

Thông tin về hình thức và cách nộp bài vòng thi chung kết

Nội dung thi vòng chung kết:

Yêu cầu truy vấn vòng chung kết được thể hiện theo một trong các dạng cụ thể sau:

- **Yêu cầu truy vấn dạng văn bản (Textual KIS):** Ban giám khảo cung cấp mô tả bằng ngôn ngữ tự nhiên của một sự kiện. Các đội dự thi cần tìm ra chính xác đoạn video của sự kiện này. Đoạn mô tả có thể gồm nhiều ý, nhiều câu văn. Ở vòng chung kết: **nội dung đoạn mô tả sẽ được cung cấp dần dần** trong thời gian dành cho câu truy vấn. Nếu đội dự thi tự tin vào kết quả tìm kiếm video từ những gợi ý ban đầu của đoạn mô tả, đội dự thi có thể nộp kết quả ngay để có thể được điểm rất cao cho câu truy vấn này (nếu kết quả đúng). Tuy nhiên, đội dự thi có thể thận trọng để chờ thêm các thông tin mô tả về sự kiện để kiểm chứng kết quả tìm được. Các đội có thể tham khảo cách tính điểm ở vòng chung kết để có chiến thuật nộp bài hợp lý.
- **Yêu cầu truy vấn dạng hình ảnh hay video (Video KIS):** thay vì mô tả bằng ngôn ngữ tự nhiên, các đội dự thi sẽ xem một đoạn video ngắn (không quá 20 giây) được trích ra từ một sự kiện nào đó trong kho dữ liệu video đã cho. Các đội **không được phép chụp ảnh, ghi hình đoạn video này** bằng bất kỳ phương tiện điện tử nào để đưa vào công cụ của mình mà phải tìm cách phù hợp để diễn tả yêu cầu tìm kiếm, ví dụ như mô tả bằng ngôn ngữ tự nhiên nội dung video hay vẽ lại bối cảnh mình nhìn thấy...
- **Yêu cầu truy vấn dạng Q&A:** Đây là nhiệm vụ tìm kiếm sự kiện và trích xuất thông tin cụ thể từ video. Ban giám khảo cung cấp một mô tả bằng ngôn ngữ tự nhiên của một sự kiện và một câu hỏi về thông tin trong sự kiện này. Các đội dự thi cần tìm ra chính xác đoạn video của sự kiện và trả lời câu hỏi. Ở **vòng chung kết**: Câu hỏi sẽ được cung cấp ngay từ đầu. Nội dung đoạn mô tả sẽ được **cung cấp lần lượt** theo thời gian dành cho câu truy vấn. Nếu đội dự thi tự tin vào kết quả tìm kiếm video từ những mô tả ban đầu, đội dự thi có thể nộp kết quả sớm để có thể được điểm cao cho câu truy vấn này (nếu kết quả đúng).
- **Yêu cầu truy vấn dạng Truy xuất và Căn chỉnh sự kiện video theo thời gian (TRAKE):** Đây là một nhiệm vụ phức hợp đòi hỏi độ chính xác cao trong cả việc truy xuất video và căn chỉnh thời gian của các khoảnh khắc quan trọng. Nhiệm vụ yêu cầu hệ thống không chỉ tìm kiếm một video phù hợp từ một kho dữ liệu lớn mà còn phải xác định chính xác các **khoảnh khắc ngữ nghĩa (semantic keyframe)** của một chuỗi sự kiện có cấu trúc bên trong video đó.
 - **Giai đoạn 1: Truy xuất (Retrieval):** Từ một thư viện video lớn, tìm ra một video duy nhất chứa chuỗi sự kiện khớp nhất với truy vấn.
 - **Giai đoạn 2: Căn chỉnh (Alignment):** Đối với video đã truy xuất, xác định chính xác một khung hình (**semantic keyframe**) duy nhất cho mỗi giai đoạn của chuỗi

sự kiện.

- Ở **vòng chung kết**: Mô tả về các sự kiện sẽ được cung cấp một lần từ đầu.
- *Lưu ý: "Keyframe ngữ nghĩa" trong truy vấn này là khoảnh khắc mang ý nghĩa về nội dung, khác với "I-Frame" là keyframe kỹ thuật trong các thuật toán nén video.*

Cách tính điểm ở vòng chung kết:

- Mỗi câu truy vấn được tối đa 100 điểm, và có thời gian xử lý tối đa là:
 - 4 phút đối với câu truy vấn dạng **Video KIS**
 - 5 phút đối với câu truy vấn dạng **Textual KIS, Q&A** và **TRAKE**.
- Tổng điểm của một đội trong vòng chung kết là tổng điểm của đội trong tất cả câu truy vấn.
- Điểm cuối cùng của một câu truy vấn được tính dựa trên thời điểm của **lần nộp bài đúng đầu tiên**.

Công thức tính điểm

Gọi:

- P_{max} : Điểm tối đa cho một truy vấn (100 điểm).
- P_{base} : Điểm nền khi hết giờ (50 điểm).
- $P_{penalty}$: Điểm phạt cho mỗi lần nộp sai (10 điểm).
- T_{task} : Tổng thời gian cho phép của truy vấn.
- t_{submit} : Thời gian từ lúc bắt đầu đến khi nộp bài.
- k : Số lần nộp sai trước lần nộp đúng đầu tiên.

Hàm điểm thời gian $f_T(t)$ được định nghĩa là:

$$f_T(t) = 1 - \frac{t_{submit}}{T_{task}}$$

Điểm cuối cùng ($Score$) được tính như sau:

1. Trường hợp có câu trả lời đúng hoàn toàn (áp dụng cho tất cả các task):

$$Score_{full} = \max(0, P_{base} + (P_{max} - P_{base}) \times f_T(t) - k \times P_{penalty})$$

2. Trường hợp có câu trả lời đúng một phần (chỉ áp dụng cho **TRAKE**):

Nếu không có câu trả lời đúng hoàn toàn, điểm sẽ được tính tại lần nộp đúng một phần đầu tiên.

$$Score_{partial} = \frac{\max(0, P_{base} + (P_{max} - P_{base}) \times f_T(t) - k \times P_{penalty})}{2}$$

Diễn giải

- **Đối với KIS và Q&A:** Các đội sẽ nhận được điểm nếu có một câu trả lời đúng hoàn toàn. Công thức (1) sẽ được áp dụng.
- **Đối với TRAKE:**

- Nếu đội thi đạt **đúng hoàn toàn** (đúng 100% keyframe), công thức (1) được áp dụng.
- Nếu đội thi không đạt đúng hoàn toàn nhưng có một lần đạt **đúng một phần** (đúng từ 50% đến dưới 100% keyframe), công thức (2) sẽ được áp dụng. Điểm số tính như trường hợp đúng hoàn toàn nhưng sau đó chia cho 2.
- Mọi lần nộp bài trước khi có kết quả đúng (hoàn toàn hoặc một phần) đầu tiên đều bị tính là sai và bị trừ điểm.

Hệ thống nộp bài và chấm điểm trực tuyến


Trong vòng chung kết AI Challenge 2025, các đội sẽ nộp bài theo đúng thể thức chuẩn quốc tế trong cuộc thi VBS (Video Browser Showdown, <https://videobrowsershowdown.org/>)

Hệ thống server để nộp bài và chấm điểm trực tuyến được sử dụng là DRES (Distributed Retrieval Evaluation Server):

- Nguồn: <https://github.com/dres-dev/DRES>
- Rossetto L., Gasser R., Sauter L., Bernstein A., Schuldt H. (2021) A System for Interactive Multimedia Retrieval Evaluations. In: Lokoč J. et al. (eds) MultiMedia Modeling. MMM 2021. Lecture Notes in Computer Science, vol 12573. Springer, Cham.
- Lưu ý: thí sinh không cần tự deploy hệ thống DRES, chỉ cần đảm bảo mình có thể nộp bài được vào hệ thống DRES theo đúng chuẩn giao tiếp quy định.

Để hỗ trợ các đội làm quen với hệ thống, BTC sẽ mở server thử nghiệm và thi thử cho các đội tại địa chỉ <https://eventretrieval.oj.io.vn> từ **12g ngày 1/11/2025** đến **23h59 ngày 7/11/2025**. Thông tin tài khoản sẽ được gửi cho các đội thi vào ngày **1/11**.

Hướng dẫn xem thông tin truy vấn trên DRES

Đội thi đăng nhập vào web và bấm vào biểu tượng  để xem thông tin truy vấn hiện tại.

Hướng dẫn submit kết quả vào hệ thống chấm điểm tự động DRES

Đội thi xem hướng dẫn cơ bản dưới đây để nộp bài. Ngoài ra, có thể tham khảo thêm tại <https://github.com/dres-dev/Client-Examples>

Các đội thi chưa hiểu rõ cách nộp bài có thể tham gia buổi tập huấn của BTC để hiểu rõ hơn về cách thức nộp.

Bước 1:

- Các đội cần đăng nhập vào hệ thống DRES để lấy `sessionId` của đội trong cuộc thi.
- `sessionId` có thể được xem như 1 **token định danh** của đội trong cuộc thi. Có 2 cách các đội có thể thực hiện để lấy `sessionId`.

- **Cách 1:**

1. Vào đường dẫn <https://eventretrieval.oj.io.vn/login> (hoặc đường dẫn mà BTC cung cấp trong buổi thi chính thức), điền thông tin username và password của đội được cung cấp từ BTC.
2. Sau khi đăng nhập thành công, các đội vào trang <https://eventretrieval.oj.io.vn/user> để lấy thông tin sessionId.

- **Cách 2:** Thực hiện POST request theo API:

Client Request:

- **URL:** <https://eventretrieval.oj.io.vn/api/v2/login>
- **Method:** POST
- **Body:**

```
{
  "username": "<username>",
  "password": "<password>"
}
```

Server Response:

```
{
  "id": "<id>",
  "username": "<username>",
  "role": "PARTICIPANT",
  "sessionId": "<sessionID>"
}
```

Bước 2:

- Thực hiện GET request để lấy evaluationID:

Client Request:

- **URL:** <https://eventretrieval.oj.io.vn/api/v2/client/evaluation/list>
- **Method:** GET
- **Param:**

```
{
  "session": "<sessionID>"
}
```

Server Response:

```
[
  {
    "id": "<evaluationID>",
    "name": "<evaluation_name>",
    "type": "SYNCHRONOUS",
    "status": "ACTIVE"
  }
]
```

Bước 3:

- Thực hiện POST request dùng **token định danh** theo API sau để submit đáp án:
- Kết quả của đoạn video được nộp dưới định dạng thời gian xuất hiện của frame được tìm thấy trong video gốc, **đơn vị là milisecond (ms)**.
- **Lưu ý:**
 - Item được tính là tên file video **không** có phần đuôi định dạng.
 - Bạn không được nộp trùng kết quả cho cùng một truy vấn.

Client Request:

- **URL:** <https://eventretrieval.oj.io.vn/api/v2/submit/{evaluationID}>
- **Method:** POST
- **Params:**

```
{
  "session": "<sessionID>"
}
```

Body (KIS):

```
{
  "answerSets": [{
    "answers": [{
      "mediaItemName": "<VIDEO_ID>",
      "start": "<TIME(ms)>",
      "end": "<TIME(ms)>"
    }]
  }]
}
```

Body (QA):

```
{
  "answerSets": [{
    "answers": [{
      "text": "QA-<ANSWER>-<VIDEO_ID>-<TIME(ms)>"
    }]
  }]
}
```

Body (TRAKE):

```
{
  "answerSets": [{
    "answers": [{
      "text": "TR-<VIDEO_ID>-<FRAME_ID1>,<FRAME_ID2>,... "
    }]
  }]
}
```