

Mô tả cuộc thi

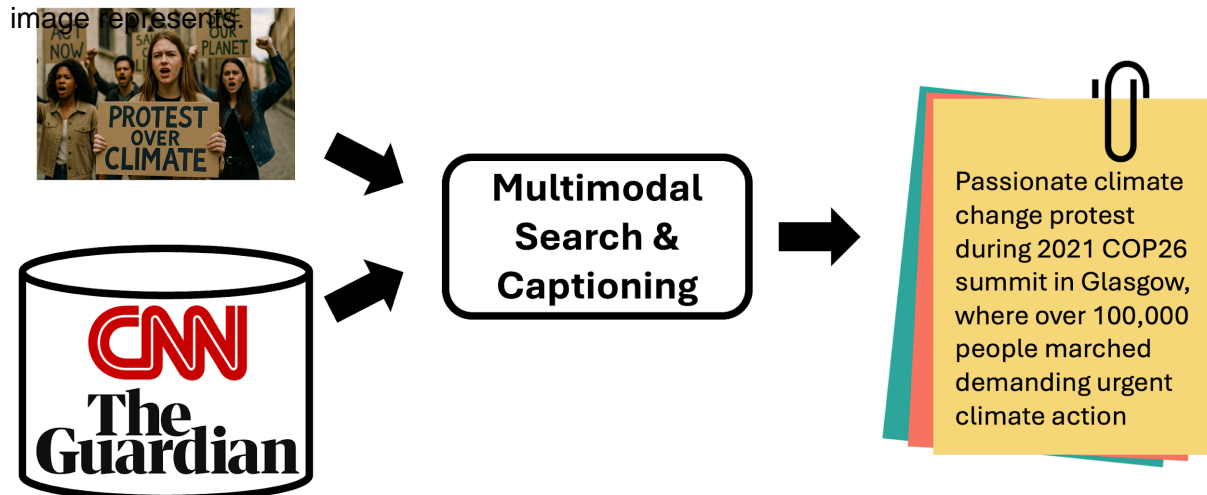
Mô tả cuộc thi
To participate in the EVENTA 2025 Grand Challenge, please first register by submitting [the form](#).

To participate in the EVENTA 2025 Grand Challenge, please first register by submitting

Overview

This track aims to generate captions that provide richer, more comprehensive information about an image. These captions go beyond simple visual descriptions, offering deeper insights, including the names and attributes of objects, the timing, context, outcomes of events, and other crucial details. These captions go beyond simple visual descriptions, offering deeper insights, including the names and attributes of objects, the timing, context, outcomes of events, and other crucial details. Participants are required to search relevant articles in a provided internal database and extract necessary information to generate an enriched image caption. This retrieval augmentation generation track facilitates the creation of more coherent and detailed narratives, capturing not only the visible elements but also the underlying context. This retrieval augmentation generation track facilitates the creation of more coherent and detailed image representations, capturing not only the visible elements but also the underlying context and significance of the scene, ultimately offering a more complete understanding of what the image represents.

Participants are required to search relevant articles in a provided internal database and extract necessary information to generate an enriched image caption. This retrieval augmentation generation track facilitates the creation of more coherent and detailed narratives, capturing not only the visible elements but also the underlying context. This retrieval augmentation generation track facilitates the creation of more coherent and detailed image representations, capturing not only the visible elements but also the underlying context and significance of the scene, ultimately offering a more complete understanding of what the image represents.



Submission Instructions

Participants must submit a CSV file named using the following format:

`TeamName_EVENTA2025_Track1.csv`. This file must be compressed into a ZIP archive named `submission.zip` before uploading to [CodaBench platform](#).

This file must be compressed into a ZIP archive

The CSV file should include predictions for all images in the query set. It must contain 12 columns, separated by commas (,), with the following structure:

The CSV file should include predictions for all images in the query set. It must contain 12 columns, separated by commas (,), with the following structure:

- Column 1: Query image ID
- Column 2-11: Top-10 retrieved article IDs, listed in descending order of relevance (from top-1 to top-10). If an article cannot be retrieved, use # as a placeholder.
- Column 12: Generated caption

- Column 12: The generated caption corresponding to the query image, enclosed in CSV Row Format Template: `"", "", ..., "", ""`.

We provide a submission "exampled_caption"

Submission Requirements

- Each image id in your submission must be unique and must come from the official ImageNet database.

- Each generated caption must not exceed 8,192 characters in length.

- The submission.csv must include all headers and columns, in the exact order as

• For published submissions, each team is allowed a maximum of 10 submissions per day and a total of 150 submissions.

- For private test submissions, each team is allowed a maximum of 10 submissions

Participants also require to submit a detailed paper through [this link](#) to validate their solutions.

Evaluation

To fairly and comprehensively evaluate submissions in our challenge, which includes both **Retrieval** and **Captioning** tasks, we define an **Overall Score** that balances key evaluation criteria across both domains and tasks, we define an **Overall Score** that balances key evaluation criteria across both domains.

The **Overall Score** integrates the following six metrics:

The Overall Score integrates the following six metrics:		
Metric	Description	Domain
Metric	Description	Domain
AP	Average Precision – measures retrieval precision across thresholds	Retrieval
AP	Average Precision · measures retrieval precision across thresholds	
Recall@1	Whether the correct item is ranked at the top	Retrieval
Recall1	Whether the correct item is ranked at the top	
Recall@10	Whether the correct item is within the top 10 results	Retrieval
Recall10	Whether the correct item is within the top 10 results	
CLIPScore	Semantic alignment between image and caption	Captioning
CLIPScore	Semantic alignment between image and caption	
CIDEr	Agreement between generated and reference captions	Captioning

CIDEr Agreement between generated and reference captions

These metrics are combined using a **weighted harmonic mean**, which emphasizes balanced performance and penalizes poor performance more heavily than a simple average.

The formula is:

$$\text{Overall Score} = r \cdot \left(\frac{\sum w_i (\sum (w_i m_i))}{\epsilon} \right)$$

Where:

Overall Scorer (wi(wimi.))

Where:

m_i is the value of the i -th metric
 w_i is the weight for the i -th metric
 ϵ is a small constant (e.g., 10^{-8}) to prevent division by zero
 N is the number of the #total ground-truth queries
 n is the number of the #total ground-truth queries
 Default Weights: [1, 2, 2, 3, 2] for AP, Recall@1, Recall@10, CLIPScore, CIDEr
 r (valid queries) (total ground-truth queries)
 Default Weights: 1, 2, 2, 3, 2 for AP, Recall1, Recall10, CLIPScore, CIDEr
 This **Overall Score** will be used as the **official ranking metric** for evaluating all participant submissions on the leaderboard.
 This will be used as the for evaluating all participant submissions on the leaderboard.

Leaderboard

TBA

Leaderboard

Tui có dataset như sau:

Tui có dataset nh· sau:

Cấu trúc Dataset EVENTA

Tổng quan

C·u trúc Dataset EVENTA

EVENTA-DATASET/

T·ng quan

EVENTA-DATASET

Data database json

data_base.json/ ch·a id.jpg

image_public_set/ ch·a id.jpg

public_image/ ch·a id.jpg

image_query.csv ch·a id.jpg

query_train_set/

train_image/ ch·a id.jpg

train_query.csv ch·a id.jpg

train.csv

Chi tiết

Database

Chi tiết

Database: Chứa thông tin về các bài viết (articles) và cấu trúc:

- database.json: Chứa thông tin về các bài viết (articles) và cấu trúc:
 - id
 - title
 - content
 - date
 - image_id

- image/: Thư mục chứa ảnh minh họa được sử dụng trong các bài viết
 - image_id

Public_Set: Thư mục chứa các hình ảnh ...c s- d- ng trong các bài viết

Public_Set: Thư mục chứa các ảnh minh họa cần caption

- queries.csv: File chứa thông tin về image_index
 - image : Th- m- c ch- a các hình -nh c- n d- -oán caption

Train_Set: File ch- a thông tin v- image_index

image/: Thư mục chứa các ảnh minh họa hu- n- luyện

- train.csv: File chứa 4 cột:
 - image_id: Th- m- c ch- a các hình -nh hu- n- luyện
 - caption: File ch- a 4 c- t:
 - image_index
 - image_id
 - caption
 - article_id
 - retrieved_image_id
 - retrieved_article_id