

Giới thiệu môn học

- Tổng số tiết: 30 tiết lý thuyết, 30 tiết thực hành
- Tóm tắt môn học: Học phần bao gồm các nội dung chính: tổng quan về dữ liệu đa phương tiện và hệ cơ sở dữ liệu đa phương tiện, các kỹ thuật truy xuất thông tin đối với dữ liệu văn bản, hình ảnh, âm thanh, video; kiến trúc cơ sở dữ liệu đa phương tiện.
- Đánh giá học phần:
 - Giữa kỳ (40%): chuyên cần + bài tập + bài thực hành + kiểm tra giữa kỳ
 - Cuối kỳ (60%): trắc nghiệm

Chuẩn đầu ra học phần

- Kiến thức
 - Trình bày các khái niệm về đa phương tiện, hệ cơ sở dữ liệu đa phương tiện, hệ quản trị cơ sở dữ liệu đa phương tiện, các thuật toán truy xuất thông tin, thao tác dữ liệu
 - Phân biệt được các thao tác, phương pháp lập chỉ mục và kỹ thuật truy xuất cho các loại đa phương tiện
 - Giải thích được kiến trúc cơ sở dữ liệu đa phương tiện

Chuẩn đầu ra học phần

- Kỹ năng

- Áp dụng được các kỹ thuật nén cơ bản đối với các dữ liệu âm thanh, hình ảnh và video
- Thể hiện được các thuật toán truy xuất thông tin chính đối với dữ liệu văn bản và hình ảnh dưới dạng một đoạn chương trình/ứng dụng cụ thể
- Ứng dụng các phương pháp truy xuất thông tin vào một số tập dữ liệu đa phương tiện

- Mức tự chủ và trách nhiệm

- Báo cáo, bảo vệ ý kiến; phân tích giải quyết vấn đề hợp lý và hiệu quả; có kỷ luật và phối hợp làm việc

Nội dung học phần

- Tổng quan về cơ sở dữ liệu đa phương tiện
- Thiết kế cơ sở dữ liệu đa phương tiện
- Dữ liệu văn bản
- Dữ liệu hình ảnh
- Dữ liệu âm thanh và video

Tài liệu tham khảo

- Phạm Thị Ngọc Diễm (cb), Lê Đức Thắng (2015), Giáo trình Hệ cơ sở dữ liệu đa phương tiện, NXB Đại học Cần Thơ, Cần Thơ.
- V. S. Subrahmanian, Principles of Multimedia Database Systems, Morgan Kaufmann Publishers, 2013
- P. Muneesawang, N. Zhang & L. Guan, Multimedia Database Retrieval Technology and Applications, Springer 2014
- B. Thuraisingham, Managing and Mining Multimedia Databases, CRC 2001

TỔNG QUAN

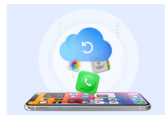
Hệ cơ sở dữ liệu đa phương tiện

HK1, 2023 - 2024

Giới thiệu chung



Dữ liệu lớn



Khả năng tính toán/
lưu trữ lớn

Nhu cầu chia
sẻ/ khai thác



Khái niệm

- **Truyền thông**

- truyền dữ liệu theo những quy tắc và cách thức nhất định
- hoạt động truyền thông tin từ một đối tượng này đến một đối tượng khác

- ***Hoạt động truyền thông tin***: sự tác động lẫn nhau qua một *trung gian* (phương tiện), giữa ít nhất hai tác nhân cùng chia sẻ một nội dung (content) mang một ý nghĩa nào đó

- truyền dữ liệu theo những quy tắc và cách thức nhất định
- hoạt động truyền thông tin từ một đối tượng này đến một đối tượng khác

Khái niệm

• Thông tin (Information)

- Tổng quát: *kiến thức (knowledge)* truyền đi hoặc nhận được liên quan đến một sự kiện (fact) hoặc một tình huống (circumstance) đặc biệt nào đó
- Kỹ thuật: *dãy các ký hiệu* thường được gọi là thông điệp (message)
- Những hiểu biết của con người về một thực thể nào đó, có thể thu thập, lưu trữ, xử lý được

• Dữ liệu (Data)

- là những đại lượng (quantities), ký tự (characters) hoặc những ký hiệu (symbols) mà máy tính thực hiện được các phép toán
- biểu diễn thông tin

Khái niệm

- **Phương tiện truyền thông (Media)**

- các công cụ được sử dụng để lưu trữ và phân phối thông tin hoặc dữ liệu
- thiết bị lưu trữ dữ liệu (storage) và các kênh truyền tín hiệu (transmission channel)
- digital media, published media, social media, multimedia, hypermedia

Khái niệm

Kiểu media

- Các kiểu thông tin và biểu diễn thông tin: văn bản (text), hình ảnh (image), đồ họa (graphic), video... và tất cả các thông tin có thể biểu diễn, lưu trữ, truyền, xử lý dưới dạng số

It was the best of times, it was the worst of times, it was the age of wisdom, it was the age of foolishness...

Text



Still image



Audio



Animation



Video footage



Interactivity

Khái niệm

Kiểu media: Phân loại

- Theo định dạng vật lý
- Theo mối liên hệ với thời gian
 - **Media tĩnh (static media)**: nội dung và ý nghĩa không phụ thuộc vào biểu diễn thời gian; ví dụ: văn bản, ảnh tĩnh, đồ họa



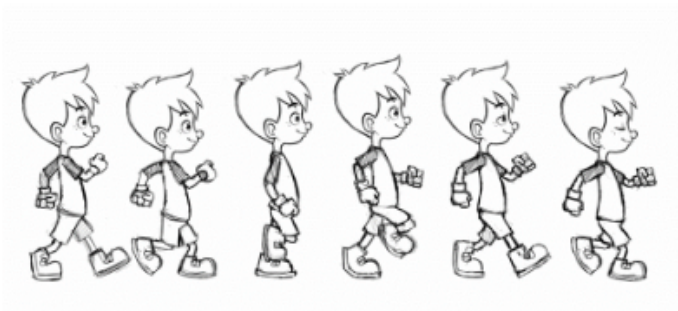
This is a lot of 12 point text to test the ocr code and see if it works on all types of file format.
The quick brown dog jumped over the lazy fox. The quick brown dog jumped over the lazy fox. The quick brown dog jumped over the lazy fox. The quick brown dog jumped over the lazy fox. The quick brown dog jumped over the lazy fox.



Khái niệm

Kiểu media: Phân loại

- Theo mối liên hệ với thời gian
 - **Media động (dynamic media, continuous media):** các dữ liệu có chiều thời gian; ví dụ: hoạt hình, âm thanh video

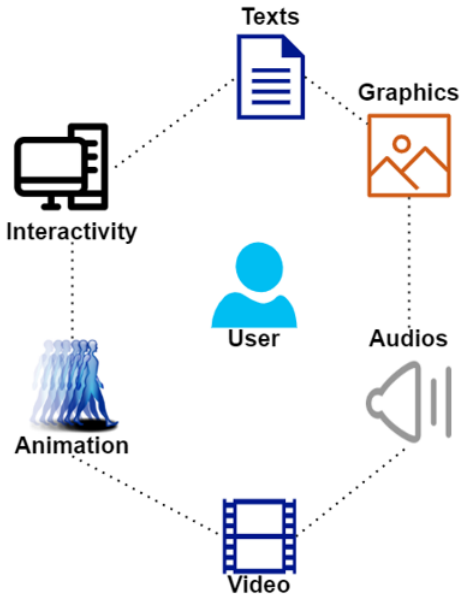


Khái niệm

Đa phương tiện (multimedia): Phân loại

- *tập hợp các kiểu media được sử dụng cùng nhau và trong đó có ít nhất một kiểu dữ liệu không phải là dữ liệu dạng văn bản*
- *có thể coi là tổ hợp của văn bản, âm thanh, ảnh tĩnh, ảnh động, video và các hình thức tương tác nội dung*

Các thành phần của đa phương tiện



Khái niệm

Dữ liệu đa phương tiện (multimedia data):

- biểu diễn của các kiểu dữ liệu media khác nhau mà máy tính có thể đọc được

CSDL đa phương tiện (multimedia database):

- là tập có cấu trúc nhất định các dữ liệu đa phương tiện

Hệ đa phương tiện (multimedia system):

- là hệ thống có khả năng xử lý dữ liệu đa phương tiện và các ứng dụng

Multimedia data

Dữ liệu dạng văn bản (text):

- chứa thông tin chủ đạo
- Input: bàn phím, các chương trình nhận dạng âm thanh và ký tự, dữ liệu lưu trên đĩa, phụ đề phim...



Multimedia data

Dữ liệu dạng văn bản (text):

- Định dạng: đa dạng
 - Text thường (file ASCII) hoặc text đã được định dạng (màu sắc, độ bóng, ...) (html, xml, RTF, Word, mã nguồn của chương trình C#, Java...)
- Kích thước lưu trữ: không đáng kể so với các dữ liệu đa phương tiện khác

Multimedia data

Dữ liệu dạng đồ họa (graphic):

- Gồm các cấu trúc đặc biệt được xây dựng bởi các đối tượng cơ bản (primitive)
 - Đường cong, đường thẳng, đa giác, đường tròn... để tạo ra các đối tượng 2D, 3D
- Dễ sửa đổi
- Input: các trình soạn thảo đồ họa (Adobe Illustrator, CorelDraw, Autocad...) hoặc bởi các chương trình khác (Postscript)
- Chuẩn đồ họa: OpenGL, GKS
- Lưu trữ: file lưu trữ tập các đối tượng cơ bản (primitive), kích thước không quá lớn

Multimedia data

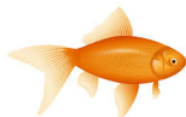
Dữ liệu ảnh (image):

- Thông tin, định dạng, mức độ chi tiết đa dạng
- Loại ảnh: tự nhiên, nhân tạo, từ các thiết bị đặc biệt
- Ảnh số là một chuỗi các điểm ảnh để biểu diễn 1 vùng sẽ được hiển thị trên màn hình của người sử dụng
- Input: camera, scanner, các chương trình mô phỏng, các phần mềm tạo và xử lý ảnh
- Định dạng: jpg, png, bmp, tiff. . .

Multimedia data

Dữ liệu ảnh (image):

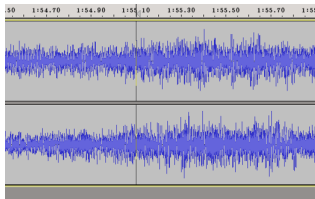
- Kích thước lưu trữ: phụ thuộc vào kích thước ảnh, độ phân giải, kỹ thuật nén (nếu có)
 - 1 bit / 1 pixel (ảnh nhị phân)
 - 8 bits / 1 pixel (ảnh đa mức xám)
 - 24 bits / 1 pixel (ảnh màu)
- Ảnh thường được nén để giảm không gian lưu trữ



Multimedia data

Dữ liệu âm thanh (audio):

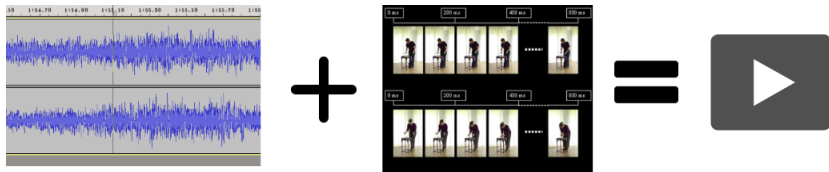
- Tiếng động, tiếng nói (văn bản đi kèm), nhạc, phim, các chương trình dịch tự động từ văn bản ...
- Tín hiệu âm thanh là tín hiệu tương tự và liên tục
- Input: microphone → số hóa và lưu trữ
- Không gian lưu trữ lớn
- Thường được nén để giảm kích thước (mp3, aac, Flac, ...)



Multimedia data

Dữ liệu video:

- Video số gồm một chuỗi các khung hình (frames) (25, 30, 50 khung hình/giây)
- Input: video camera → số hóa
- Định dạng: đa dạng (mp4, avi. . .)



Multimedia data

Dữ liệu video:

- Không gian lưu trữ: lớn
 - Tùy thuộc vào độ phân giải và kích thước, 1 khung hình có thể cần 1MB
 - Video 512×512 đơn sắc: $25 \times 0.25 = 6.25$ Mb/s (chưa nén)
 - PAL video (720×576 pixel / khung màu):
 $25 \times 1.2 = 30$ Mb/s (chưa nén)
 - High Definition DVD ($1440 \times 1080 = 1.5$ Megapixel/khung): $4.5 \times 25 = 112.5$ Mb/s (chưa nén)
- Dữ liệu phải được nén

Multimedia data: Đặc điểm

- Kích thước dữ liệu lớn → cấu trúc dữ liệu đặc biệt: lưu trữ và đánh chỉ mục
- Có chiều thời gian (audio, video)
- Dữ liệu được biểu diễn thông qua chuỗi các giá trị riêng lẻ, thiếu cấu trúc ngữ nghĩa rõ ràng để máy tính có thể “hiểu” nội dung
- Nhiều ứng dụng yêu cầu biểu diễn nhiều kiểu dữ liệu đồng thời có thông số thời gian và không gian

Multimedia data: Đặc điểm

- Nhiều ứng dụng yêu cầu biểu diễn nhiều kiểu dữ liệu đồng thời có thông số thời gian và không gian
- Giàu thông tin: cần nhiều tham số để có thể biểu diễn được nội dung

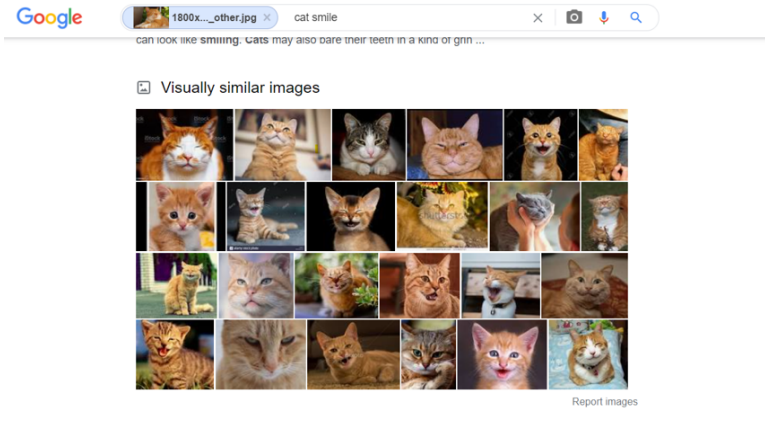
Multimedia data: Thách thức

- Biểu diễn nội dung
 - Độ tin cậy, ngữ nghĩa → phương pháp: tự động hoặc/và thủ công
- Biểu diễn câu hỏi/ câu trả lời: biểu diễn dữ liệu phức hợp
- Dữ liệu lớn → vấn đề lưu trữ, truy nhập và truyền
- Thời gian truy vấn (audio, video)
- Trích chọn thuộc tính tự động và đánh chỉ mục

Truy vấn dữ liệu

- Hệ CSDL thông thường (conventional database system): quản lý, tìm kiếm dữ liệu có cấu trúc
- **Hệ thống truy vấn thông tin (Information Retrieval System)**
 - Tìm kiếm dữ liệu trong tập văn bản lớn
 - Biểu diễn nội dung: từ khóa, tóm tắt
 - Truy vấn: từ khóa, ngôn ngữ tự nhiên
- **Hệ thống dựa trên nội dung (Content-based Retrieval System)**
 - Dựa trên đặc trưng của dữ liệu: màu sắc, hình dạng, kết cấu...
 - Hiệu năng của IR \gg CBR do keyword có thể diễn tả ngữ nghĩa

Truy vấn dữ liệu



Content-based Image Retrieval

Truy vấn dữ liệu

The screenshot shows a Google search for "information retrieval". The search bar at the top contains the text "information retrieval" with a search icon to its right. Below the search bar, there are tabs for "All", "Images", "Videos", "Books", "News", and "More". The "All" tab is selected. The search results show "About 72,900,000 results (0.43 seconds)". The first result is from "en.wikipedia.org" and is titled "Information retrieval - Wikipedia". The snippet below the title states: "Information retrieval is the science of searching for information in a document, searching for documents themselves, and also searching for the metadata that describes data, and for databases of texts, images or sounds." Below the snippet are links for "Category:Information retrieval", "Boolean model", "Evaluation measures", and "Applications". The second result is from "nlp.stanford.edu" and is titled "Introduction to Information Retrieval - Stanford NLP Group". The snippet below the title states: "Introduction to Information Retrieval. By Christopher D. Manning, Prabhakar Raghavan & Hinrich Schütze. Website: <http://informationretrieval.org/>, Cambridge ...". Below the search results is a section titled "People also ask" with four questions: "What do you mean by information retrieval?", "Why is information retrieval important?", "What is information retrieval tools?", and "What are the components of information retrieval?". To the right of the search results is a knowledge panel titled "Information retrieval" with a sub-header "Information retrieval". The panel contains a definition: "Information retrieval is the process of obtaining information system resources that are relevant to an information need from a collection of those resources. Searches can be based on full-text or other content-based indexing. [Wikipedia](#)". Below the definition are two expandable sections: "Psychology" and "Types". At the bottom of the panel is a section titled "People also search for" with five icons and labels: "Memory", "Information", "Semantics", "Machine Learning", and "Natural Language".

Keyword-based Information Retrieval

MIRS

Dữ liệu đa phương tiện:

- Thu thập và lưu trữ ngày càng nhiều
 - Máy tính cá nhân
 - Internet: Flickr, Facebook, Instagram, Youtube, ...
 - Đặc điểm khác biệt so với dữ liệu số truyền thống
- Hệ quản trị CSDL truyền thống không còn phù hợp để xử lý
- Các kỹ thuật IR có thể hỗ trợ nhưng không đủ để xử lý hiệu quả

MIRS: Multimedia Indexing and Retrieval System

MIRS = DBMS + IR + CBR techniques

MIRS

DBMS

- Sử dụng cho dữ liệu có cấu trúc liên quan đến dữ liệu đa phương tiện (ngày, tác giả, ...)
- **Object-Relational DBMS**: hỗ trợ cho dữ liệu đa phương tiện

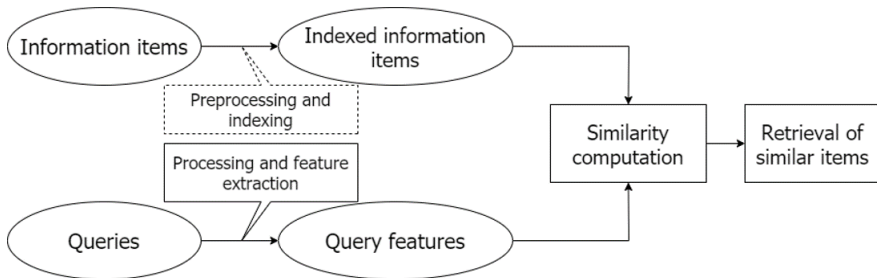
IR: text-based retrieval

- Dữ liệu văn bản chiếm phần lớn
- Chú thích cho dữ liệu đa phương tiện

Một điểm quan trọng nhất trong MIRS: ***trích chọn đặc trưng/biểu diễn nội dung***

MIRS hoàn chỉnh \sim MM - DBMS

Mô hình IR tổng quát



MM-DBMS

MM-DBMS: Multimedia Database Management System

- Framework quản lý các *kiểu dữ liệu khác nhau* với *định dạng phong phú* và được lưu trên nhiều *nguồn phương tiện khác nhau*

Các yêu cầu tương tự DBMS truyền thống

- Tích hợp (Integration)
- Độc lập dữ liệu (Data independence)
- Điều khiển tương tranh (Concurrency control)
- Tính bền vững (Persistence)

MM-DBMS

Các yêu cầu tương tự DBMS truyền thống

- Tính riêng tư (Privacy)
- Toàn vẹn (Integrity control)
- Phục hồi dữ liệu (Recovery)

Ngoài ra, phải đảm bảo

- Truy vấn dữ liệu đồng nhất với các dữ liệu có định dạng khác nhau
- Truy vấn đồng thời từ nhiều nguồn

→ hỗ trợ truy vấn (query support)

MM-DBMS

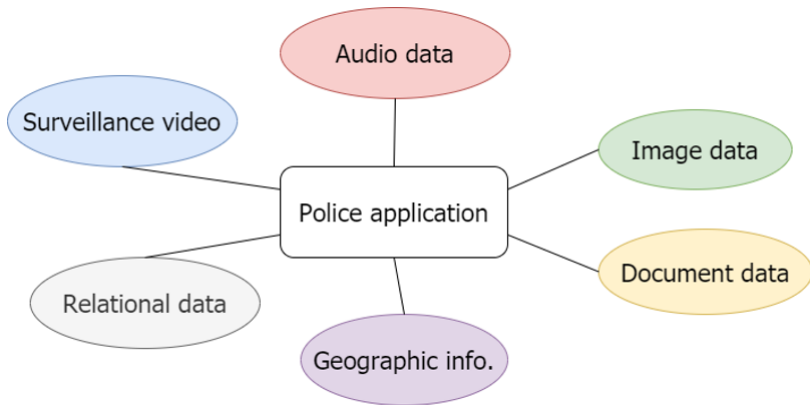
Ngoài ra, phải đảm bảo

- Truy xuất các đối tượng từ các thiết bị lưu trữ mà không bị rung/giật (video, audio)
- ➔ hỗ trợ lưu trữ (storage support)
- Có thể biểu diễn, truyền tải câu trả lời dưới dạng phương tiện nghe nhìn, đảm bảo các yêu cầu QoS
- ➔ hỗ trợ trình diễn và truyền dữ liệu (presentation and delivery support)

Một số vấn đề chính

- Hỗ trợ truy vấn
 - Ngôn ngữ truy vấn cho phép dễ dàng truy vấn đến CSDL đa phương tiện
 - Cho phép truy vấn hiệu quả
- Mô tả nội dung
 - Trích chọn đặc trưng và biểu diễn nội dung thể nào từ dữ liệu đa phương tiện
 - Đánh chỉ mục (indexing) các nội dung thể nào?
- Hỗ trợ lưu trữ
- Trình diễn và truyền dữ liệu

Ví dụ một ứng dụng đa phương tiện

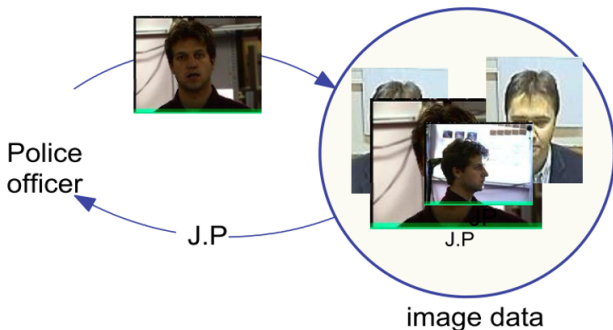


Điều tra của công an về hoạt động ma túy trên diện rộng

Truy vấn trong thư viện ảnh

Ngữ cảnh 1

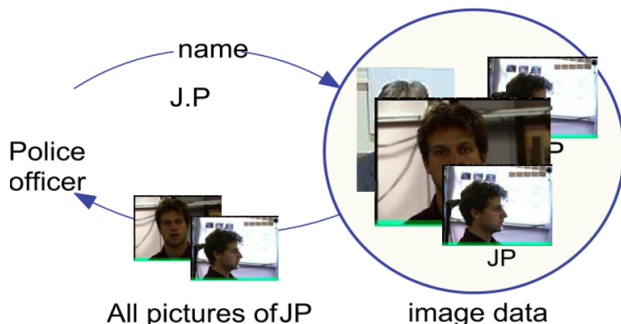
- Cảnh sát có ảnh của 01 đối tượng tình nghi và muốn tìm ra danh tính
- Truy vấn Q1: tìm tất cả các ảnh từ thư viện giống với ảnh đang có



Truy vấn trong thư viện ảnh

Ngữ cảnh 2

- Cảnh sát cần kiểm tra các bức ảnh của JP
- Truy vấn Q2: Tìm tất cả các bức ảnh trong thư viện có hình của JP



Truy vấn trong thư viện ảnh

Truy vấn sử dụng ảnh

- Similarity: độ tương tự giữa 02 ảnh
- Ranking: xếp loại

Truy vấn sử dụng từ khóa

- Đối tượng ảnh liên kết với giá trị thuộc tính
- Đánh chỉ mục và tìm kiếm hiệu quả các thuộc tính

Truy vấn dữ liệu âm thanh

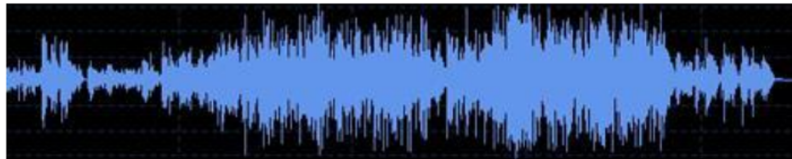
Ngữ cảnh 1

- Đoạn ghi âm chứa đoạn đối thoại của 02 người A và B. A là Denis Dopeman
- Q1: Xác định danh tính của B

Ngữ cảnh 2

- Nghe tất cả các đoạn băng có Denis Dopeman trong 1 khoảng thời gian nào đó
- Q2: tìm tất cả các băng ghi âm mà Denis Dopeman có tham dự

Truy vấn dữ liệu âm thanh



Dialogue between Denis and B



Who is B ?

All audio tapes in which Denis was a participant

Truy vấn dữ liệu văn bản

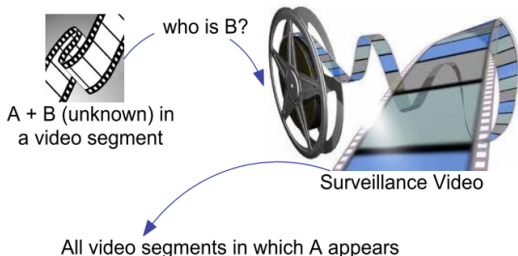
Ngữ cảnh

- Tra cứu tài liệu để điều tra 1 vụ việc chưa được giải quyết
- Q1: tìm tất cả các tài liệu có giao dịch tài chính giữa công ty X và công ty Y

Truy vấn dữ liệu video

Ngữ cảnh

- Cảnh sát đang xem đoạn video theo dõi 01 vụ tấn công, mặt của kẻ tấn công (B) bị che khuất 1 phần → giải thuật nhận dạng không cho kết quả tốt
- Phỏng đoán: có thể người bị tấn công (A) quen biết thủ phạm
- Q1: Tìm tất cả các đoạn video mà có mặt người bị tấn công



Truy vấn hỗn hợp

Truy vấn phức tạp: liên quan đến nhiều nguồn đa phương tiện khác nhau

→ Khó, ngay cả trong trường hợp chỉ có dữ liệu văn bản

Truy vấn hỗn hợp

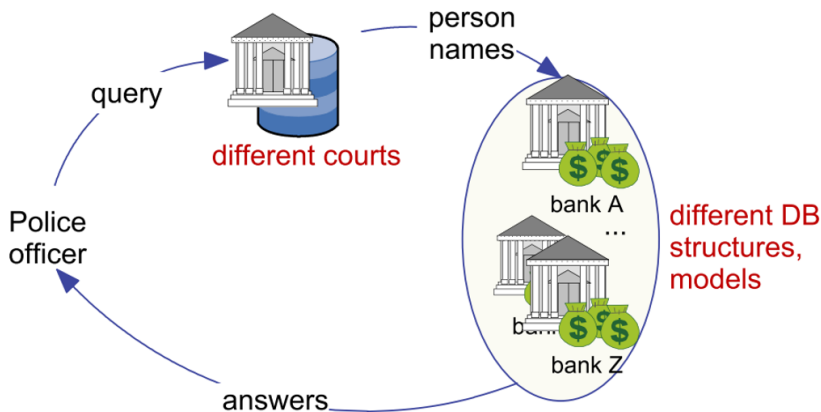
Ví dụ HQ1 truy vấn dữ liệu văn bản: tìm tất cả những kẻ

- Đã bị kết án trong vụ tấn công ở Nam Phi và
- Đã có chuyển khoản vào tài khoản của họ từ công ty ABC

Vấn đề:

- Việc tìm kiếm tất cả những kẻ đã bị kết án do các tội khác nhau → cần truy cập đến nhiều CSDL đa dạng thuộc vào các phán xử, tòa án khác nhau
- Công ty ABC có thể có nhiều tài khoản ở rất nhiều ngân hàng trên thế giới (định dạng khác nhau, hệ CSDL khác nhau)

Truy vấn hỗn hợp



Truy vấn hỗn hợp

Ví dụ truy vấn HQ2: Tìm tất cả những người

- Đã bị kết án trong vụ tấn công ở Nam Phi
- Đã có chuyển khoản vào tài khoản của họ từ công ty ABC
- Có chụp chung hình với Jose

Thực hiện truy vấn:

- Giống HQ1
- Truy nhập dữ liệu ảnh có gán nhãn để tìm ảnh của người thỏa HQ1 với tên đã được biết
- Tìm trong CSDL ảnh tĩnh, trong video đối tượng xuất hiện cùng Jose
- Xác định đối tượng trên ảnh, video nhờ vào giải thuật xử lý ảnh

Truy vấn hỗn hợp

