedición
29 Jan 2009
Authority / Autorité / Autoridad
United States
Department of State
Date of expiration / Date d'expiration / Fecha de caducidad
31 Jan 2029
Endorsements / Mentions Speciales / Anotaciones
SEE PAGE 27
P<USAROSALEZ< <ALEJANDRO< < < <

Architecture overview



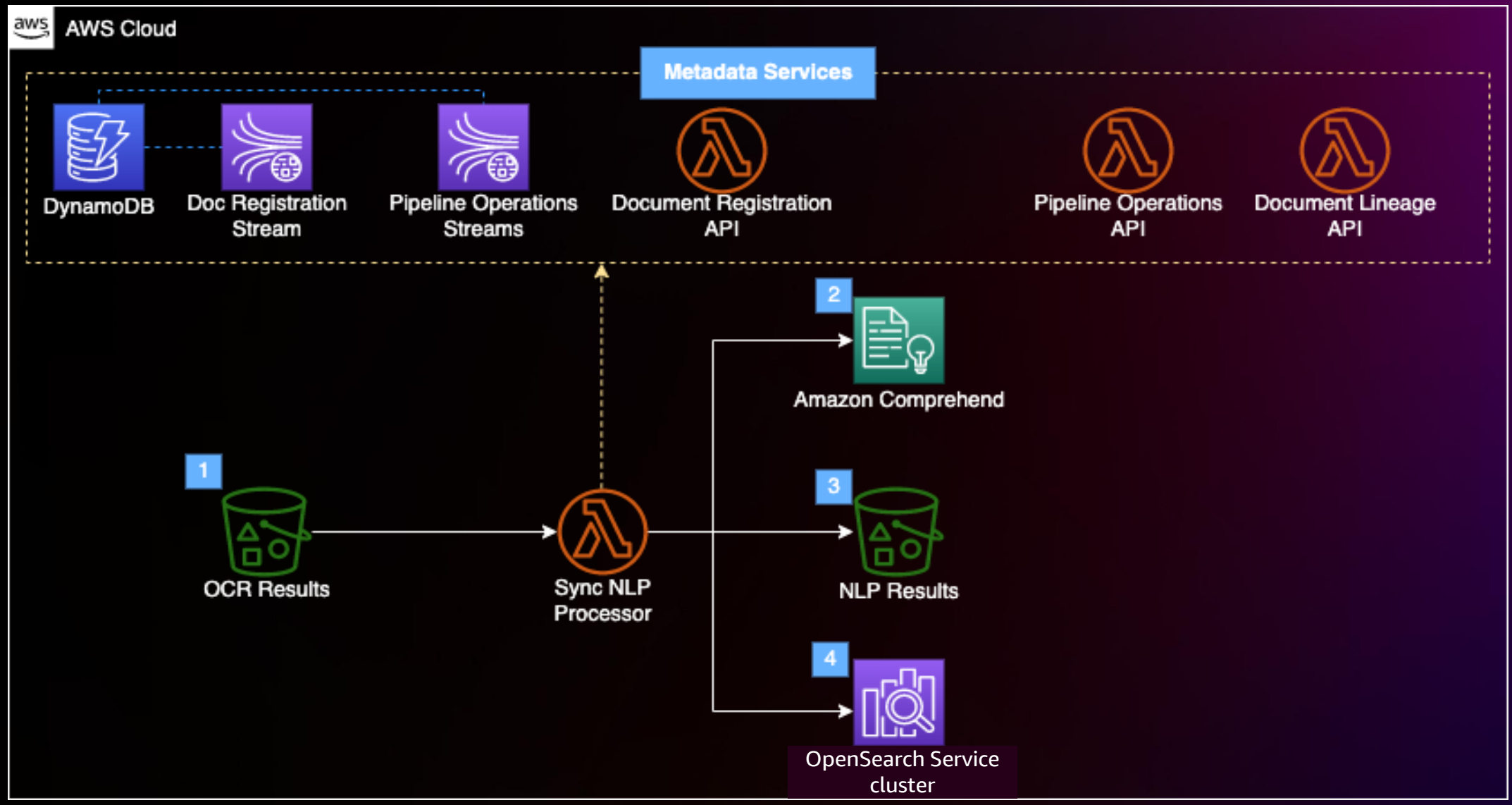
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

OCR module



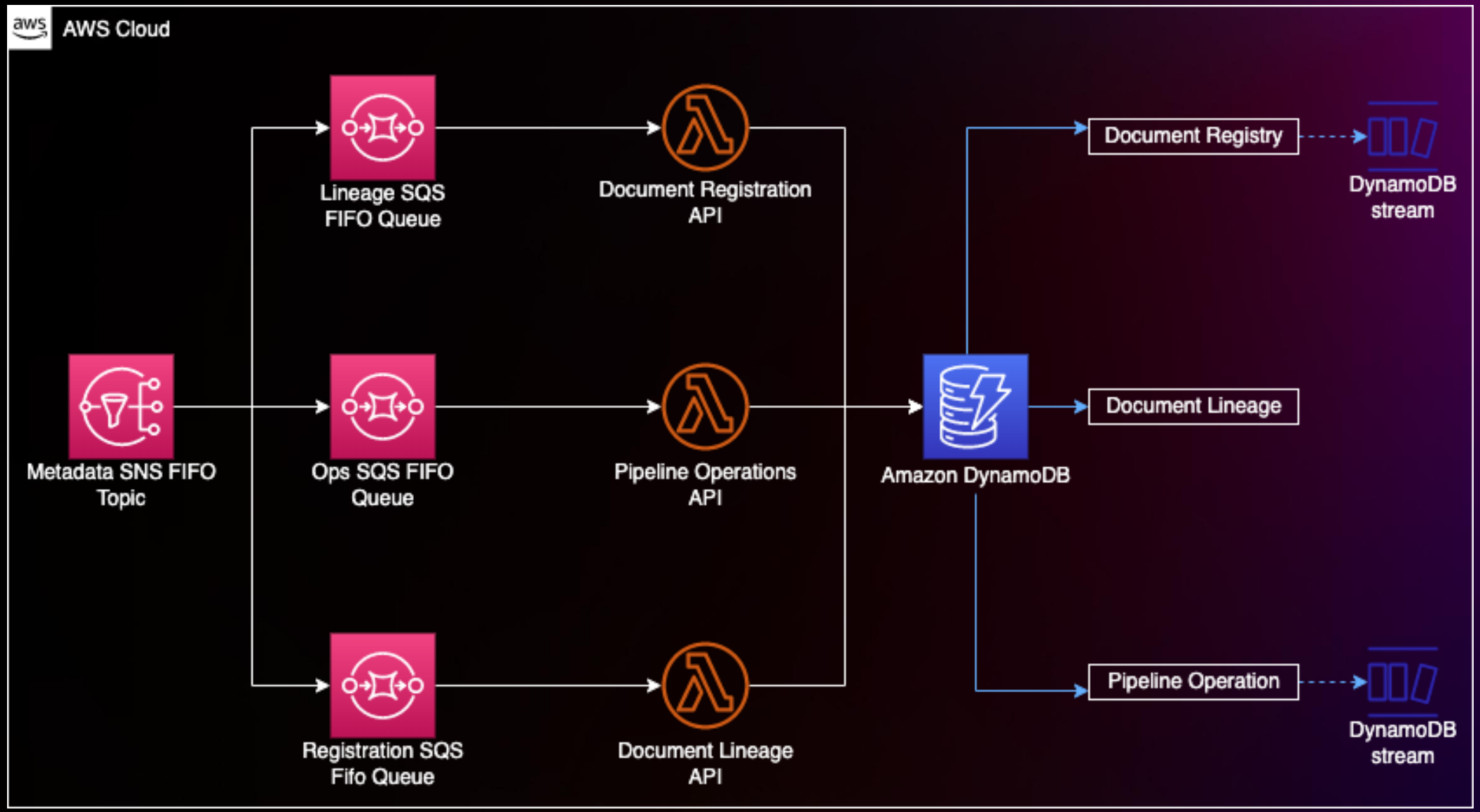
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

NLP module



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

Metadata services



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

Workshop

URL: https://bit.ly/3XF3LwG

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

Next steps

- Use Amazon Polly to convert the text documents to speech
  - Accessibility
  - Adding another step in the pipeline to interact with the Polly API
- Establish PII redaction
- Retrofit the pipeline for your own environment

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

References

- Blogs
  - Intelligent governance of document processing pipelines for regulated industries



- Additional sessions
  - **AIM306**: Automate content moderation and compliance with AI
  - **AIM204:** Automate insurance document processing with AI
  - **AIM202:** Automate your mortgage document processing with AWS AI/ML
  - **FSI205:** How Moody’s uses serverless and microservices for ESG scores

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

Thank you!

Daniel Sezanoff
dsezanof@amazon.com

Pierre Nkola
pknkola@amazon.com

Radhika Gupta
radhigup@amazon.com

Daniela Vargas
awsdaniv@amazon.com

Madison Paulik
mlpaulik@amazon.com



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

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