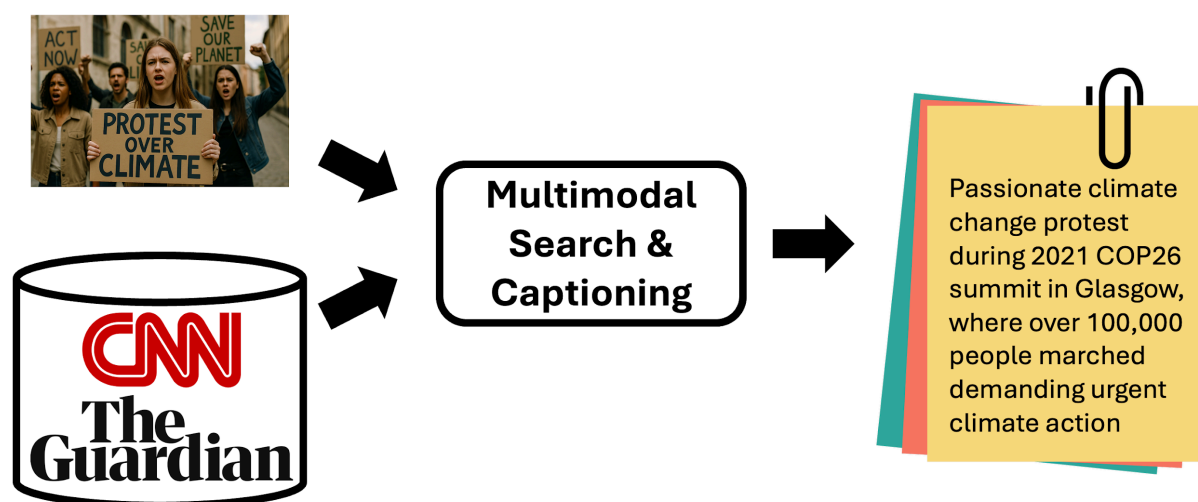


## Mô tả cuộc thi

Để tham gia vào Thử thách lớn Eventa 2025, vui lòng đăng ký trước bằng cách [gửi bài mẫu](#).

## Tổng quan

Ca khúc này nhằm tạo ra các chú thích cung cấp thông tin phong phú hơn, toàn diện hơn về một hình ảnh. Các chú thích này vượt xa các mô tả trực quan đơn giản bằng cách đưa ra những hiểu biết sâu sắc hơn, bao gồm các tên và thuộc tính của các đối tượng, thời gian, bối cảnh, kết quả của các sự kiện và các chi tiết quan trọng khác thông tin không thể lược lặt được từ việc tạo ra một cơ sở để tìm kiếm một cơ sở. Tạo ra các câu chuyện mạch lạc và chi tiết hơn, nắm bắt không chỉ các yếu tố có thể nhìn thấy mà còn cả bối cảnh và ý nghĩa cơ bản của cảnh, cuối cùng mang đến sự hiểu biết đầy đủ hơn về những gì hình ảnh thể hiện.



## Hướng dẫn đệ trình

Người tham gia phải gửi tệp CSV có tên bằng định dạng sau: TEAMNAME\_EVE NTA2025\_TRACK1.CSV. Tệp này phải được nén vào kho lưu trữ ZIP có tên là Submission.zip trước khi tải lên nền tảng [Codabench](#).

Tệp CSV nên bao gồm các dự đoán cho tất cả các hình ảnh trong bộ truy vấn. Nó phải chứa 12 cột, được phân tách bằng dấu phẩy (,), với cấu trúc sau:

Cột 1: ID hình ảnh truy vấn    Cột 2 đến 11: ID bài viết Top-10 được truy xuất, được liệt kê theo thứ tự giảm dần liên quan (từ Top-1 đến Top-10). Nếu một bài viết không thể được truy xuất, hãy sử dụng # làm người giữ chỗ.

Cột 12: Chú thích được tạo tương ứng với hình ảnh truy vấn, được đặt trong các dấu ngoặc kép ("").

Mẫu định dạng hàng CSV:

query\_id, bài báo\_id\_1, bài báo\_id\_2, ..., bài báo\_id\_10, đã tạo\_caption

<query\_id>, < article\_id\_1>, < article\_id\_2>, ..., < article\_id\_10>, "<

<query\_id>, < article\_id\_1>,#, ...,#, "< created\_caption >"

Chúng tôi cung cấp một ví dụ gửi:

query\_id, bài báo\_id\_1, bài báo\_id\_2, ..., bài báo\_id\_10, đã tạo\_caption

12312,56712,56723.56734.56745,56756.56767.56778.56789,56790,56701, "Một nhóm trẻ chơi bóng đá vào một buổi chiều đầy nắng."

12334,56712,##,##,##,##,##,##, "Một người đàn ông đi xe đạp qua một con đường thành phố bận rộn."

12345,56712,56723.56734.56745,56756,##,##,##,##, "Một con mèo ngồi trên bệ cửa sổ nhìn ra ngoài mưa."

Yêu cầu đề trình:

Mỗi query\_id trong bài nộp của bạn phải là duy nhất và phải đến từ danh sách chính thức được cung cấp. Tất cả các giá trị Image\_id phải tương ứng với các mục hợp lệ trong cơ sở dữ liệu hình ảnh của chúng tôi. Không có yêu cầu sắp xếp các hàng theo query\_id, điều này sẽ được xử lý tự động trong quá trình đánh giá. Mỗi lần tạo\_caption không được vượt quá 8.000 ký tự. Tập gửi RAW của bạn.csv không được chứa bất kỳ dòng trống nào.

Gửi tập Submission.csv dưới dạng một tập ZIP có tên là SubMed.zip - Không bao gồm bất kỳ thư mục chứa nào. Việc gửi.csv phải bao gồm tất cả các tiêu đề và cột, theo thứ tự chính xác như trong tập ví dụ.

Giới hạn đề trình:

Đối với bài kiểm tra công khai, mỗi nhóm được phép tối đa 10 bài nộp mỗi ngày và tổng cộng 150 bài nộp. Đối với các bài nộp thử nghiệm riêng, mỗi nhóm được phép tối đa 10 bài nộp mỗi ngày và tổng cộng 150 bài nộp.

Những người tham gia cũng yêu cầu gửi một bài báo [chi tiết thông qua nền tảng CMT](#) để xác nhận các giải pháp của họ.

# Sự đánh giá

Để đánh giá công bằng và toàn diện các bài nộp trong thách thức của chúng tôi, bao gồm cả các nhiệm vụ truy xuất và chú thích, chúng tôi xác định điểm tổng thể cân bằng các tiêu chí đánh giá chính trên cả hai miền.

Điểm tổng thể tích hợp sáu số liệu sau:

Metric	Description	Domain
AP	Average Precision – measures retrieval precision across thresholds	Retrieval
Recall@1	Whether the correct item is ranked at the top	Retrieval
Recall@10	Whether the correct item is within the top 10 results	Retrieval
CLIPScore	Semantic alignment between image and caption	Captioning
CIDEr	Agreement between generated and reference captions	Captioning

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:

$$\text{Điểm tổng thể} = r \cdot (\sum w_i \sum (w_i m_i + \epsilon))$$

Ở đâu:

$M_i$  là giá trị của số liệu thứ  $i$   
 $w_i$  là trọng lượng cho số liệu thứ  $i$   
 $\epsilon$  là một hằng số nhỏ (ví dụ: 10<sup>-8</sup>) để ngăn chặn sự phân chia theo 0  
 $R = (\text{truy vấn \#valid}) / (\text{truy vấn-biến số mặt đất \#total})$   
Trọng số mặc định: [1, 2, 2, 3, 2] cho [AP, nhớ lại@1, nhớ lại@10, clipscore, rượu táo]

Điểm tổng thể này sẽ được sử dụng làm số liệu xếp hạng chính thức để đánh giá tất cả các bài nộp của người tham gia trên bảng xếp hạng.

## Bảng xếp hạng

TBA

Tôi có dataset như sau:

## CẤU TRÚC DataSet Eventa

### Tổng Quan

Deta-DataSet/

Cơ sở dữ liệu/ cơ sở dữ liệu C  
ơ sở dữ liệu.

# Chi tiết

## Cơ sở dữ liệu

A

Id id	Tiêu đề
	Nội dung
	Ngày Im
	age_id

Hình ảnh/

## Public\_set

Hình ảnh/

## Train\_set

Hình ảnh/