

# Amazon SageMaker Ground Truth

## 1. Create workforce

**Amazon SageMaker**

Amazon SageMaker Studio

대시보드

검색

Ground Truth

- 레이블링 작업
- 데이터 세트에 레이블 지정
- 레이블링 인력

노트북

- 노트북 인스턴스
- 수명 주기 구성
- Git 리포지토리

처리 중

- 작업 처리 중

훈련

- 알고리즘
- 훈련 작업
- 하이퍼파라미터 튜닝 작업

추론

- 컴파일 작업
- 모델 패키지
- 모델

의견 한국어

Amazon SageMaker > 레이블 지정 작업 인력 > 프라이빗 팀 만들기

**프라이빗 팀 만들기**

**Private team creation**

Create a private team with AWS Cognito  
Create a private work team by sending email invitations to new workers or importing workers from existing Amazon Cognito user groups.

Create a private team with OpenID Connect (OIDC)  
Create a private work team with your own identity provider (IdP). Your IdP must support OIDC user groups.

**팀 세부 정보**

팀 이름  
작업팀에 알아보기 쉬운 이름을 지정합니다. 이 이름은 나중에 변경할 수 없습니다.

최대 63자의 알파벳 문자 및 숫자, 하이픈을 포함할 수 있지만 공백은 포함할 수 없습니다. AWS 리전의 계정에서 고유해야 합니다.

**작업자 추가 정보**

작업자 이메일 주소를 추가하거나, 기존 Amazon Cognito 사용자 그룹에서 작업자를 가져와서 프라이빗 작업팀에 작업자를 추가합니다.

이메일로 새 작업자 초대  
기존 Amazon Cognito 사용자 그룹에서 작업자 가져오기

**이메일 주소**

초대장 미리 보기

© 2008 - 2020, Amazon Web Services, Inc. 또는 계열사. All rights reserved. 개인 정보 보호 정책 이용 약관

- SageMaker의 Labeling workforce에서 Create private team을 선택해 줍니다.
- 각 항목들을 입력하여 private 팀을 생성해 줍니다.

**프라이빗 작업 인력 요약 정보**

작업 인력을 Amazon Cognito 사용자 풀에서 지원합니다. 사용자 풀을 사용하여 작업자는 Amazon Cognito를 통해 레이블 지정 포털에 로그인할 수 있습니다.

Amazon Cognito 사용자 풀  
us-east-1\_8OqZg77p8 [ ]

액 클라이언트  
1d3272645r7ej94hr18uaedque [ ]

레이블 지정 포털 로그인 URL  
<https://rgai4phalo.labeling.us-east-1.sagemaker.aws> [ ]

**프라이빗 팀 (1) 정보**

프라이빗 작업 인력의 작업 팀입니다. 레이블 지정 작업을 위해 여러 프라이빗 팀을 만들 수 있습니다.

작업자	삭제	프라이빗 팀 만들기			
이름	▼	ARN	▼	생성 시간	▼
private-team-1	arn:aws:sagemaker:us-east-1:556463194331:workteam/private-crowd/private-team-1	2020년 7월 29일 오전 8:57 UTC			

**작업자 정보**

프라이빗 작업 인력의 모든 작업자

작업자	활성화	비활성화	삭제	새 작업자 초대					
작업자 검색	< 1 ... >	[ ]							
이메일	▲	상태	▼	Cognito 상태	▼	활성	▼	사용자 이름	▼
ggamini@hmail.hanyang.ac.kr		초대장 보냄		Force_change_password	예	ggamini@hmail.hanyang.ac.kr			

<https://console.aws.amazon.com/sagemaker/groundtruth?region=us-east-1#>

© 2008 - 2020, Amazon Web Services, Inc. 또는 계열사. All rights reserved. 개인 정보 보호 정책 이용 약관

- private team을 생성하면 다음과 같이 나옵니다.
- 메일로 들어가서 임시 password로 authorize를 해 줍니다.

The screenshot shows the AWS SageMaker Ground Truth interface. On the left, a sidebar lists various options like 'Ground Truth', 'Notebook', 'Training', etc. The main area displays 'Private Team' settings. It includes fields for 'Amazon Cognito 사용자 풀' (us-east-1\_BOqZg77p8) and '액 클라이언트' (1d3272645r7ej94hr18uaedque). Below this, a table shows the 'Private Team (1)' with one entry: 'private-team-1' (arn:aws:sagemaker:us-east-1:556463194331:workteam/private-crowd/private-team-1), created on 2020년 7월 29일 오전 8:57 UTC. Another table below shows 'Workers' with one entry: 'ggamini@hmail.hanyang.ac.kr' (Cognito 상태: Confirmed, 활성: 예). The bottom of the screen shows the URL https://console.aws.amazon.com/sagemaker/groundtruth?region=us-east-1# and standard AWS footer links.

- 위와 같이 authorized로 바뀝니다.

## 2. Create a labeling job

이제 labeling을 해보겠습니다.

The screenshot shows the 'Create Labeling Job' interface. The sidebar is identical to the previous screenshot. The main area has a red button labeled '레이블 지정 작업 생성' (Create Labeling Job). A message below the button states '현재 생성된 레이블 지정 작업이 없습니다.' (No labeling jobs have been created yet). The bottom of the screen shows the URL https://console.aws.amazon.com/sagemaker/labeling-jobs?region=us-east-1# and standard AWS footer links.

- 레이블 지정 작업 생성으로 들어갑니다.

작업 세부 정보 지정

작업 개요

작업 이름  
job-demo-1

I want to specify a label attribute name different from the labeling job name.  
Label attribute name is the key where your labels are stored in the augmented manifest. Ground Truth uses the labeling job name as the default label attribute name.

Automated data setup  
Provide the S3 location of the dataset you want labeled and let Ground Truth automatically connect to and use this dataset for your job.

Manual data setup  
Provide the S3 location of a file (an input manifest file) that identifies the data objects you want labeled

S3 location for input datasets 정보  
This is the location in S3 where your dataset objects are stored. Ground Truth will use all data objects in this location for your labeling job.  
s3://input-dataset-bucket/optional-prefix

S3 location for output datasets 정보  
This is the location in S3 where your labeling job output data is stored.  
 Same location as input dataset  
 Specify a new location

Data type

의견 한국어

2008 - 2020, Amazon Web Services, Inc. 또는 계열사. All rights reserved. 개인 정보 보호 정책 이용 약관

- s3에 있는 파일로 할 수도, manifest file을 만들 수도 있습니다.

버킷 만들기

① 이름 및 지역 ② 옵션 구성 ③ 권한 설정 ④ 검토

이름 및 지역

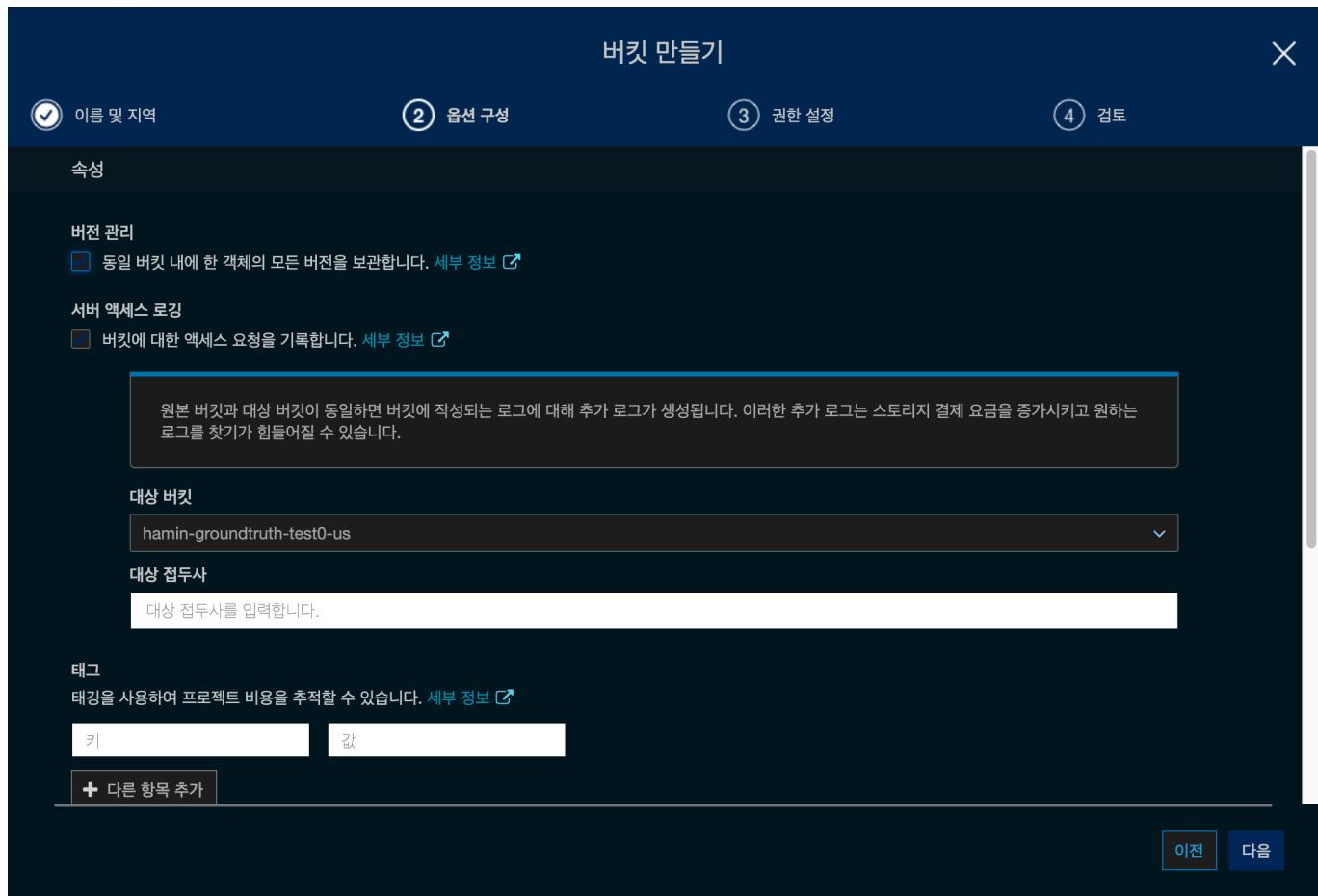
버킷 이름 ⓘ  
hamin-groundtruth-test0-us

리전  
미국 동부(버지니아 북부)

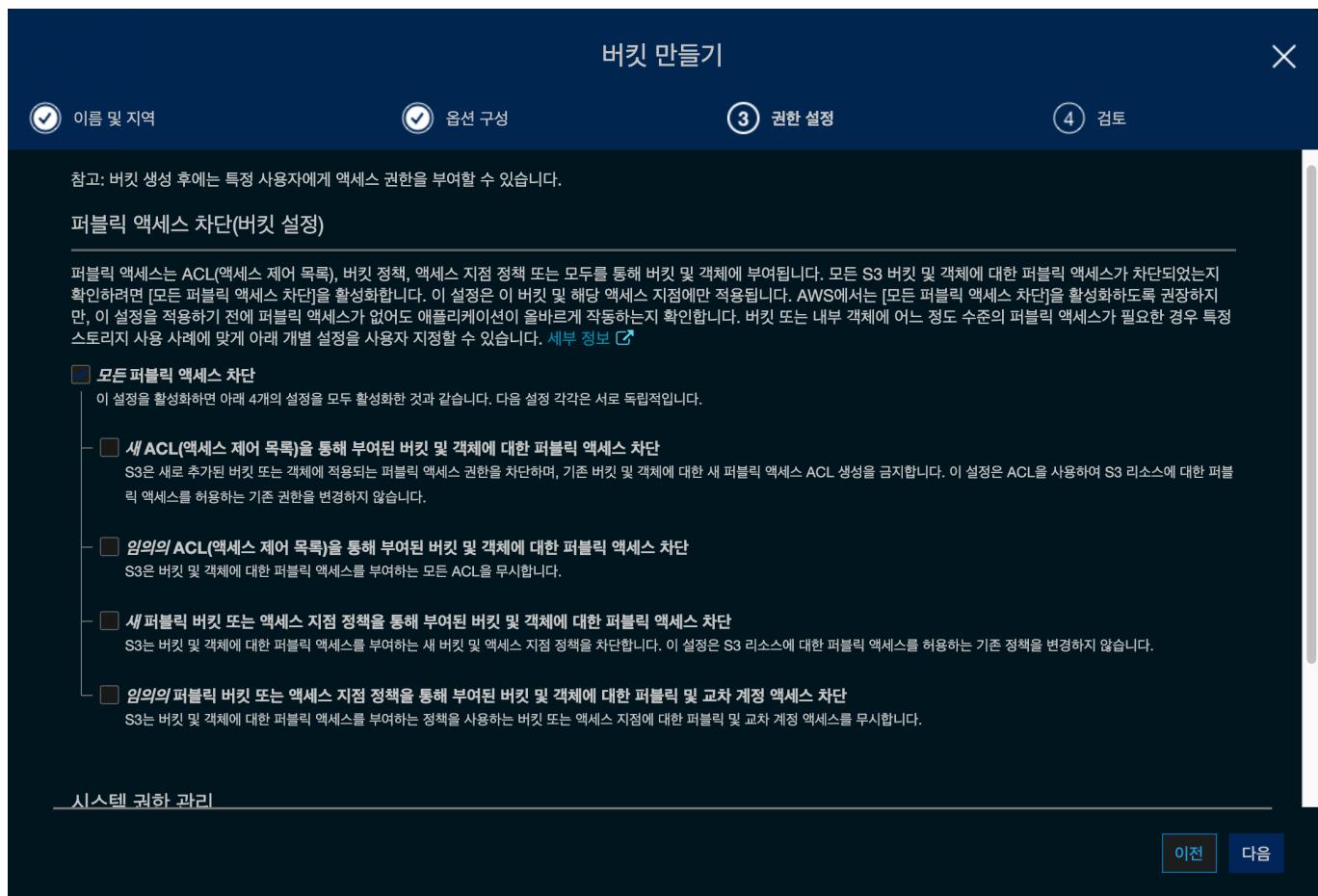
기존 버킷에서 설정 복사  
버킷을 선택합니다(선택 사항). 4버킷

생성 취소 다음

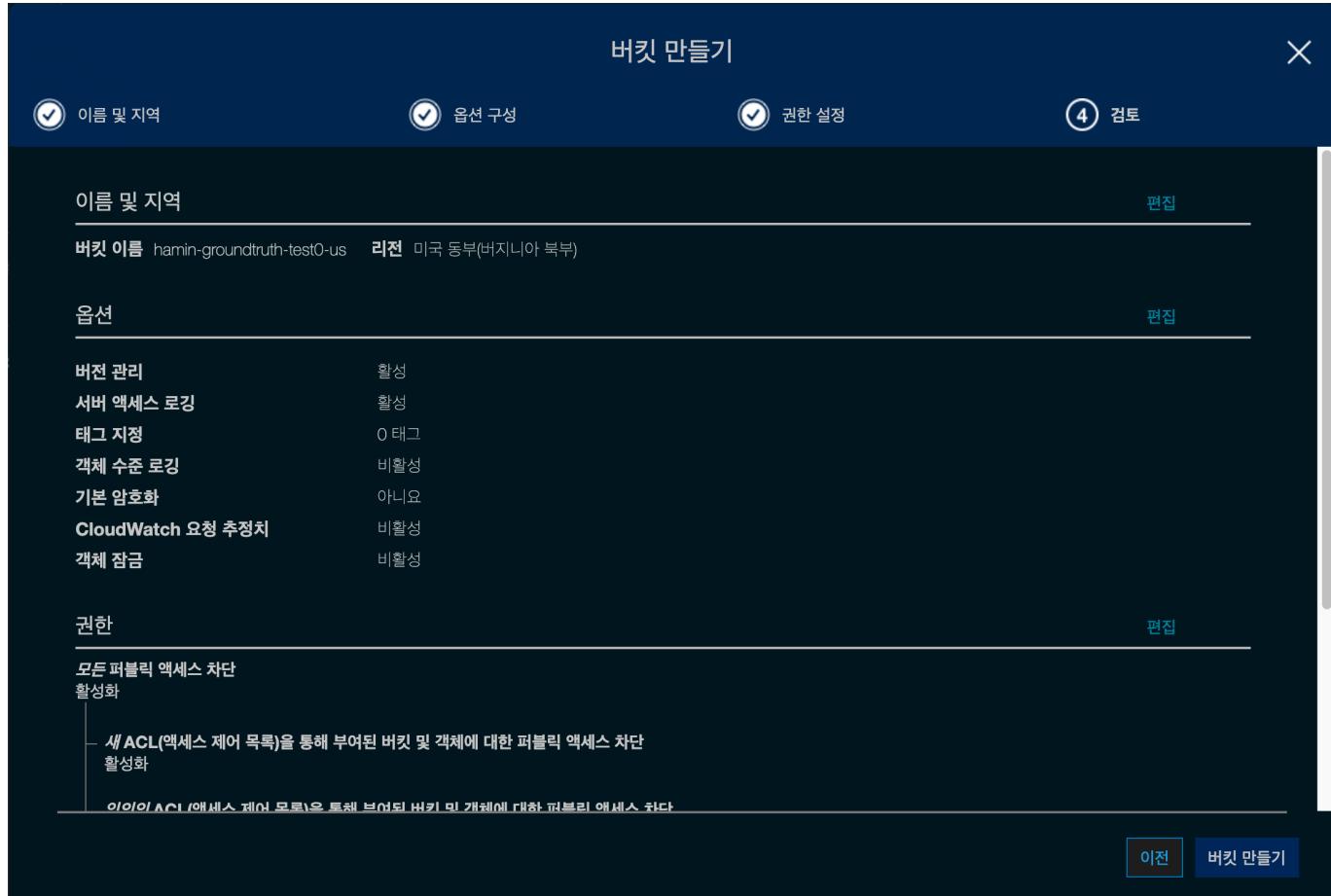
- S3는 다음과 같은 세팅을 하였습니다.



- 해당 버킷에 대한 액세스 요청을 모두 기록합니다.



- 권한 설정도 우선 기본 사항으로 하였습니다.



- 이렇게 S3 버킷을 생성했습니다.
- 그리고 샘플 dataset은 제 마음대로 넣어봤습니다.
- S3 버킷에 파일을 넣고 manifest를 생성해주어야 합니다.
- manifest 만들기는 다음을 참고합니다

[https://docs.aws.amazon.com/ko\\_kr/sagemaker/latest/dg/sms-data-input.html#sms-console-create-manifest-file](https://docs.aws.amazon.com/ko_kr/sagemaker/latest/dg/sms-data-input.html#sms-console-create-manifest-file)

<https://docs.aws.amazon.com/sagemaker/latest/dg/sms-getting-started-step1.html>

- manifest는 샘플들을 리스트하는 json 파일인데 docs와 다르게 manifest 없이도 할 수 있도록 되어있습니다.

그리고 다시 labeling으로 돌아가

Amazon SageMaker Studio

Ground Truth

Input data setup

S3 location for input datasets

S3 location for output datasets

Data type

- 이런식으로 세팅을 해 줍니다

Select the type of data being labeled to view available task templates for it or select 'Custom' to create your own.

Image

작업 선택

Image Classification (Single Label)

Image Classification (Multi Label)

Object Detection

Semantic Segmentation

- dataset이 이미지를 이므로 task category는 image, task selection은 semantic segmentation으로 해보겠습니다.

다음으로 넘어가서

The screenshot shows the Amazon SageMaker Studio interface. On the left sidebar, there's a navigation menu with sections like 'Amazon SageMaker Studio', 'Ground Truth', 'Notes', 'Checklist', 'Training', 'Deployment', and 'Models'. The main content area is titled '작업자와 구성 도구 선택' (Worker and Configuration Tools Selection). It includes a '작업자 유형' (Worker Type) section with three options: '퍼블릭' (Public), '프라이빗' (Private), and '공급업체 관리' (Managed by Vendor). The 'Private' option is selected. Below this, there's a '프라이빗 팀' (Private Team) section where 'private-team-1' is chosen. There are also sections for '테스크 제한 시간' (Task Duration Limit) set to 1 minute, '데이터 레이블 지정 자동화' (Automated Labeling), and '추가 구성 - 선택 사항' (Additional Configuration - Options). At the bottom right, there are links for '개인 정보 보호 정책' (Privacy Policy) and '이용 약관' (Terms of Service).

- private으로 세팅해 줍니다.

The screenshot shows the 'Ground Truth Labeling' interface. It features a 'Good example' section with a placeholder 'Add image here' and a 'Bad example' section with a placeholder 'Add image here'. To the right, there's a text input field containing the instruction 'please segment people and traffic light and bus, tree, car'. Below this, there's a preview image of a street scene with several people and vehicles. To the right of the image is a '레이블' (Labels) section with a list of items: 'people', 'traffic light', 'bus', 'car', and 'tree'. A '레이블 추가' (Add Label) button is also present. At the bottom, there are buttons for '취소' (Cancel), '이전' (Previous), and a large red '생성' (Create) button. The footer contains links for '개인 정보 보호 정책' (Privacy Policy) and '이용 약관' (Terms of Service).

- 이런식으로 예시가 나옵니다
- 구분해야할 물체를 최대한 구체적으로 명시해줍니다.

Amazon SageMaker Studio

Amazon SageMaker > 레이블 지정 작업

작업	레이블 지정 작업			
<input type="button" value="C"/> 작업	레이블 지정 작업 생성			
< 1 >				
이름	상태	작업 유형	레이블이 지정된 객체/전체	생성 시간
job-demo-1	진행 중	의미 체계 분할	-	2020년 7월 30일 오전 9:19 UTC

Search labeling jobs

© 2008 - 2020, Amazon Web Services, Inc. 또는 계열사. All rights reserved. 개인 정보 보호 정책 이용 약관

- 생성 버튼을 눌러주면 다음과 같이 labeling job이 생성 됩니다.

Amazon SageMaker Studio

Amazon SageMaker > 레이블 지정된 데이터 세트 객체 (-)

출력	태그
< 1 ... >	
pic1.jpg	pic2.jpg

© 2008 - 2020, Amazon Web Services, Inc. 또는 계열사. All rights reserved. 개인 정보 보호 정책 이용 약관

- 들어가보면 작업 중인 사진들이 보입니다.
- 시간이 좀 걸리니 기다려야 합니다.

The screenshot shows the Amazon SageMaker Studio interface. On the left, there is a sidebar with various navigation options like Amazon SageMaker Studio, Dashboard, Search, Ground Truth (Labeling Job, Data Set Labeling), Notebook Instances, Notebooks, Pending Tasks, Algorithms, Tuning Jobs, and Model Packages. The main content area is titled 'Amazon SageMaker > 레이블 지정 작업' (Labeling Job). It displays a table with one row of data:

작업 유형	레이블이 지정된 객체/전체	생성 시간
의미 체계 분할	0 / 8	2020년 7월 30일 오전 9:19 UTC

At the bottom right of the main area, there are buttons for '작업' (Job) and '레이블 지정 작업 생성' (Create Labeling Job). The footer contains links for ' 의견' (Feedback), ' 한국어' (Korean), and copyright information: '© 2008 - 2020, Amazon Web Services, Inc. 또는 계열사. All rights reserved.' and links for '개인 정보 보호 정책' (Privacy Policy) and '이용 약관' (Terms of Service).

- 굉장히 오래걸리는군요..4시간 정도 경과했습니다.