



# 클라우드에서 AI 탐험기 with AWS

어반데이터랩 강호현



# 목차

---

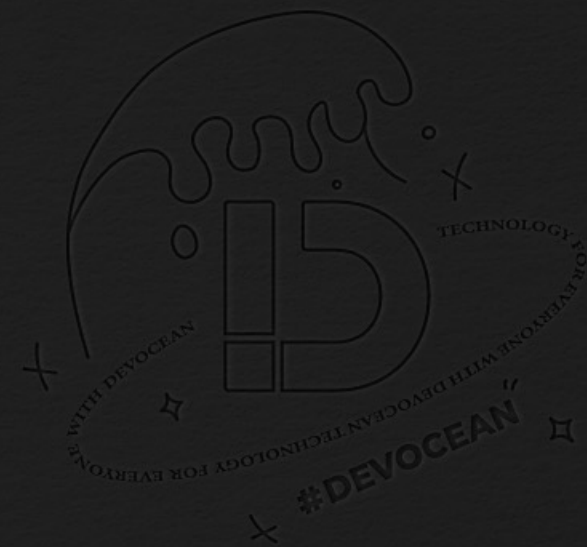
01 스터디 소개

02 AWS Certified MLS-C01 자격증

03 AWS AI 서비스 1: AI 애플리케이션 서비스

04 AWS AI 서비스 2: AI 개발 및 배포 플랫폼

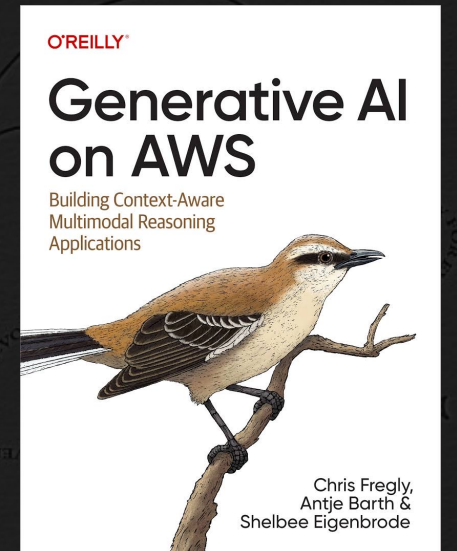
05 마무리



# 01 스터디 소개 - Cloud에서 AI 활용하기

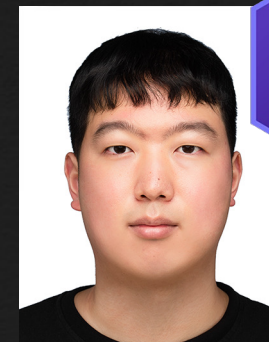
- Cloud에서 AI를 활용해보고자 참여했습니다. AWS 클라우드의 AI 서비스 기능을 사용해보고, 나아가 AWS의 MLS-C01 자격증을 취득해보는 스터디입니다.
- 목표
  - 1) AWS에서 Data를 수집, 처리, 분석하고 AI 모델 적용까지 해보기
  - 2) AWS Certified Machine Learning - Specialty (MLS-C01) 자격증 취득
- 활용 교재
  - 1) AWS Certified Machine Learning - Specialty (MLS-C01) Certification Guide
  - 2) Generative AI on AWS

|         |                                                                                                                                                              |                   |
|---------|--------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------|
| 9월 9일   | AWS Certified Machine Learning - Specialty (MLS-C01)<br>Chapter 1: Machine Learning Fundamentals<br>Chapter 2: AWS Services for Data Storage                 | 박찬병(1장), 김건(2장)   |
| 9월 23일  | Chapter 3: AWS Services for Data Migration and Processing<br>Chapter 4: Data Preparation and Transformation                                                  | 문대한(3장), 조인환(4장)  |
| 10월 7일  | Chapter 5: Data Understanding and Visualization<br>Chapter 6: Applying Machine Learning Algorithms                                                           | 강호현(5장), 박병영(6장)  |
| 10월 21일 | Chapter 7: Evaluating and Optimizing Models<br>Chapter 8: AWS Application Services for AI/ML                                                                 | 강형근(7장) 정경일(8장)   |
| 11월 4일  | Chapter 9: Amazon SageMaker Modeling<br>Chapter 10: Model Deployment                                                                                         | 강민정(9장), 정승환(10장) |
| 11월 18일 | Generative AI on AWS<br>1. Generative AI Use Cases, Fundamentals, and Project Life Cycle<br>2. Prompt Engineering and In-Context Learning                    | 정현진(1장), 오주원(2장)  |
| 12월 2일  | 3. Large-Language Foundation Models<br>9. Context-Aware Reasoning Applications Using RAG and Agents<br>12. Amazon Bedrock: Managed Service for Generative AI | 정경일, 오주원, 박찬병     |

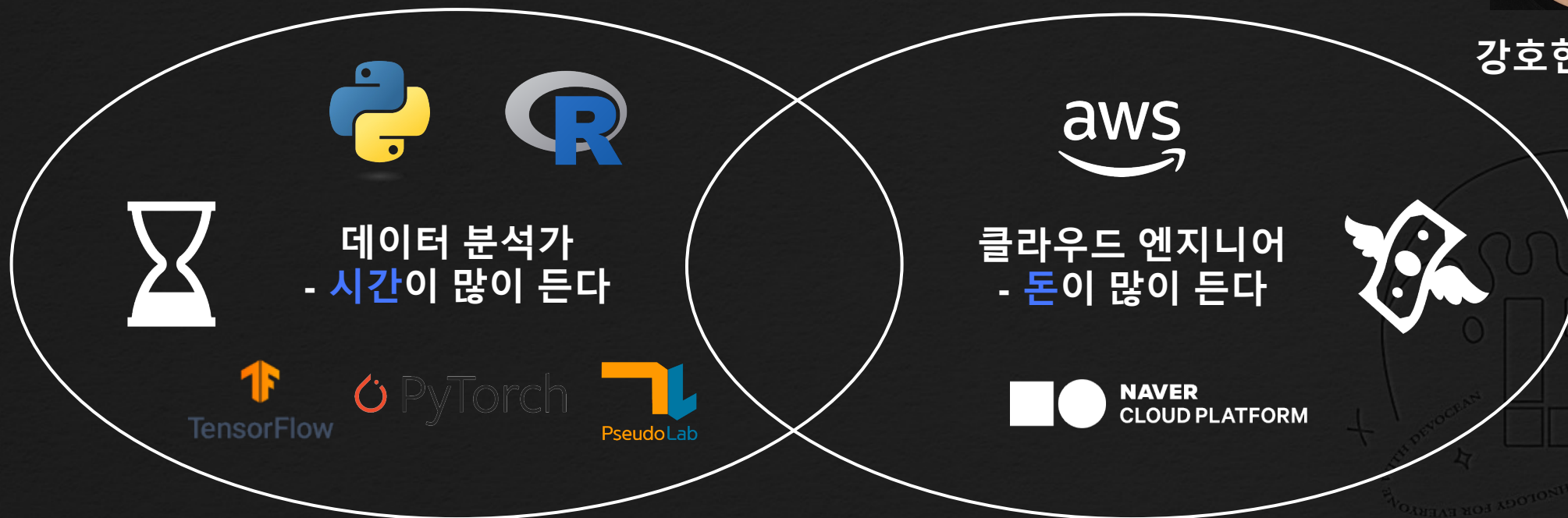


# 01 스터디 소개 - 발표자 소개

- 열정 넘치는 발표자 강호현을 소개합니다.
- 아무도 지원하지 않을 때 도전하는 사람 + MLS-C01 자격 취득한 사람
- 최근 AWS 클라우드 공부하고 있습니다. 문제는 포트폴리오 만들기에는 비쌉니다.
- 하나씩 체험해보는 재미로 살고 있습니다. 돈 쓰는 맛!



강호현 발표자

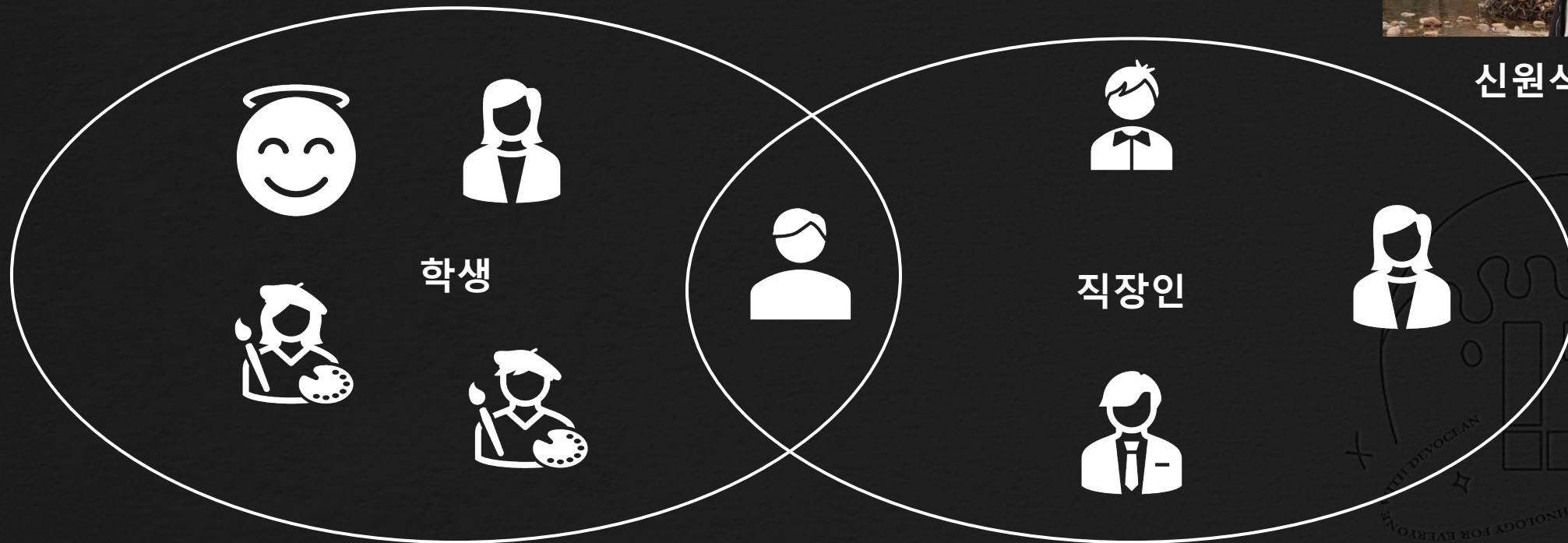


# 01 스터디 소개 - 리더와구성원 소개

- 열정 넘치는 리더 및 구성원 12명을 소개합니다.
- 배경이 아주 다양합니다. 학생, 직장인과 그 사이의 누군가와 함께합니다.
  - 대학생, 대학원생, 데보션 영, 취업준비생, 직장인, 이직준비생
- 모두 **클라우드에서 AI 활용**에 관심을 갖고 참여했습니다.



신원석 리더



## 02 AWS Certified MLS-C01 자격증

- AWS Certified Machine Learning - Specialty
- AWS에서 기계 학습(ML) 모델을 구축, 훈련, 튜닝 및 배포 관련 전문성을 보여줍니다.
- 이 자격증은 비쌉니다. 응시료: 394,575 KRW (약 40만원)
- 그러나 목표로 삼았다면, 그냥 **도전**하는 것입니다. 재미있지 않을까요?



### 시험 개요

#### 시험 개요

등급: Specialty

시간: 시험 완료까지 180분 소요

비용: 300 USD. 환율을 포함한 추가 비용 정보는 [시험 응시료](#)를 참조하세요.

형식: 65개 문항, 다항 선택 또는 복수 응답.

제공 방법: Pearson VUE 테스트 센터 또는 온라인 감독 시험

### AWS 공식 연습 문제 맛보기

기계 학습 - Specialty(MLS-C01)

샘플 시험 문항

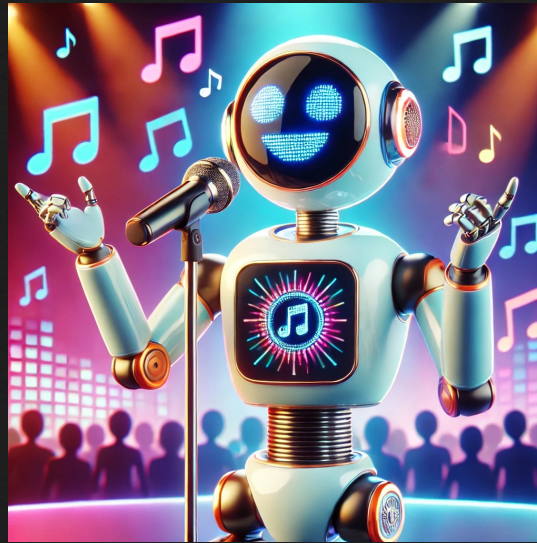
1) 한 기계 학습 팀이 Amazon S3에 여러 개의 대규모 CSV 데이터 집합을 보유하고 있다. 원래는 Amazon SageMaker 선형 학습자 알고리즘으로 구축된 모델에 비슷한 규모의 데이터 집합을 훈련시키는 데 몇 시간이 걸린다. 팀 리더는 훈련 프로세스를 가속화해야 한다.

기계 학습 전문가가 이 문제를 해결하기 위해 할 수 있는 일은 무엇인가?

- A) Amazon SageMaker 파이프 모드를 사용한다.
- B) Amazon Machine Learning 을 사용하여 모델을 훈련시킨다.
- C) Amazon Kinesis 를 사용하여 데이터를 Amazon SageMaker 로 스트리밍한다.
- D) AWS Glue 를 사용하여 CSV 데이터 집합을 JSON 형식으로 변환한다.

### 03 AWS AI 서비스 1 - 생각해보기

- 우리는 AI를 왜 사용할까요?
- 사람이 (     ) 하는 일을 대신해주기 때문입니다.
  - 1) 즉시 사용 가능한
  - 2) 나만의 AI 모델을 개발, 훈련, 배포할 수 있는



TECHNOLOGY FOR EVERYONE WITH DEVOTION  
#DEVOCEAN

# 03 AWS AI 서비스 1 - AI 애플리케이션 서비스

- 오늘은 **말하기** 귀찮아요! > Amazon Polly
- Party Rock : “말하기 귀찮을 때 쓰는 AWS AI 서비스에 대해 알려줘”



Amazon Polly

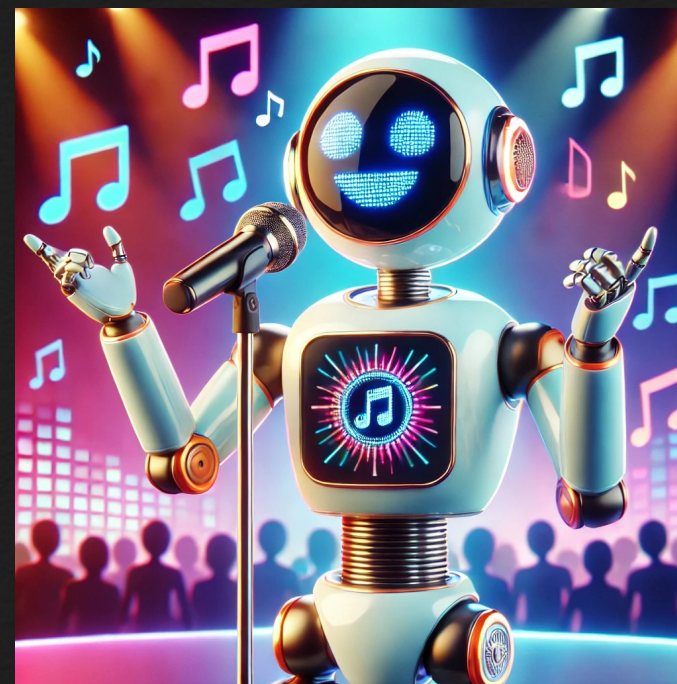
## AWS Expert Response



안녕하세요! 저는 폴리에요. Amazon Polly 서비스의 마스코트입니다.

저는 여러분의 목소리가 되어 드릴 수 있어요. 말하기 귀찮으시죠? 걱정 마세요! 제가 여러분을 대신해 말해드릴게요. 여러분이 텍스트만 입력하시면, 저는 그것을 자연스러운 음성으로 바꿔드립니다.

저는 다양한 언어와 억양을 구사할 수 있어요. 한국어는 물론이고, 영어, 일본어, 중국어 등 여러 나라 말을 할 수 있습니다. 남자 목소리, 여자 목소리, 어린이 목소리 등 다양한 음색으로 말할 수도 있어요.



Amazon Polly > 텍스트 투 스피치

Amazon Polly

텍스트 투 스피치

여러집

S3 합성 작업

텍스트 투 스피치 정보

S3에 저장 다운로드 듣기

엔진 | 정보

음성 | 정보

언어 | 정보

한국어

Seoyeon, 여성

신경망

표준

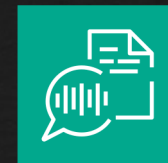
생성형

음성

언어

### 03 AWS AI 서비스 1 - AI 애플리케이션 서비스

- 오늘은 **글쓰기** 귀찮아요! 말하면 글로 써주는 서비스 없나요?
- Party Rock : “말하면 글로 써주는 AWS의 AI 서비스에 대해 알려줘”



Amazon Transcribe

#### AWS Expert Response



안녕하세요! 저는 AWS Transcribe예요. 음성을 텍스트로 바꾸는 AI 비서입니다.

제 특기는 사람들이 말하는 것을 귀 기울여 듣고 그대로 글로 옮기는 거예요. 여러분이 회의나 강의, 인터뷰 등에서 말한 내용을 일일이 받아 적기 귀찮을 때 저를 불러주세요.

작업 설정

이름

transcribetest

이름은 최대 200자가

언어 설정

오디오 파일을 저장한

☒ 특정 언어 정보

☐ 자동 언어 식별

☐ 자동 다국어 식

소스 오디오에서, 언어마다 각기 다.

오디오 파일에서, 이 적용됩니다.

오디오 파일에서, 적은 수의 추가 처

언어

입력 오디오의 언어를

한국 한국어(ko-KR)

작업 세부 정보

이름

transcribetest

상태

☒ 완료

언어

한국 한국어(ko-KR)

언어 설정

특정 언어

파일 정보

트랜스크립션은 89일 더 사용할 수 있습니다.

입력 파일 형식

mp3

오디오 샘플링 속도

24000 Hz

번역할 언어

-

안녕하세요. 저는 AW 블 유 S 트랜스 크 라이브 예요. 음성을 텍스트로 바꾸는 AI 비서 랍니다. 제 특기는 사람들이 말하는 것을 귀 기울여 듣고 그대로 글로 옮기는 거예요. 여러분이 회의나 강의 인터뷰 등에서 말한 내용을 일일이 받아 적기 귀찮을 때 저를 불러주세요.



AI

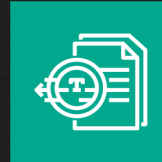
DEVOCATION

### 03 AWS AI 서비스 1 - AI 애플리케이션 서비스

- 일반화를 해봅시다.
- 오늘은 ( ) 하기가 귀찮아요! 대신 ( ) 해주는 서비스 없나요?
- Party Rock : “대신 ( ) 해주는 AWS의 AI 서비스에 대해 알려줘”
- “즉각적인” 서비스가 가능한 도구



“추출해줘”



Amazon Textract

“인식해줘”



Amazon Rekognition  
(Image, Video)



Amazon Comprehend

“통찰해줘”



Amazon Forecast

“예측해줘”  
(신규고객 중단)



Amazon Lex

“답변해줘”



Amazon Translate

“번역해줘”



# 03 AWS AI 서비스 1 - AI 애플리케이션 서비스

안녕하세요 데보선, 저는 Cloud입니다. AnyCompany Financial Services, LLC 신용 카드 계좌 1111-0000-1111-0008의 최소 결제 금액은 \$24.53이며 7월 31일까지 납부해야 합니다. 자동 결제 설정에 따라 라우팅 번호가 있는 은행 계좌 번호 XXXX1111에서 기한 내에 결제 대금을 인출해 드립니다.

선사인 스파, 123 메인 스트리트, 애니웨어에 대한 고객 피드백. 엘리스(sunspa@mail.com)에게 의견을 보냅니다.

저는 스파를 방문하는 것이 즐거웠습니다. 매우 편안했지만 매우 비쌌습니다. 편의시설은 괜찮았지만 서비스 덕분에 스파는 좋은 경험이었습니다.

▼ Results

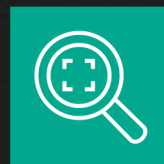
| Entity                             | Type         | Confidence |
|------------------------------------|--------------|------------|
| 데보                                 | Person       | 0.37       |
| 선                                  | Organization | 0.41       |
| Cloud                              | Title        | 0.44       |
| AnyCompany Financial Services, LLC | Organization | 0.99+      |
| 1111-0000-1111-0008                | Other        | 0.99+      |

“추출해줘”



Amazon Textract

“인식해줘”



Amazon Rekognition  
(Image, Video)



Amazon Comprehend

“통찰해줘”



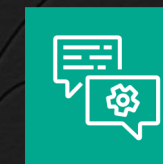
Amazon Forecast

“예측해줘”  
(신규고객 중단)



Amazon Translate

“번역해줘”



Amazon Lex

“답변해줘”



# 03 AWS AI 서비스 1 - AI 애플리케이션 서비스

문서 분석

샘플 문서를 선택하거나 직접 업로드하여 문서 분석 API의 결과를 봅니다.

health\_insurance\_card

AnyInsurance Co.

Jacob Michael

Member id

XZ 9147589652

Effective Date

11/02/2021

Office Visit

\$55/0%

Contract Code

XZ

Specialist Visit

\$65/0%

Rx Be

2035699

Oop Max

\$6000/\$12000

Rx Pcn

XZY2

Coinurance

80%

Plan

050

Dad, Coins and Ov Time Apply

AnyPlan X-EPO

샘플 문서 선택

건강 보험 카드

문서 업로드

문서 선택 또는 끌어서 놓기

결과 다운로드

데모 재설정

원시 텍스트

레이아웃

양식

테이블

쿼리

결과

검색

라인별 세그먼트

AnyInsurance Co.

Jacob Michael

Member id

SILVER

XZ 9147589652

Effective Date

11/02/2021

Office Visit

\$55/0%

Contract Code

XZ

Specialist Visit

\$65/0%

Rx Be

2035699

Oop Max

\$6000/\$12000

Rx Pcn

XZY2

Coinurance

30%

Plan

050

Dad, Coins and Ov Time Apply

MAVA

AnyPlan X-EPO

PPO

Rx

“추출해줘”



Amazon Textract

“인식해줘”



Amazon Rekognition  
(Image, Video)



Amazon Comprehend

“통찰해줘”



Amazon Forecast

“예측해줘”  
(신규고객 중단)



Amazon Lex

“답변해줘”



Amazon Translate

“번역해줘”

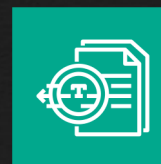


# 03 AWS AI 서비스 1 - AI 애플리케이션 서비스



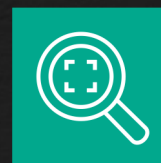
Amazon Comprehend

“통찰해줘”



Amazon Textract

“추출해줘”



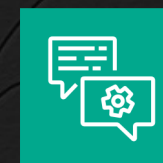
Amazon Rekognition  
(Image, Video)

“인식해줘”



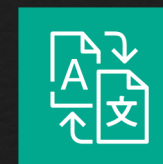
Amazon Forecast

“예측해줘”  
(신규고객 중단)



Amazon Lex

“답변해줘”



Amazon Translate

“번역해줘”



# 03 AWS AI 서비스 1 - AI 애플리케이션 서비스

## 실시간 번역 정보

실시간 번역을 사용하면 입력한 텍스트나 업로드한 문서에 대한 온디맨드 번역을 제공할 수 있습니다.

### 번역

Text

Document

소스 언어

한국어(ko)

대상 언어

영어(en)

텍스트 입력

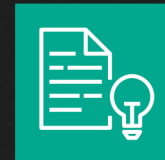
당신은 아마존을 어떻게 사용하십니까?

번역된 텍스트

How do you use Amazon?

20 문자, 52/10000 바이트를 사용했습니다. [정보](#)

▶ 추가 설정



Amazon Comprehend

“통찰해줘”



Amazon Forecast

“예측해줘”  
(신규고객 중단)



Amazon Lex

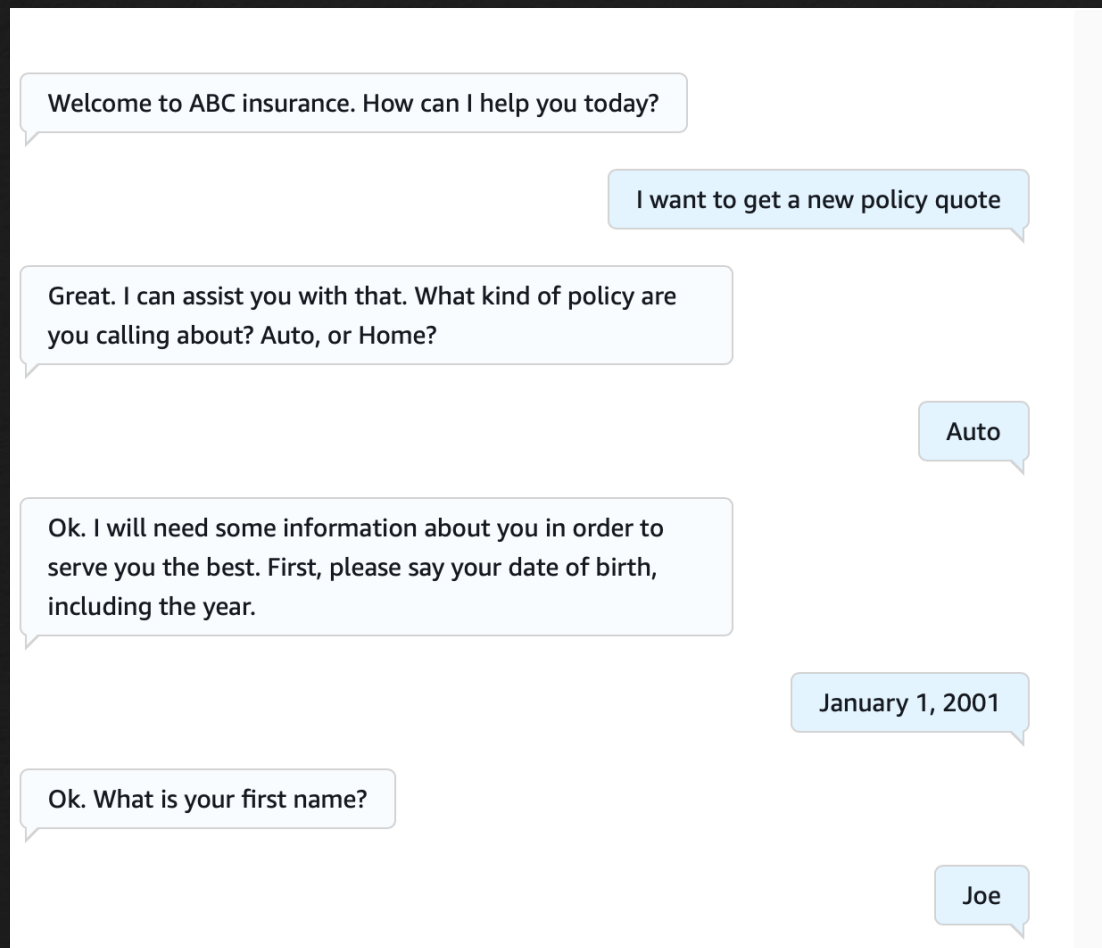
“답변해줘”



Amazon Translate

“번역해줘”

## 03 AWS AI 서비스 1 - AI 애플리케이션 서비스

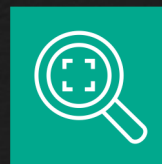


Amazon Comprehend

“통찰해줘”



Amazon Textract



Amazon Rekognition  
(Image, Video)



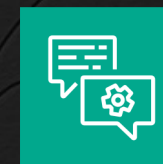
Amazon Translate

“번역해줘”



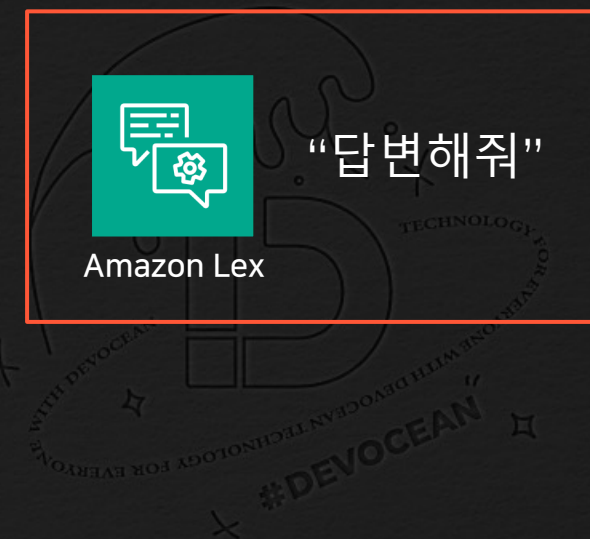
Amazon Forecast

“예측해줘”  
(신규고객 중단)

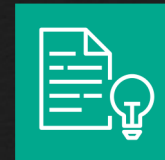


Amazon Lex

“답변해줘”



## 03 AWS AI 서비스 1 - AI 애플리케이션 서비스



Amazon Comprehend

### Transition your Amazon Forecast usage to Amazon SageMaker Canvas

by Nirmal Kumar, Biswanath Hore, Dan Sinnreich, and Davide Gallitelli | on 29 JUL 2024 | in [Amazon Forecast](#), [Amazon SageMaker Canvas](#), [Technical How-to](#) | [Permalink](#) | [Comments](#) | [Share](#)

*After careful consideration, we have made the decision to close new customer access to Amazon Forecast, effective July 29, 2024. Amazon Forecast existing customers can continue to use the service as normal. AWS continues to invest in security, availability, and performance improvements for Amazon Forecast, but we do not plan to introduce new features.*

[Amazon Forecast](#) is a fully managed service that uses statistical and machine learning (ML) algorithms to deliver highly accurate time series forecasts. Launched in August 2019, Forecast predates [Amazon SageMaker Canvas](#), a popular low-code no-code AWS tool for building, customizing, and deploying ML models, including time series forecasting models.

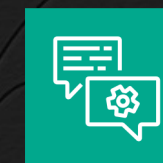
(image, video)

“예측해줘”



Amazon Forecast

“예측해줘”  
(신규고객 중단)



Amazon Lex

“답변해줘”



Amazon Translate

“번역해줘”

## 03 AWS AI 서비스 2 - AI 개발 및 배포 플랫폼

- AI 모델을 구축, 훈련, 배포하고 싶지만, 환경 설정이 귀찮다면?  
"Amazon SageMaker"



Amazon SageMaker

- 1) 도메인 : 팀이 SageMaker 리소스에 액세스할 수 있는 환경
- 2) 완전관리형 : "완전관리형 인프라, 도구 및 워크플로를 사용하여 모든 사용 사례에 대해 데이터를 준비하고 기계 학습(ML) 모델을 구축, 훈련 및 배포하는 완전관리형 서비스입니다."

The screenshot displays the Amazon SageMaker console interface. On the left, a sidebar lists various applications: '애플리케이션', '이름', 'JupyterLab', 'Code Editor', 'Canvas', 'Studio Classic', 'RStudio', and 'MLflow'. The main area shows the 'stats' notebook instance configuration, including buttons for '삭제', '중지', '시작', 'Jupyter 열기', and 'JupyterLab 열기'. Below this, the '노트북 인스턴스 설정' (Notebook Instance Settings) section is visible, showing the instance name 'stats'. The right side of the image shows the JupyterLab interface, which includes a file browser, a code editor, and a launcher for various notebooks like 'conda\_pytorch\_p310', 'conda\_python3', 'conda\_tensorflow2\_p310', 'R', 'Sparkmagic (PySpark)', and 'Sparkmagic (Spark)'.

# 03 AWS AI 서비스 2 - AI 개발 및 배포 플랫폼

- 데이터 없이, AI 모델을 쉽게 구축하려면?

“Amazon SageMaker **JumpStart**”

- “기초 모델, 컴퓨터 비전 모델, 자연어 처리 모델 모두 **사전 학습된** 것을 제공합니다.”
- 아마존이 만든 허깅페이스 모델 허브



Amazon SageMaker

Amazon SageMaker > 기초 모델

### 기초 모델

기초 모델은 대량의 데이터에 대해 사전 학습된 것으로, 텍스트, 이미지 또는 비디오 생성과 같은 광범위한 작업을 수행할 수 있습니다. 이 페이지에는 JumpStart에서 사용할 수 있는 일부 기초 모델이 있습니다. 사용 가능한 모든 모델을 보려면 스튜디오의 JumpStart를 방문하세요. [설명서](#)

#### 기초 모델

**Stable Diffusion XL 1.0**  
기초 Stability AI | Ver 20230726

**PROFESSIONAL: COMPARED TO PREVIOUS VERSIONS, SDXL 1.0 GENERATES MORE VIBRANT...**

The official foundation model for image generation from Stability. Deploy this optimized instance and serve generative AI within minutes.

[보기 모델](#)

**Meta Llama 2 7B Chat**  
기초 Meta | Ver 1.0.0

**CHAT OPTIMIZED, TEXT GENERATION, LLAMA 2**

7B dialogue use case optimized variant of Llama 2 models. Llama 2 is an auto-regressive language model that uses an optimized transformer architecture. Llama 2 is intended for commercial...

[보기 모델](#)

### 작업별 컴퓨터 비전 모델

**Image Classification - TensorFlow**  
기초 TensorFlow Hub

**IMAGE CLASSIFICATION, TENSORFLOW, MOBILENET, RESNET, TRANSFORMERS,...**

Fine-tunable image classification model from TensorFlow hub that classifies an image into one of the 1001 classes.

[보기 작업](#)

### 자연어 처리 모델

JumpStart는 고객이 모델을 더 빠르게 구축할 수 있도록 공개적으로 사용 가능하고 사전 훈련된 자연어 처리(NLP) 모델과 SageMaker 내장 스크립트를 제공합니다. 모델은 공개적으로 사용 가능한 이미지 데이터 세트를 이용하여 사전 훈련됩니다. 입력 데이터로 이미지를 사용하고 작업 유형에 따라 다른 출력을 제공합니다. 고객은 자체 데이터를 사용하여 세부적으로 조정하고 훈련을 거듭하여 사용 사례에 맞춰 모델을 사용자 지정할 수 있습니다. 이 페이지에서는 가장 일반적인 작업이 제공됩니다. [제공되는 알고리즘 및 모델의 전체 목록은](#) 이 설명서 페이지를 참조하세요.

#### 작업별 자연어 처리 모델

**Text Classification - TensorFlow**  
기초 TensorFlow

**TEXT CLASSIFICATION, TENSORFLOW, BERT**

Fine-tunable text classification model built upon text embedding models from TensorFlow Hub. It takes a text string as input and classifies the input text as either a positive or negative movie review...

[보기 작업](#)

**Text Summarization - HuggingFace**  
기초 HuggingFace

**TEXT SUMMARIZATION, HUGGINGFACE, DISTILBART**

Text summarization model from HuggingFace that takes a text string as input and returns a summary of the text.

[보기 작업](#)

**Machine Translation - HuggingFace**  
기초 HuggingFace

Machine Translation model from HuggingFace that takes a text string as input and predicts its translation.

[보기 작업](#)

## 03 AWS AI 서비스 2 - AI 개발 및 배포 플랫폼

- 데이터 없이, AI 모델을 쉽게 구축하려면?

“Amazon SageMaker **JumpStart**”



Amazon SageMaker

```
1 from sagemaker.jumpstart.estimator import JumpStartEstimator
2 from datasets import load_dataset
3 # Load dataset and convert each row to an instruction prompt
4 dataset = load_dataset(...)
5 dataset = dataset.map(convert_row_to_instruction)
6 # Define and load the model for fine-tuning
7 model_checkpoint = "<choose a model>"
8 # Save training data to a local file to be uploaded to s3
9 local_data_file = "train.jsonl"
10 dataset.to_json(local_data_file)
11 # Specify S3 location and upload the local dataset file
12 train_data_s3_location = "s3://<your-private-s3-location>/"
13 S3Uploader.upload(local_data_file, train_data_s3_location)
14 # Configure the estimator including instance type and count
15 estimator = JumpStartEstimator(
16     model_id=model_checkpoint,
17     instance_type="ml.p4de.24xlarge",
18     instance_count=2, # increase this value for a larger cluster
19     # Set the hyper-parameters including instruction_tuned="True"
20     estimator.set_hyperparameters(
21         instruction_tuned="True"
22     )
23 # Specify S3 location of training data and start fine-tuning!
24 estimator.fit({"training": train_data_s3_location})
```

```
14 # Configure the estimator including instance type and count
15 estimator = JumpStartEstimator(
16     model_id=model_checkpoint,
17     instance_type="ml.p4de.24xlarge",
18     instance_count=2, # increase this value for a larger cluster
19     # Set the hyper-parameters including instruction_tuned="True"
20     estimator.set_hyperparameters(
21         instruction_tuned="True"
22     )
23 # Specify S3 location of training data and start fine-tuning!
24 estimator.fit({"training": train_data_s3_location})
```



## 04 AWS AI 서비스 2 - AI 개발 및 배포 플랫폼

- 생성형 AI 플랫폼을 만들고 싶은데, 어떻게 해야하죠?

“Amazon Bedrock”



Amazon Bedrock

- 가장 쉽게 쓰는 법: 모델 액세스 > 승인 > 플레이그라운드(chat/text) > 모델 선택 > 채팅하기
- 그 외에도 오케스트레이션(프롬프트, 에이전트), 세이프가드(가드레일), 추론, 평가 및 배포가 콘솔에서 가능합니다.

모델 액세스란 무엇입니까?

Bedrock을 사용하려면 올바른 IAM 권한 [\(클\)](#) 보유한 계정 사용자가 사용 가능한 Bedrock 파운데이션 모델(FM)에 대한 액세스를 설정. [Bedrock 모델 용어](#)에 대한 모든 Bedrock FM을(를) 보세요.

[모든 모델 활성화](#) [특정 모델 활성화](#)

Visit [Amazon Bedrock Quotas](#) for a quick guide to the default quotas and limits that apply to Amazon Bedrock.

기반 모델 (4)

Not seeing a model you're interested in? Check out all supported models by region [여기](#).

| 모델                                                     | 액세스 상태                |
|--------------------------------------------------------|-----------------------|
| ▼ Amazon (1)                                           | 0/1 access granted    |
| Titan Text Embeddings V2                               | <a href="#">요청 가능</a> |
| ▼ Anthropic (3)                                        | 0/3 access granted    |
| Claude 3.5 Sonnet                                      | <a href="#">요청 가능</a> |
| Claude 3 Sonnet <a href="#">Cross-region inference</a> | <a href="#">요청 가능</a> |
| Claude 3 Haiku                                         | <a href="#">요청 가능</a> |

**AI Claude 3.5 Sonnet v1** [정보](#)

온디맨드

▼ 시스템 프롬프트 [정보](#)

당신은 AWS AI 서비스의 전문가입니다. 한국어로 설명합니다.

▼ 무작위성과 다양성 [정보](#)

온도

Amazon Bedrock은 AWS의 완전 관리형 서비스로, 다양한 기반 모델(foundation models, FM)을 API를 통해 쉽게 사용할 수 있게 해주는 서비스입니다. 특히 서울 리전에서 사용할 수 있는 기능들을 강조하여 설명드리겠습니다.

- 서울 리전 지원:  
Amazon Bedrock은 서울 리전(ap-northeast-2)에서 사용 가능합니다. 이는 한국 기업들이 데이터 주권과 지연 시간 요구사항을 충족하면서 AI 서비스를 활용할 수 있음을 의미합니다.
- 사용 가능한 모델:  
서울 리전에서는 다음과 같은 모델들을 사용할 수 있습니다:
  - Amazon Titan 텍스트 모델
  - Anthropic의 Claude 모델
  - Stability AI의 Stable Diffusion 모델
- 한국어 지원:  
서울 리전에서 사용 가능한 모델들은 한국어를 포함한 다국어를 지원합니다. 이를 통해 한국 기업들은 자연스러운 한국어 텍스트 생성 및 처리가 가능합니다.

Hallucination



# 04 AWS AI 서비스 2 - AI 개발 및 배포 플랫폼

- 생성형 AI 플랫폼을 만들고 싶은데, 어떻게 해야하죠?  
"Amazon Bedrock"
- 리전마다 가능한 모델이 다릅니다.
- 서울 리전 : 2024년 10월 2일자로 사용 가능합니다. Titan 1개, Claude 3개
- 버지니아 북부 리전 : AI21 Labs Jamba 5개, Titan 8개 등 많습니다.



Amazon Bedrock

Amazon Bedrock, 이제 아시아 태평양(서울) 및 미국 동부(오하이오) 리전에서 사용 가능

게시된 날짜: 2024년 10월 2일

오늘부터 고객은 아시아 태평양(서울)과 미국 동부(오하이오) 리전에서 Amazon Bedrock을 통해 다양한 파운데이션 모델(FM)과 함께 생성형 AI 애플리케이션 구축을 위한 강력한 도구를 사용하여 생성형 AI 애플리케이션을 손쉽게 구축하고 규모를 조정할 수 있습니다.

## 서울

|                                                                                                      |                                                                                                          |
|------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------|
|  Titan<br>Amazon 기준 |  Claude<br>Anthropic 기준 |
|------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------|

## 버지니아 북부

|                                                                                                               |                                                                                                                |                                                                                                                           |                                                                                                          |
|---------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------|
|  Jamba 1.5<br>AI21 Labs 기준 |  Titan<br>Amazon 기준         |  Claude<br>Anthropic 기준                |  Command<br>Cohere 기준 |
|  Llama<br>Meta 기준        |  Mistral<br>Mistral AI 기준 |  Stable Diffusion<br>Stability AI 기준 |                                                                                                          |



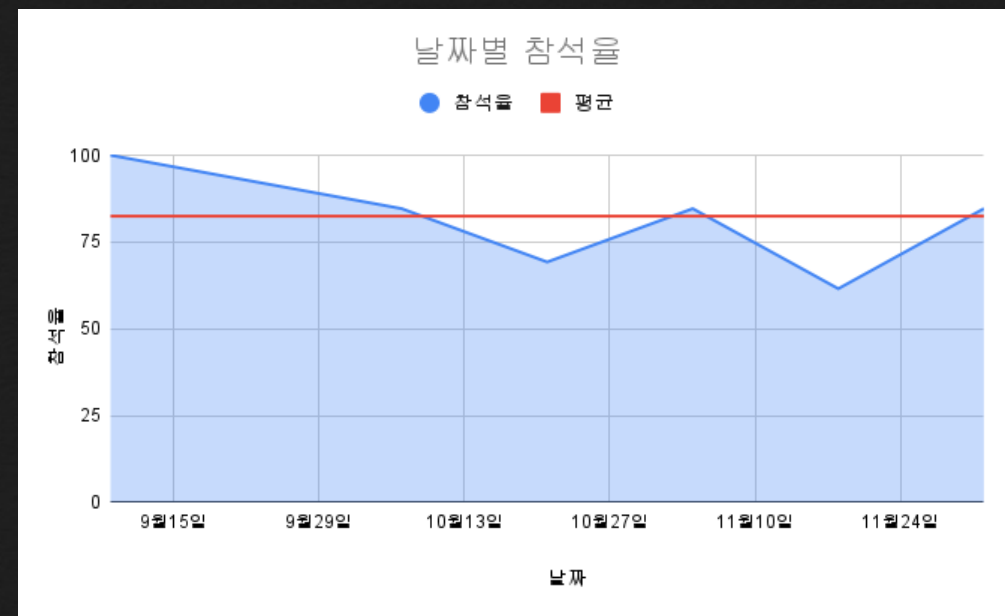
## 05 마무리

### 후기

- MLS 자격증 취득하니 뿌듯합니다.
- 다음에는 돈을 많이 벌어서 여러 기능을 마음껏 써보고 싶습니다. 돈 쓰는 맛이 있습니다.
- API를 활용한 다양한 생성형 AI 어플리케이션을 만들어볼 예정입니다.
- 차후 Machine Learning을 사용하는 분야로 가고 싶습니다.

### 성과

- 스터디 참석율이 매우 우수 (평균 82.05%)
- 스터디 리더님의 자격증 취득 성공
- 강호현, 강형근 님의 자격증 취득 성공
- 다른 멤버 분들도 자격증 응시 예정



## 05 마무리

### 깜짝 퀴즈 (6 문제)

- 재미있는 퀴즈를 준비했습니다. 5등까지 상품을 드립니다.
    - 귀여운 스티커와 무언가를 준비했습니다.
    - 랜덤 선물은 발표 후 등수대로 하나씩 가져가시면 됩니다.
    - 양보 가능합니다.
  - 상품
    - 1등: 커피 기프티콘 + AWS 큐브, 스티커, 선물
    - 2~5등 : 커피 기프티콘 + 선물
- \* 기프티콘: 오픈 카톡방(참여코드: mls1206) 에서 지급 (인증샷)



## 05 마무리

---

감사합니다

