

- Bài toán: truy xuất, phân tích, giải thích các sự kiện tin tức từ tập dataset video lớn.
 - Có 2 hình thức của bài toán:
 - + Chế độ tương tác: người dùng nhập vào một queries và hệ thống sẽ tinh chỉnh và xuất ra thông tin (dạng như VBS: <https://videobrowsershowdown.org/> và LSC: <http://lsc.dcu.ie/>).
 - + Chế độ tự động: AI sẽ tự động truy xuất và giải thích sự kiện trong dataset.
 - Lưu ý về 2 chế độ:
 - + Chế độ tương tác: thời gian thực tương tác giữa con người - AI (maybe con người phải có tốc độ nhập thông tin nhanh từ BTC)
 - + Chế độ tự động: kiểm tra tính tinh chỉnh và giải thích tự động.
 - > Cả 2 đều check độ chính xác, hiệu quả
 - Nguồn dataset: tập dữ liệu video lớn tin tức từ các nguồn BBC, CNN, Reuters và Việt Nam. Gồm:
 - + Transcript lời nói
 - + Văn bản trích xuất OCR
 - + Mốc thời gian
 - > Có vẻ giống nhưng đơn giản hơn cuộc thi VBS 2025
 - Yêu cầu: sử dụng model AI để phân tách từng đoạn, truy xuất và giải thích các sự kiện. (ko giới hạn model).
 - Các tiêu chí chấm:
 - + Độ chính xác
 - + Thời gian
 - + Độ trễ của hệ thống giao tiếp người - AI
 - + Maybe người ta thêm phần bộ nhớ
 - Kết hợp giữa NLP và CV: NLP để truy xuất phần transcript lời nói sau đó gen ra phần explain và CV để truy xuất phân đoạn video.
- > Tạo một link liên kết giữa đoạn video với transcript lời nói và lưu trữ dưới dạng embedding vector để có thể truy xuất nhanh chóng. Có thể liên kết giữa các thông tin trong đoạn video vì nó là 1 video về tin tức nên có thể có cùng chủ đề --> Có thể dùng kỹ thuật GraphRAG.
- > Truy xuất phân đoạn video: dạng câu hỏi text-to-video (dùng trong giao tiếp giữa người với AI). Khi nhập vào 1 text xuất ra phân đoạn video cần tìm(chỗ này BTC chưa nêu rõ queries rõ cuộc là gì?? có thể là nhập text xuất video hoặc có thể là nhập video thứ xxx rồi bắt giải thích nội dung video đó hoặc có thể nhập image rồi bắt xuất video chứa image đó)
- Một số paper có thể liên quan:
 - + GraphRAG: <https://arxiv.org/pdf/2501.00309>
 - + Text-to-Video: <https://arxiv.org/abs/2401.03177>, <https://arxiv.org/abs/2204.02874>
 - + Image-toVideo: <https://web3.arxiv.org/abs/1604.07939>
 - + Explanation Video: <https://arxiv.org/html/2312.17432>, <https://arxiv.org/abs/2406.19528>