ahyarmuawwal / Automl-Translation Public

Notifications Fork 0 Star 0

Code Issues Pull requests Actions Projects Security Insights

master 1 Branch 0 Tags

Go to file

Code

ahyarmuawwal dataset untul cloud translation

.gitattributes	Initial commit
en-zh.tsv	dataset untul cloud
id-en.tsv	dataset untul cloud
id-en2.tsv	dataset untul cloud
id-en3.tsv	dataset untul cloud

Clone

Which remote URL should I use?

HTTPS GitHub CLI

https://github.com/ahyarmuawwal/Automl-Translat

Use Git or checkout with SVN using the web URL.

Open with GitHub Desktop

Download ZIP

About

Dataset untuk layanan google cloud translation

Activity
0 stars
2 watching
0 forks
Report repository

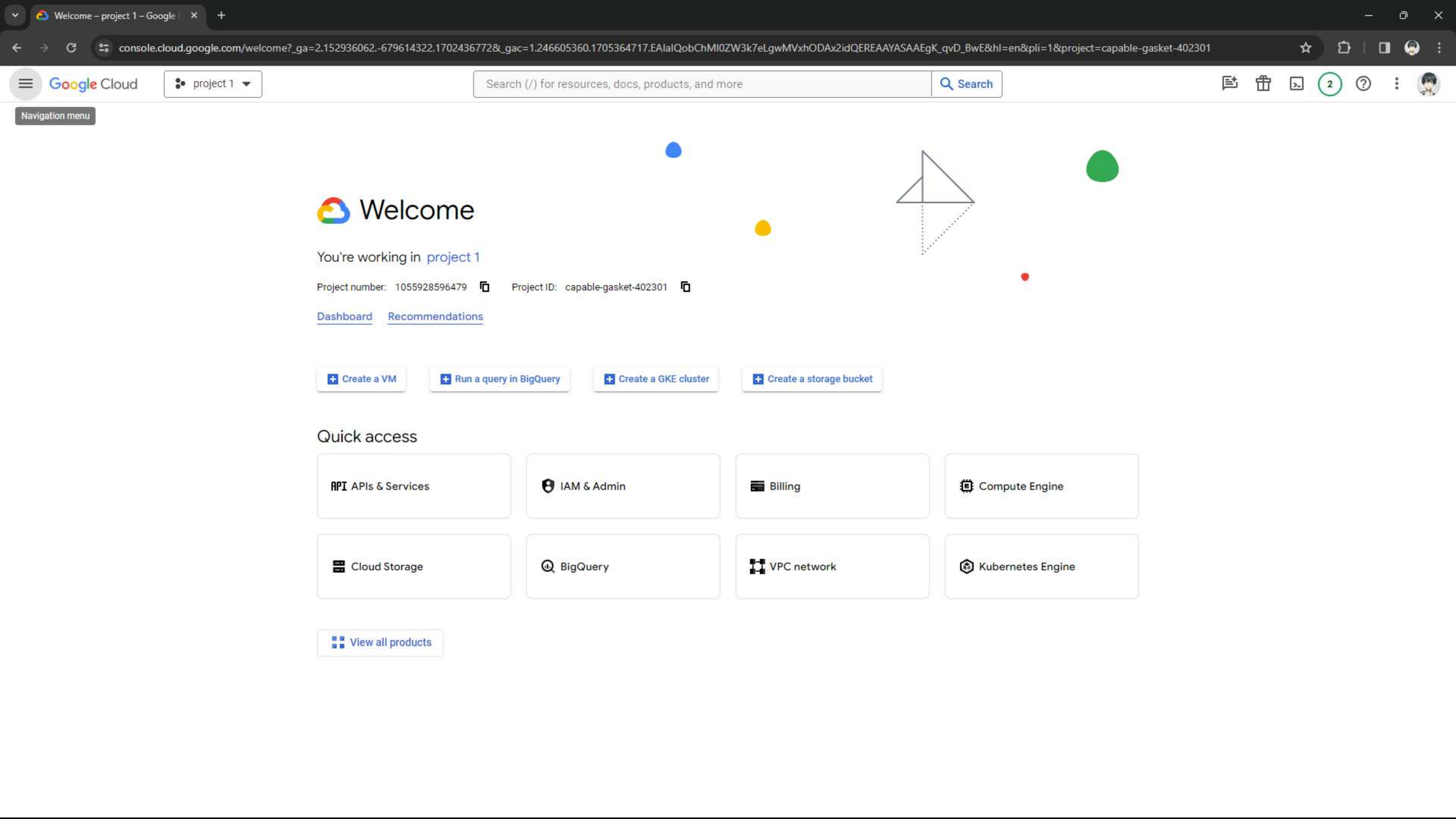
Releases

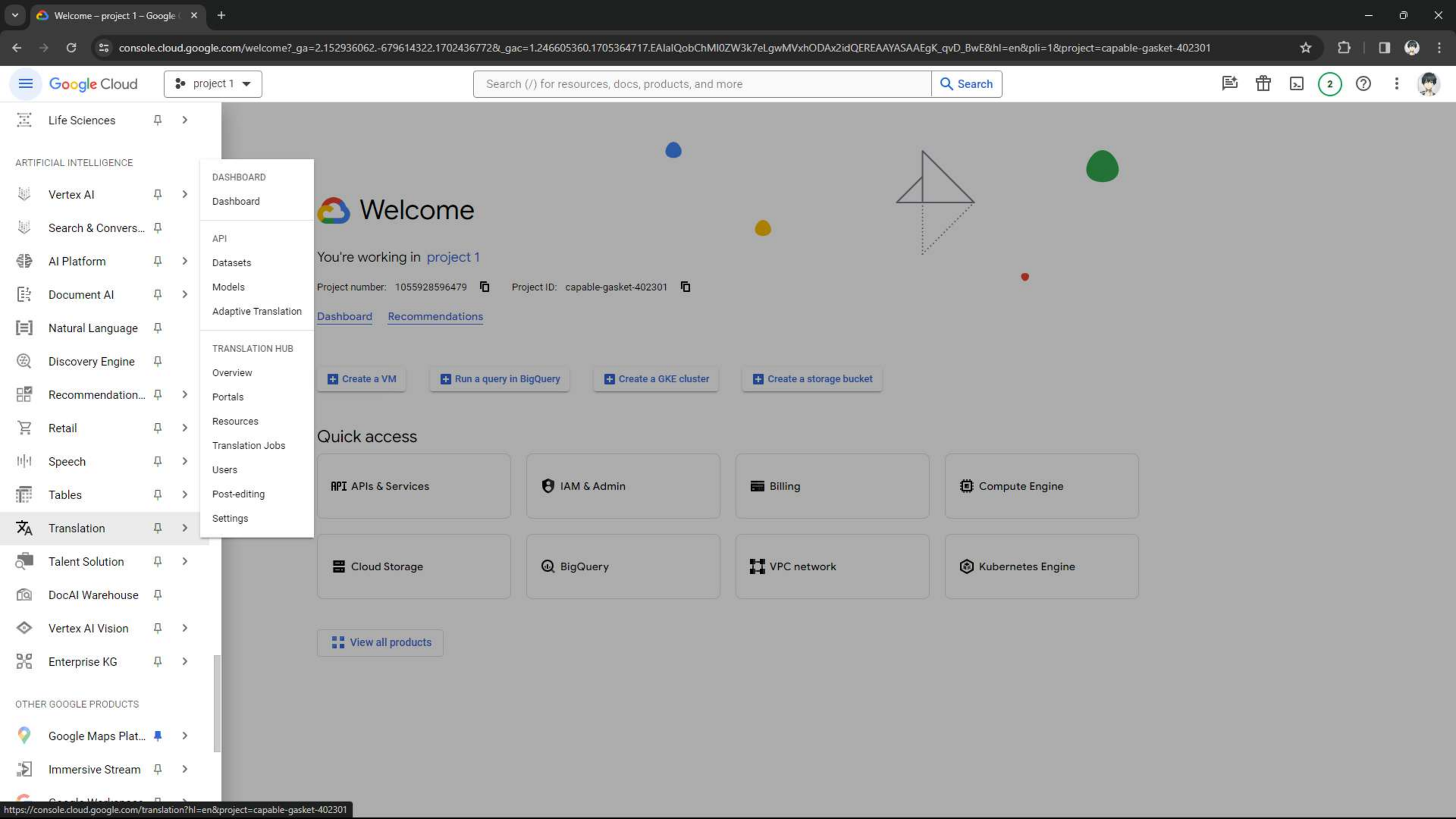
No releases published

Packages

No packages published

https://docs.github.com/articles/which-remote-url-should-i-use


Welcome – project 1 – Google
console.cloud.google.com/welcome?_ga=2.152936062.-679614322.1702436772&_gac=1.246605360.1705364717.EAIaIQobChMI0ZW3k7eLgwMVxhODAx2idQEREAAYASAAEgK_qvD_BwE&hl=en&pli=1&project=capable-gasket-402301
Google Cloud
project 1
Search (/) for resources, docs, products, and more
Search
Navigation menu
Welcome
You're working in project 1
Project number: 1055928596479
Project ID: capable-gasket-402301
Dashboard
Recommendations
Create a VM
Run a query in BigQuery
Create a GKE cluster
Create a storage bucket
Quick access
APIs & Services
IAM & Admin
Billing
Compute Engine
Cloud Storage
BigQuery
VPC network
Kubernetes Engine
View all products


Welcome – project 1 – Google
console.cloud.google.com/welcome?_ga=2.152936062.-679614322.1702436772&_gac=1.246605360.1705364717.EAIaIQobChMI0ZW3k7eLgwMVxhODAx2idQEREAAYASAAEgK_qvD_BwE&hl=en&pli=1&project=capable-gasket-402301
Google Cloud
project 1
Search (/) for resources, docs, products, and more
Search
Life Sciences
ARTIFICIAL INTELLIGENCE
Vertex AI
Search & Convers...
AI Platform
Document AI
Natural Language
Discovery Engine
Recommendation...
Retail
Speech
Tables
Translation
Talent Solution
DocAI Warehouse
Vertex AI Vision
Enterprise KG
OTHER GOOGLE PRODUCTS
Google Maps Plat...
Immersive Stream
DASHBOARD
Dashboard
API
Datasets
Models
Adaptive Translation
TRANSLATION HUB
Overview
Portals
Resources
Translation Jobs
Users
Post-editing
Settings
Welcome
You're working in project 1
Project number: 1055928596479
Project ID: capable-gasket-402301
Dashboard
Recommendations
Create a VM
Run a query in BigQuery
Create a GKE cluster
Create a storage bucket
Quick access
APIs & Services
IAM & Admin
Billing
Compute Engine
Cloud Storage
BigQuery
VPC network
Kubernetes Engine
View all products
https://console.cloud.google.com/translation?hl=en&project=capable-gasket-402301

Translation

Dashboard
Dashboard

API
Datasets
Models
Adaptive Translation

Translation Hub
Overview
Portals
Resources
Translation Jobs
Users
Post-editing
Settings

AutoML

Translation Models

Train a custom model using your own dataset of sentence pairs. No machine learning experience required. Learn more

To get started, enable the AutoML API

ENABLE API

Translation

Dashboard

Dashboard
Dashboard
API
Datasets
Models
Adaptive Translation
Translation Hub
Overview
Portals
Resources
Translation Jobs
Users
Post-editing
Settings

Translation products

AutoML Translation

Build on top of Google's powerful Translation API with the words, phrases, and idioms that matter most to you. No machine learning experience needed. Learn more

Get started

Cloud Translation API

Use Google's proven pre-trained model to get translations for 100+ languages. Learn more

Take Tutorial

Translation Hub

Enables translation admins to create and manage an end-user, self-serve document translation portal with document format preservation and high translation quality, and amend with human post-editing when necessary.

Get started

Datasets – Translation – project

console.cloud.google.com/translation/datasets?hl=en&project=capable-gasket-402301

Google Cloud | project 1 | Search (/) for resources, docs, products, and more | Search

Translation

Datasets | CREATE DATASET | VIEW RECENT ACTIVITIES

Dashboard
- Dashboard

API
- Datasets
- Models
- Adaptive Translation

Translation Hub
- Overview
- Portals
- Resources
- Translation Jobs
- Users
- Post-editing
- Settings

Name	Source	Target	Total pairs	Last updated	Status

No Translation datasets created yet.

Translation
Dashboard
Dashboard
API
Datasets
Models
Adaptive Translation
Translation Hub
Overview
Portals
Resources
Translation Jobs
Users
Post-editing
Settings

Datasets
CREATE DATASET
VIEW RECENT ACTIVITIES

Name	Source	Target	Total pairs	Last updated	Status

No Translation datasets created yet.

Create dataset

Dataset name *
angel_translate_dataset
Use letters, numbers and underscores up to 32 characters.

Translate from ... *
Indonesian (ID)

Translate to ... *
English (EN)

Create Legacy Dataset

CANCEL
CREATE

Datasets – Translation – project

console.cloud.google.com/translation/datasets?hl=en&project=capable-gasket-402301

Google Cloud | project 1 | Search (/) for resources, docs, products, and more | Search

Translation

Datasets | CREATE DATASET | VIEW RECENT ACTIVITIES

- Dashboard
  - Dashboard
- API
  - Datasets
  - Models
  - Adaptive Translation
- Translation Hub
  - Overview
  - Portals
  - Resources
  - Translation Jobs
  - Users
  - Post-editing
  - Settings

Name	Source	Target	Total pairs	Last updated	Status
angel_translate_dataset 0127c374457b093f	Indonesian (ID)	English (EN)	0	1/16/24, 8:08 AM	

angel_translate_dataset EXPORT DATA VIEW RECENT ACTIVITIES

IMPORT SENTENCES TRAIN EVALUATE PREDICT Translation (ID — EN)

Select files to import

To build a custom model, you first need to import a set of sentence pairs to train it.
Generally the more sentence pairs, the better. TSV and TMX files are currently supported.
You can add more files later. More Importing tips

- (•) Upload files from your computer
- ( ) Select files on Cloud Storage

☐ Use separate files for training, validation, and testing (advanced)

Upload files from your computer

Your files will be automatically split into training, test, and validation sets. If you have more than 100,000 sentence pairs, use the separate files option.

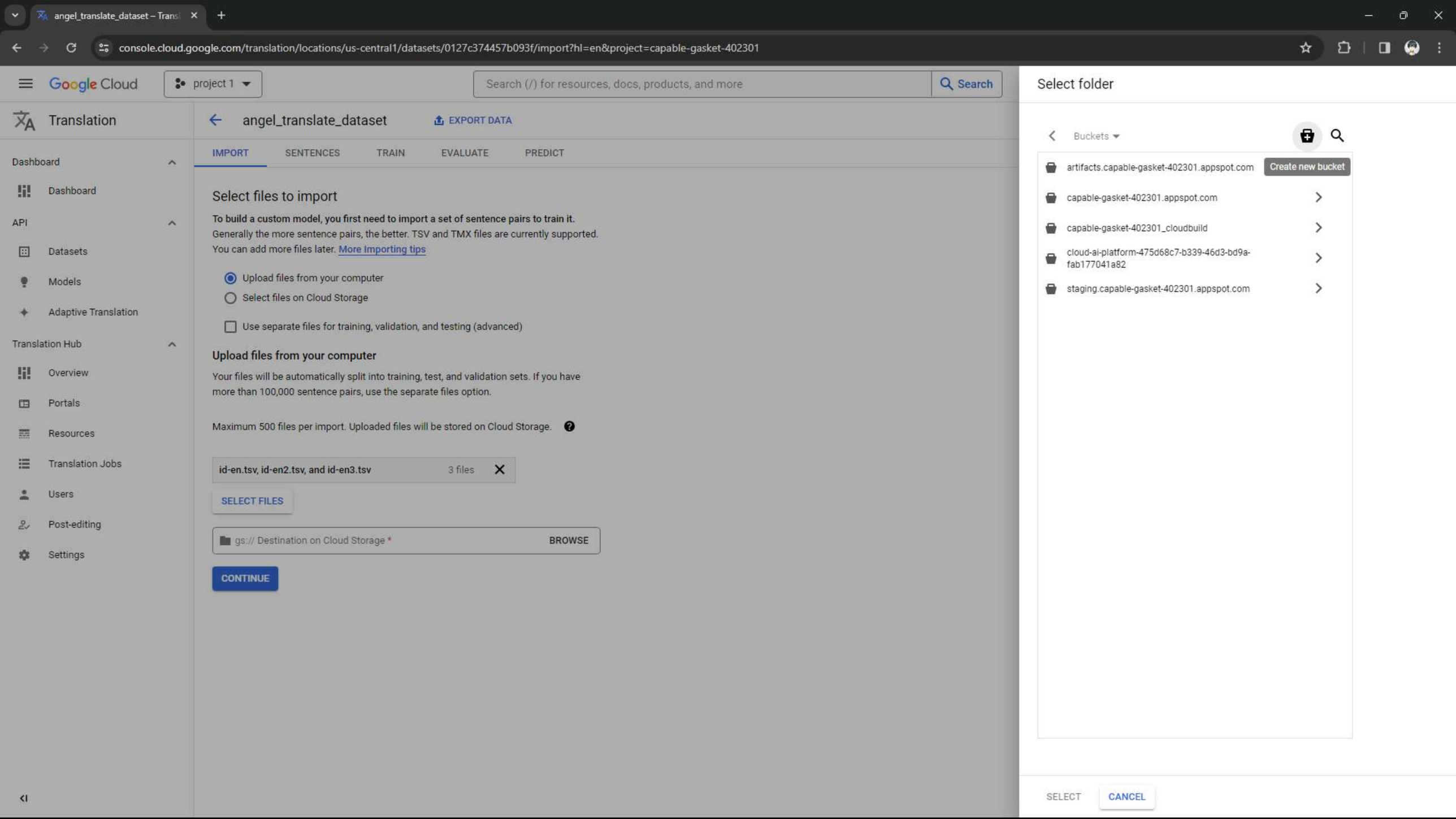
Maximum 500 files per import. Uploaded files will be stored on Cloud Storage.

id-en.tsv, id-en2.tsv, and id-en3.tsv 3 files ✕

SELECT FILES

gs:// Destination on Cloud Storage * BROWSE

CONTINUE


angel_translate_dataset – Transl
console.cloud.google.com/translation/locations/us-central1/datasets/0127c374457b093f/import?hl=en&project=capable-gasket-402301
Google Cloud
project 1
Search (/) for resources, docs, products, and more
Search
Translation
angel_translate_dataset
EXPORT DATA
Dashboard
Dashboard
API
Datasets
Models
Adaptive Translation
Translation Hub
Overview
Portals
Resources
Translation Jobs
Users
Post-editing
Settings
IMPORT
SENTENCES
TRAIN
EVALUATE
PREDICT
Select files to import
To build a custom model, you first need to import a set of sentence pairs to train it.
Generally the more sentence pairs, the better. TSV and TMX files are currently supported.
You can add more files later. More Importing tips
Upload files from your computer
Select files on Cloud Storage
Use separate files for training, validation, and testing (advanced)
Upload files from your computer
Your files will be automatically split into training, test, and validation sets. If you have more than 100,000 sentence pairs, use the separate files option.
Maximum 500 files per import. Uploaded files will be stored on Cloud Storage.
id-en.tsv, id-en2.tsv, and id-en3.tsv
3 files
SELECT FILES
gs:// Destination on Cloud Storage *
BROWSE
CONTINUE
Select folder
Buckets
Create new bucket
artifacts.capable-gasket-402301.appspot.com
capable-gasket-402301.appspot.com
capable-gasket-402301_cloudbuild
cloud-ai-platform-475d68c7-b339-46d3-bd9a-fab177041a82
staging.capable-gasket-402301.appspot.com
SELECT
CANCEL

Google Cloud | project 1 | Search (/) for resources, docs, products, and more | Search

Translation

angel_translate_dataset | EXPORT DATA

IMPORT | SENTENCES | TRAIN | EVALUATE | PREDICT

Dashboard
- Dashboard

API
- Datasets
- Models
- Adaptive Translation

Translation Hub
- Overview
- Portals
- Resources
- Translation Jobs
- Users
- Post-editing
- Settings

Select files to import

To build a custom model, you first need to import a set of sentence pairs to train it.
Generally the more sentence pairs, the better. TSV and TMX files are currently supported.
You can add more files later. More Importing tips

- (•) Upload files from your computer
- ( ) Select files on Cloud Storage

[ ] Use separate files for training, validation, and testing (advanced)

Upload files from your computer

Your files will be automatically split into training, test, and validation sets. If you have more than 100,000 sentence pairs, use the separate files option.

Maximum 500 files per import. Uploaded files will be stored on Cloud Storage.

id-en.tsv, id-en2.tsv, and id-en3.tsv — 3 files ✕

SELECT FILES

gs:// Destination on Cloud Storage * BROWSE

CONTINUE

Create a bucket

Name your bucket

Pick a **globally unique**, permanent name. Naming guidelines

translate_database

Tip: Don't include any sensitive information

LABELS (OPTIONAL)

CONTINUE

Choose where to store your data

This choice defines the geographic placement of your data and affects cost, performance, and availability. Cannot be changed later. Learn more

Location type

- ( ) Multi-region
  Highest availability across largest area
- ( ) Dual-region
  High availability and low latency across 2 regions
- (•) Region
  Lowest latency within a single region

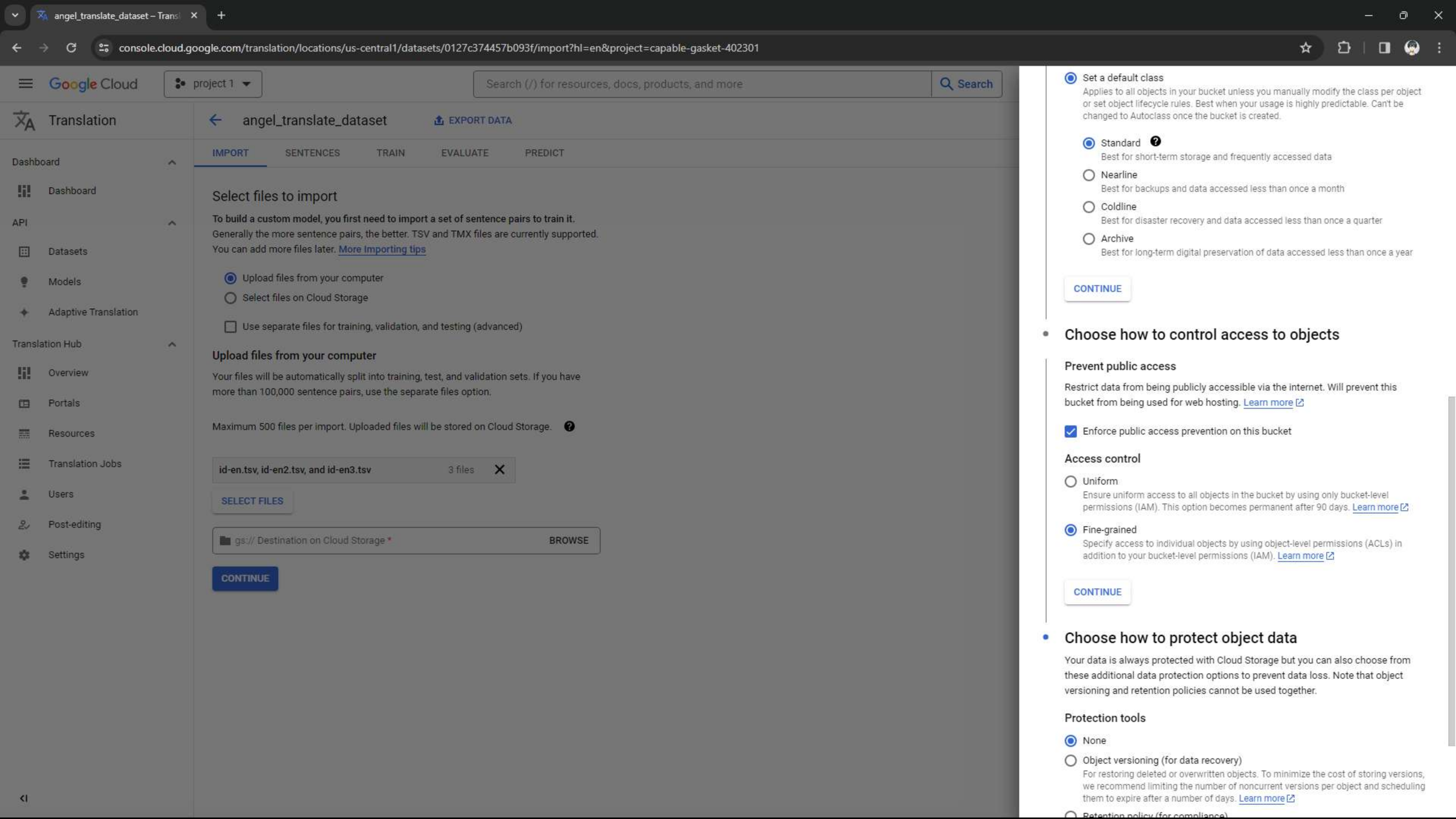
us-central1 (Iowa)

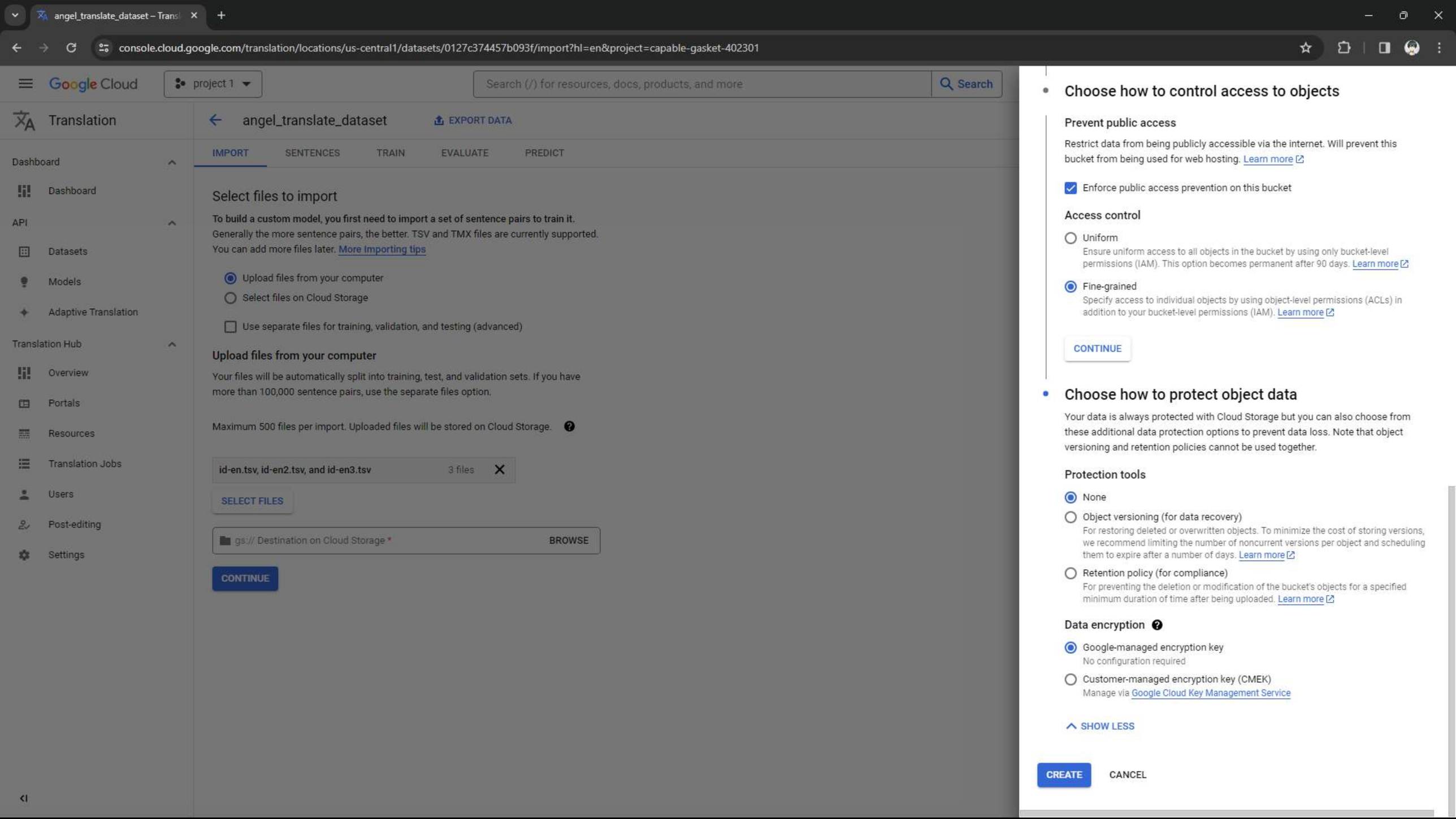
CONTINUE

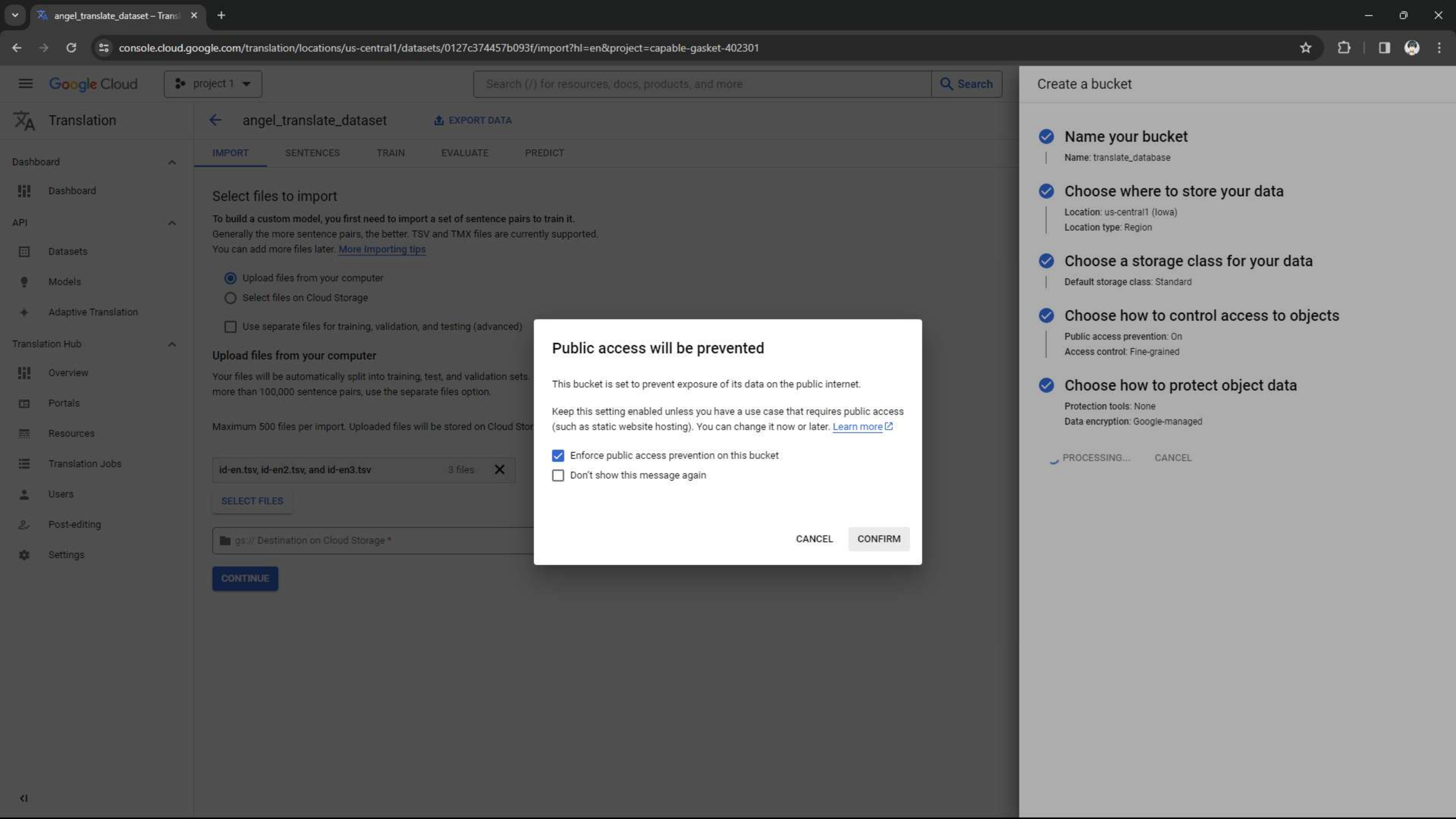
Choose a storage class for your data

A storage class sets costs for storage, retrieval, and operations, with minimal differences in uptime. Choose if you want objects to be managed automatically or specify a default storage class based on how long you plan to store your data and your workload or use case. Learn more

- ( ) Autoclass
  Automatically transitions each object to Standard or Nearline class based on object-level activity, to optimize for cost and latency. Recommended if usage frequency may be unpredictable. Can be changed to a default class at any time. Pricing details
- (•) Set a default class
  Applies to all objects in your bucket unless you manually modify the class per object or set object lifecycle rules. Best when your usage is highly predictable. Can't be changed to Autoclass once the bucket is created


angel_translate_dataset – Trans
console.cloud.google.com/translation/locations/us-central1/datasets/0127c374457b093f/import?hl=en&project=capable-gasket-402301
Google Cloud
project 1
Search (/) for resources, docs, products, and more
Search
Translation
angel_translate_dataset
EXPORT DATA
IMPORT
SENTENCES
TRAIN
EVALUATE
PREDICT
Dashboard
Dashboard
API
Datasets
Models
Adaptive Translation
Translation Hub
Overview
Portals
Resources
Translation Jobs
Users
Post-editing
Settings
Select files to import
To build a custom model, you first need to import a set of sentence pairs to train it.
Generally the more sentence pairs, the better. TSV and TMX files are currently supported.
You can add more files later. More Importing tips
Upload files from your computer
Select files on Cloud Storage
Use separate files for training, validation, and testing (advanced)
Upload files from your computer
Your files will be automatically split into training, test, and validation sets. If you have more than 100,000 sentence pairs, use the separate files option.
Maximum 500 files per import. Uploaded files will be stored on Cloud Storage.
id-en.tsv, id-en2.tsv, and id-en3.tsv
3 files
SELECT FILES
gs:// Destination on Cloud Storage *
BROWSE
CONTINUE
Set a default class
Applies to all objects in your bucket unless you manually modify the class per object or set object lifecycle rules. Best when your usage is highly predictable. Can't be changed to Autoclass once the bucket is created.
Standard
Best for short-term storage and frequently accessed data
Nearline
Best for backups and data accessed less than once a month
Coldline
Best for disaster recovery and data accessed less than once a quarter
Archive
Best for long-term digital preservation of data accessed less than once a year
CONTINUE
Choose how to control access to objects
Prevent public access
Restrict data from being publicly accessible via the internet. Will prevent this bucket from being used for web hosting. Learn more
Enforce public access prevention on this bucket
Access control
Uniform
Ensure uniform access to all objects in the bucket by using only bucket-level permissions (IAM). This option becomes permanent after 90 days. Learn more
Fine-grained
Specify access to individual objects by using object-level permissions (ACLs) in addition to your bucket-level permissions (IAM). Learn more
CONTINUE
Choose how to protect object data
Your data is always protected with Cloud Storage but you can also choose from these additional data protection options to prevent data loss. Note that object versioning and retention policies cannot be used together.
Protection tools
None
Object versioning (for data recovery)
For restoring deleted or overwritten objects. To minimize the cost of storing versions, we recommend limiting the number of noncurrent versions per object and scheduling them to expire after a number of days. Learn more


angel_translate_dataset – Trans
console.cloud.google.com/translation/locations/us-central1/datasets/0127c374457b093f/import?hl=en&project=capable-gasket-402301
Google Cloud
project 1
Search (/) for resources, docs, products, and more
Search
Translation
angel_translate_dataset
EXPORT DATA
IMPORT
SENTENCES
TRAIN
EVALUATE
PREDICT
Dashboard
Dashboard
API
Datasets
Models
Adaptive Translation
Translation Hub
Overview
Portals
Resources
Translation Jobs
Users
Post-editing
Settings
Select files to import
To build a custom model, you first need to import a set of sentence pairs to train it.
Generally the more sentence pairs, the better. TSV and TMX files are currently supported.
You can add more files later. More Importing tips
Upload files from your computer
Select files on Cloud Storage
Use separate files for training, validation, and testing (advanced)
Upload files from your computer
Your files will be automatically split into training, test, and validation sets. If you have more than 100,000 sentence pairs, use the separate files option.
Maximum 500 files per import. Uploaded files will be stored on Cloud Storage.
id-en.tsv, id-en2.tsv, and id-en3.tsv
3 files
SELECT FILES
gs:// Destination on Cloud Storage *
BROWSE
CONTINUE
Choose how to control access to objects
Prevent public access
Restrict data from being publicly accessible via the internet. Will prevent this bucket from being used for web hosting. Learn more
Enforce public access prevention on this bucket
Access control
Uniform
Ensure uniform access to all objects in the bucket by using only bucket-level permissions (IAM). This option becomes permanent after 90 days. Learn more
Fine-grained
Specify access to individual objects by using object-level permissions (ACLs) in addition to your bucket-level permissions (IAM). Learn more
CONTINUE
Choose how to protect object data
Your data is always protected with Cloud Storage but you can also choose from these additional data protection options to prevent data loss. Note that object versioning and retention policies cannot be used together.
Protection tools
None
Object versioning (for data recovery)
For restoring deleted or overwritten objects. To minimize the cost of storing versions, we recommend limiting the number of noncurrent versions per object and scheduling them to expire after a number of days. Learn more
Retention policy (for compliance)
For preventing the deletion or modification of the bucket's objects for a specified minimum duration of time after being uploaded. Learn more
Data encryption
Google-managed encryption key
No configuration required
Customer-managed encryption key (CMEK)
Manage via Google Cloud Key Management Service
SHOW LESS
CREATE
CANCEL


angel_translate_dataset – Trans
console.cloud.google.com/translation/locations/us-central1/datasets/0127c374457b093f/import?hl=en&project=capable-gasket-402301
Google Cloud
project 1
Search (/) for resources, docs, products, and more
Search
Translation
angel_translate_dataset
EXPORT DATA
IMPORT
SENTENCES
TRAIN
EVALUATE
PREDICT
Dashboard
Dashboard
API
Datasets
Models
Adaptive Translation
Translation Hub
Overview
Portals
Resources
Translation Jobs
Users
Post-editing
Settings
Select files to import
To build a custom model, you first need to import a set of sentence pairs to train it.
Generally the more sentence pairs, the better. TSV and TMX files are currently supported.
You can add more files later. More Importing tips
Upload files from your computer
Select files on Cloud Storage
Use separate files for training, validation, and testing (advanced)
Upload files from your computer
Your files will be automatically split into training, test, and validation sets
more than 100,000 sentence pairs, use the separate files option.
Maximum 500 files per import. Uploaded files will be stored on Cloud Stor
id-en.tsv, id-en2.tsv, and id-en3.tsv
3 files
SELECT FILES
gs:// Destination on Cloud Storage *
CONTINUE
Public access will be prevented
This bucket is set to prevent exposure of its data on the public internet.
Keep this setting enabled unless you have a use case that requires public access (such as static website hosting). You can change it now or later. Learn more
Enforce public access prevention on this bucket
Don't show this message again
CANCEL
CONFIRM
Create a bucket
Name your bucket
Name: translate_database
Choose where to store your data
Location: us-central1 (Iowa)
Location type: Region
Choose a storage class for your data
Default storage class: Standard
Choose how to control access to objects
Public access prevention: On
Access control: Fine-grained
Choose how to protect object data
Protection tools: None
Data encryption: Google-managed
PROCESSING...
CANCEL

angel_translate_dataset – Trans

console.cloud.google.com/translation/locations/us-central1/datasets/0127c374457b093f/import?hl=en&project=capable-gasket-402301

Google Cloud | project 1 | Search (/) for resources, docs, products, and more | Search

Translation

angel_translate_dataset | EXPORT DATA

IMPORT | SENTENCES | TRAIN | EVALUATE | PREDICT

Dashboard
- Dashboard

API
- Datasets
- Models
- Adaptive Translation

Translation Hub
- Overview
- Portals
- Resources
- Translation Jobs
- Users
- Post-editing
- Settings

Select files to import

To build a custom model, you first need to import a set of sentence pairs to train it.
Generally the more sentence pairs, the better. TSV and TMX files are currently supported.
You can add more files later. More Importing tips

- (•) Upload files from your computer
- ( ) Select files on Cloud Storage

[ ] Use separate files for training, validation, and testing (advanced)

Upload files from your computer

Your files will be automatically split into training, test, and validation sets. If you have more than 100,000 sentence pairs, use the separate files option.

Maximum 500 files per import. Uploaded files will be stored on Cloud Storage.

id-en.tsv, id-en2.tsv, and id-en3.tsv 3 files ✕

SELECT FILES

Destination on Cloud Storage *
gs:// translate_database BROWSE

CONTINUE

Created bucket translate_database ✕

Select folder

Buckets

- artifacts.capable-gasket-402301.appspot.com
- capable-gasket-402301.appspot.com
- capable-gasket-402301_cloudbuild
- cloud-ai-platform-475d68c7-b339-46d3-bd9a-fab177041a82
- staging.capable-gasket-402301.appspot.com
- translate_database

SELECT | CANCEL

Dashboard
Dashboard
API
Datasets
Models
Adaptive Translation
Translation Hub
Overview
Portals
Resources
Translation Jobs
Users
Post-editing
Settings

Select files to import

To build a custom model, you first need to import a set of sentence pairs to train it.
Generally the more sentence pairs, the better. TSV and TMX files are currently supported.
You can add more files later. More Importing tips

- (•) Upload files from your computer
- ( ) Select files on Cloud Storage

- [ ] Use separate files for training, validation, and testing (advanced)

Upload files from your computer

Your files will be automatically split into training, test, and validation sets. If you have more than 100,000 sentence pairs, use the separate files option.

Maximum 500 files per import. Uploaded files will be stored on Cloud Storage.

id-en.tsv, id-en2.tsv, and id-en3.tsv 3 files ×

SELECT FILES

Destination on Cloud Storage *
gs:// translate_database BROWSE

CONTINUE

Created bucket translate_database ×

Translation

Dashboard
Dashboard
API
Datasets
Models
Adaptive Translation
Translation Hub
Overview
Portals
Resources
Translation Jobs
Users
Post-editing
Settings

angel_translate_dataset EXPORT DATA VIEW RECENT ACTIVITIES

IMPORT SENTENCES TRAIN EVALUATE PREDICT Translation (ID → EN)

All sentences	2,109
Training	1,662
Validation	218
Testing	229

Filter Enter property name or value

Source	Target	Set
Listening for "Ok Google"	Esperando escuchar "Ok Google"	Validation
See and delete your answers to survey questions from Google.	Consulta y elimina tus respuestas a encuestas de Google.	Testing
Your activity, while you were signed in to Google.	Tu actividad con la sesión iniciada en tu cuenta de Google.	Training
Guillermo Haro's 105th birthday	105 aniversario del nacimiento de Guillermo Haro	Training
Send a message to this business.	Envía un mensaje a esta empresa.	Testing
You can change your primary list whenever you want.	Puedes modificar esta lista siempre que quieras.	Training
Anyone with the link can see this collection	Cualquier usuario que tenga el enlace puede ver esta colección.	Testing
Junko Tabei's 79th Birthday	79.º aniversario del nacimiento de Junko Tabei	Testing
Gboard is no longer in my keyboard list	Gboard ya no está en mi lista de teclados	Training
This site may be hacked.	Este sitio web puede haber sido pirateado.	Training
Add, organize, or share books	Añadir, organizar o compartir libros	Testing
Watch this video to learn more.	Para obtener más información, reproduce este vídeo .	Training
Added to your events	Se ha añadido a tus eventos	Training
May–Aug is mild and peak travel period.	Los meses de mayo a agosto son templados y coinciden con la temporada alta.	Training
Tap a description to suggest whether it applies to this place	Ayuda a describir el lugar tocando las características correspondientes	Training
Double-click to end this line	Haz doble clic para finalizar esta línea.	Validation
Popular travel periods include Dec–Jan, Apr and Jul–Aug.	La temporada alta es en diciembre, enero, abril, julio y agosto.	Training
Jun–Sep is mild and rainy.	De junio a septiembre, el tiempo es lluvioso y templado.	Training
Choose the category where the book is.	Elige la categoría donde está el libro.	Testing
You liked this album	Te ha gustado este álbum	Training
Deselect filter chips to see more products	Desmarca los filtros para ver más productos	Training
Email address contains invalid domain name	La dirección de correo electrónico contiene un nombre de dominio no válido.	Training
May–Sep is warm.	De mayo a septiembre, el tiempo es cálido.	
Clear filters for more results.	Para obtener más resultados, elimina los filtros.	
You liked this book	Te ha gustado este libro	
See or delete your search activity	Ver o eliminar la actividad de búsqueda	
What info shows up	Qué información se muestra	

Created bucket translate_database

Uploads and project 1 operations

id-en3-2024-01-16T01:17:38.935Z.tsv	Complete
id-en2-2024-01-16T01:17:38.935Z.tsv	Complete
id-en-2024-01-16T01:17:38.935Z.tsv	Complete

Train new model – Translation -

console.cloud.google.com/translation/locations/us-central1/datasets/0127c374457b093f/train?hl=en&project=capable-gasket-402301

Google Cloud | project 1 | Search (/) for resources, docs, products, and more

Translation

angel_translate_dataset | EXPORT DATA

IMPORT | SENTENCES | TRAIN | EVALUATE | PREDICT

START TRAINING

Dashboard
- Dashboard

API
- Datasets
- Models
- Adaptive Translation

Translation Hub
- Overview
- Portals
- Resources
- Translation Jobs
- Users
- Post-editing
- Settings

Created bucket translate_database

Train new model

Model name *
angel_translate_final

All custom translation models are trained based on the Google NMT model.

Base model
Google NMT

Data summary
1662 training pairs, 218 validation pairs, 229 testing pairs

See the Pricing guide for details about training time and cost.

START TRAINING | CANCEL

Translation

angel_translate_dataset EXPORT DATA VIEW RECENT ACTIVITIES

IMPORT SENTENCES TRAIN EVALUATE PREDICT Translation (ID → EN)

Dashboard
Dashboard
API
Datasets
Models
Adaptive Translation
Translation Hub
Overview
Portals
Resources
Translation Jobs
Users
Post-editing
Settings

START TRAINING

Training a model

This can take several hours or more.

Created bucket translate_database

Uploads and project 1 operations

id-en3-2024-01-16T01:17:38.935Z.tsv	Complete
id-en2-2024-01-16T01:17:38.935Z.tsv	Complete
id-en-2024-01-16T01:17:38.935Z.tsv	Complete