Giới thiệu về quản lý tri thức

Hồ Tú Bảo
Trường Khoa học Tri thức
Viện Khoa học và Công nghệ Tiên tiến Nhật bản (JAIST)
Viện Công nghệ Thông tin
Viện Khoa học và Công nghệ Việt Nam
bao@jaist.ac.jp

Japan Advanced Institute of Science and Technology (JAIST)

Viện đại học quốc gia, đào tạo thạc sĩ và tiến sĩ theo các ngành khoa học và công nghệ chọn lọc

- Trường Khoa học Thông tin (1992)
- Trường Khoa học Vật liệu (1993)
- Trường Khoa học Tri thức (1998)

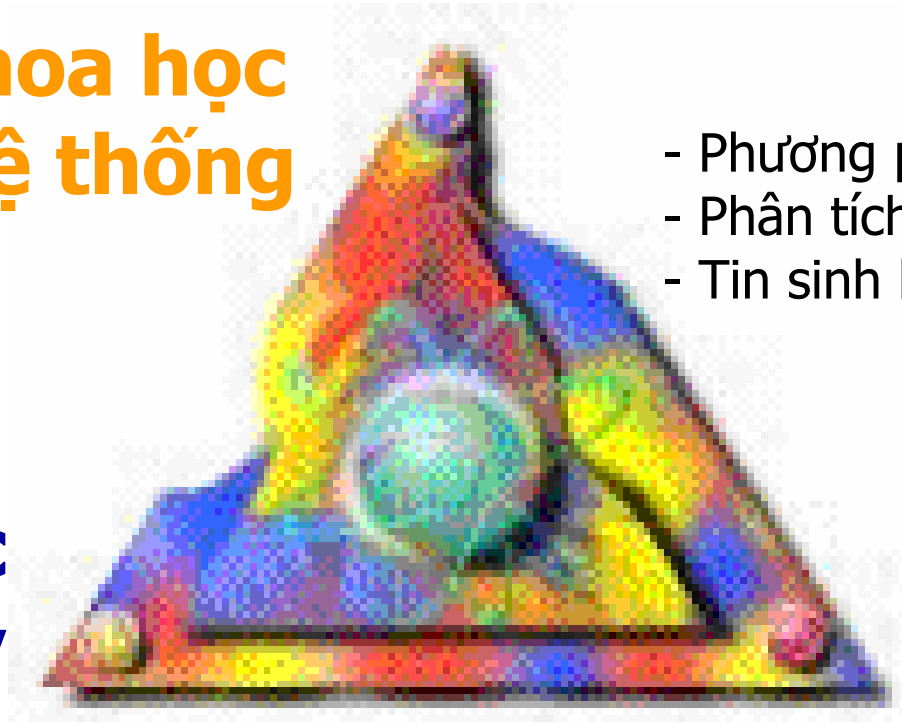
- 48 labs, 300 nhân viên
- 750 s/v master, 350 NCS tiến sĩ

Hợp tác với VKHCNQG, ĐHKHTN, ĐHBK, 25 NCS Việt nam

Trường Khoa học Tri thức (JAIST)

Khoa học Hệ thống

- Phương pháp luận về hệ thống
- Phân tích hệ thống phức tạp
- Tin sinh học (bioinformatics)



Khoa học Quản lý

- **Quản lý tri thức**
- Quá trình ghiên cứu và phát triển
- Quản lý xã hội

Khoa học Thông tin

- Phát hiện tri thức từ dữ liệu
- Cấu trúc hóa tri thức
- Hỗ trợ sáng tạo

Nội dung

Quản lý tri thức (KM)

Quản lý tri thức ở Nhật bản

CNTT và quản lý tri thức

Quản lý tri thức (knowledge management)?

- Tri thức nói nôm na là “hiểu biết”
- Quản lý tri thức là “quản lý” hiểu biết của/trong các tổ chức
- Vai trò của CNTT và truyền thông trong quản lý tri thức?
- Tri thức của thiên hạ trong sự phát triển của một quốc gia, của một tỉnh, một tổ chức?

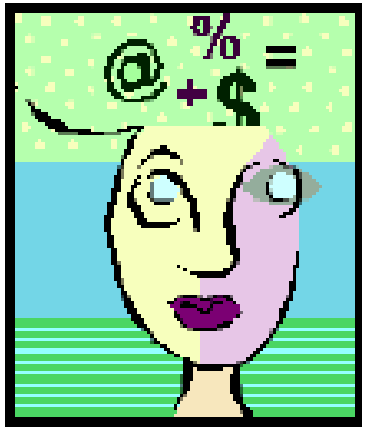
Tri thức và phát triển

Rất nhiều câu hỏi của sự phát triển:

- Đông Nam Á trong quan hệ Mỹ, Nhật, Trung quốc và ảnh hưởng đến Việt Nam?
- WTO: lợi, hại và thách thức?
- Khác biệt về hệ thống giáo dục ở Việt Nam, Mỹ và một số nước châu Âu?
- Gia công phần mềm: đặc điểm thị trường phần mềm Nhật bản?
- …

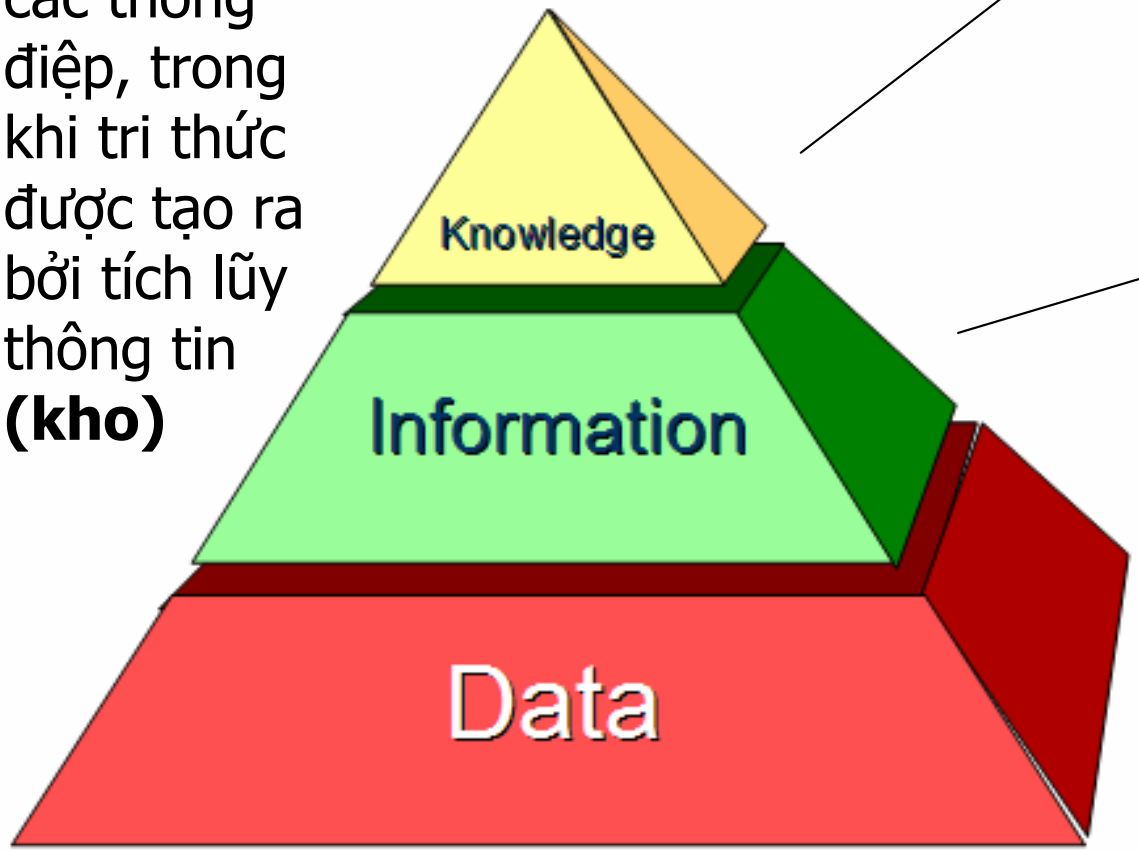
Gần đây ta thường nghe

- Tri thức đang trở nên tiềm lực cạnh tranh
- Tri thức là sức mạnh và tài nguyên
- Nâng cao dân trí
- Kinh tế tri thức
- ...



Dữ liệu, thông tin, tri thức

Thông tin là **dòng** chảy các thông điệp, trong khi tri thức được tạo ra bởi tích lũy thông tin **(kho)**



Hiểu biết đã được kiểm nghiệm, cần cho quyết định và hành động

Chuồn chuồn bay thấp thì mưa

Dữ liệu kèm theo ý nghĩa (do được xử lý)

Chuồn chuồn bay vậy là bay thấp

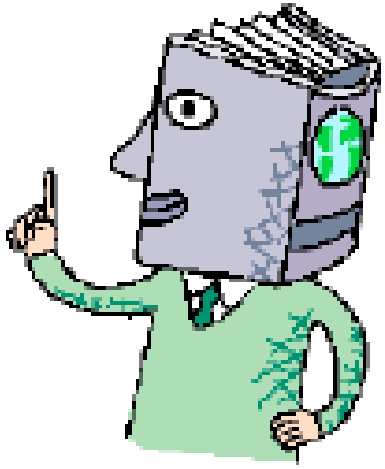
Tín hiệu quan sát, đo đạc được

0.3m, 0.5m, 0.6m, 0.4m, 0.3m, … khoảng cách chuồn chuồn bay cách mặt đất.

Quản lý tri thức là gì? Vài định nghĩa

- Các quá trình xác định, sáng tạo, nắm bắt, xử lý, chuyển giao, lưu trữ, và **sử dụng** tri thức để đạt những **giá trị chiến lược**.

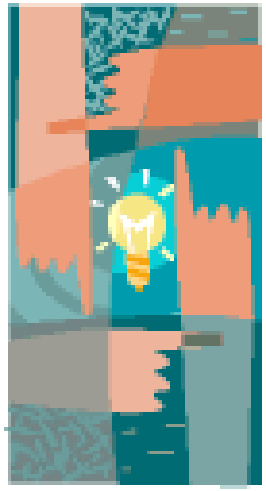
- Tạo dụng tri thức là một tập hợp các quá trình điều hành sự **sáng tạo, phổ biến, và sử dụng** tri thức.



Quản lý tri thức là gì?Vài định nghĩa

- Chiến lược có ý thức để có được **tri thức cần thiết cho đúng người cần ở đúng lúc** cần, giúp **mọi người chia sẻ và đưa tri thức vào hành động** theo nhiều cách để nâng hiệu quả hoạt động của tổ chức.

- Tạo dụng tri thức là cách các tổ chức **sáng tạo, nắm giữ, và tái sử dụng tri thức** để đạt các mục tiêu của mình.



Quản lý tri thức là gì?Vài định nghĩa

- Sử dụng và quản trị tri thức của tổ chức dựa trên thực tiễn **quản lý thông tin**, tập trung vào việc **học tập của tổ chức**, nhận thức về đóng góp và giá trị của người lao động, và có thể **thực thi nhờ vào công nghệ.**

- Một khái niệm trong đó một doanh nghiệp **thu thập, tổ chức, chia sẻ, và phân tích tri thức** như tài nguyên, tài liệu, kỹ năng của con người một cách có ý thức và toàn diện.

Quản lý tri thức là gì?Vài định nghĩa

- Quản lý tri thức tôi nói ở đây không phải là sản phẩm phần mềm hoặc một phạm trù phần mềm nào cả. Quản lý tri thức thậm chí không bắt đầu với công nghệ. Nó bắt đầu với mục tiêu kinh doanh và các quá trình và nhận thức về sự cần thiết của việc chia sẻ thông tin.

- Quản lý tri thức không là gì khác ngoài việc **quản lý dòng thông tin**, nắm lấy thông tin chính xác cho những người cần đến thông tin sao cho họ có thể hành động nhanh chóng với thông tin (Bill Gates).

Quản lý tri thức là gì? Vài định nghĩa

Quản lý tri thức liên quan đến việc nâng cao hiệu quả hoạt động của mọi tổ chức, gồm hai vấn đề chính:

- Tạo được tri thức cần thiết
- Dùng được tri thức để nâng hiệu quả

(Khái niệm "quản lý" ở đây không hoàn toàn theo nghĩa thông thường).

Nội dung

Quản lý tri thức (KM)

Quản lý tri thức ở Nhật bản

CNTT và quản lý tri thức

Con đường thành công của Nhật bản

- Chúng ta đã và đang tồn tại và phát triển trong một môi trường với duy nhất một điều chắc chắn là sự không chắc chắn của chính môi trường.
- Đối đầu với các chuyện sống còn này, các công ty Nhật bản đã tìm ra con đường thành công riêng của mình (luôn là điều bí ẩn?).
- Con đường thành công của Nhật bản trong kinh doanh có thể được nhìn nhận tóm tắt là

sáng tạo tri thức → liên tục cách tân → ưu thế cạnh tranh

Idea: Sáng tạo tri thức với sự tham gia của mọi cá nhân trong tổ chức

The Knowledge-Creating Company

Nonaka, I. and Takeuchi, H. (1995), Oxford University Press.

I. Nonaka

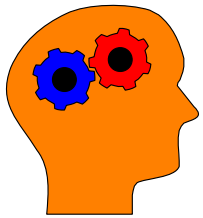
- Xerox Distinguished Professor, Hiệu trưởng sáng lập của trường Khoa học Tri thức, Viện Khoa học và Công nghệ Tiên tiến Nhật bản.
- Tác giả của lý thuyết sáng tạo tri thức, chỉ ra sự thành công của các công ty Nhật là nhờ thành công trong quản lý tri thức.
- Xây dựng khái niệm tri thức hiện (explicit) và tri thức ngầm (tacit), lý thuyết về tương tác của hai loại tri thức.
- "the Best Book of the Year in Business and Management in 1996". Các công ty Nhật bản tạo ra sự năng động trong cách tân (innovation) như thế nào?

Hai loại tri thức

Tri thức hiện (explicit knowledge)

- diễn đạt bằng ngôn ngữ hình thức, dễ trao đổi giữa các cá nhân.
- có thể biểu diễn bằng các công thức khoa học, các thủ tục tường minh, hoặc nhiều cách khác.
- bao gồm thông tin, dữ liệu, sách báo, văn bản, tài liệu đã được hệ thống bằng nhiều phương tiện.

Tri thức ngầm (tacit knowledge)



- có được và ẩn chứa trong kinh nghiệm của từng cá nhân, mang tính chủ quan, bao gồm những hiểu biết riêng thấu đáo, trực giác, linh cảm, kỹ năng, ...
- khó trao đổi hoặc chia sẻ với người khác.
- chỉ có thể học được từ người khác nhờ quan hệ gần gũi trong một khoảng thời gian nào đó.

Hai loại tri thức

Tri thức hiện (explicit)

- Tiếp cận lý thuyết
- Các giải quyết vấn đề
- Tài liệu
- Cơ sở dữ liệu
- Cơ sở tri thức

Tri thức ngầm (tacit)

- Nhận thức
  - Niềm tin
  - Quan niệm
  - Trực giác
  - Mô hình ẩn dụ
- Kỹ thuật
  - Ngón nghề (craft)
  - Bí quyết (know-how)

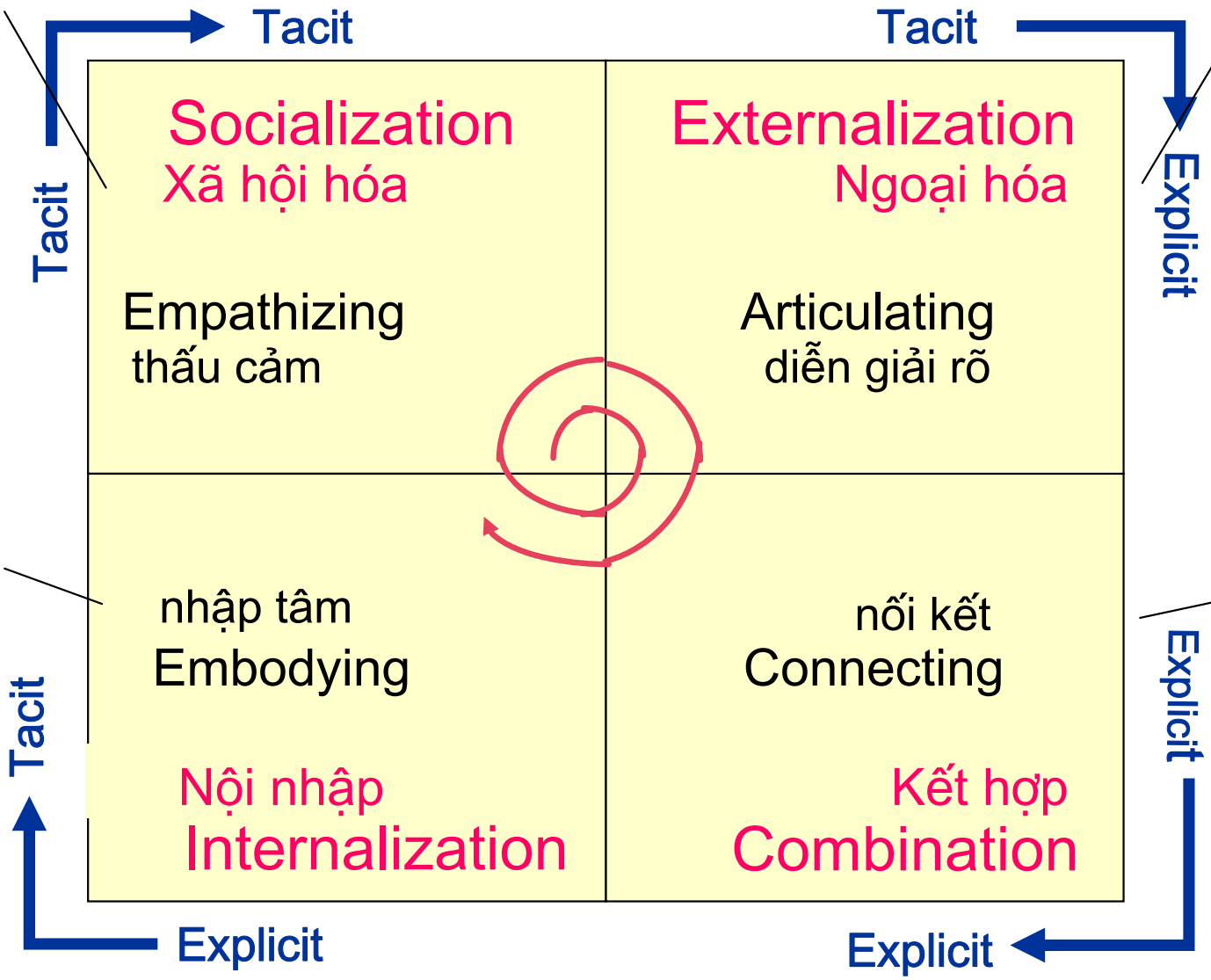
Một thí dụ từ tri thức ngầm đến tri thức hiện

- 1978: Honda muốn tạo một loại xe hơi mới, giao trách nhiệm cho một nhóm kỹ sư trẻ (trung bình 27 tuổi).
- (1) sản phẩm với khái niệm cơ bản khác trước, (2) xe phải không đắt không rẻ (mở đường cho sáng tạo)
- Khẩu hiệu “Automobile revolution”. Câu hỏi: “Nếu xe hơi là một thực thể sống, nó sẽ tiến hóa thế nào?”
- Ý tưởng: Xu hướng “cách mạng” là xe hơi phải vượt qua những quan hệ người-xe truyền thống → xe phải ngắn hơn và cao hơn, hình cầu sẽ cho nhiều chỗ hơn bên trong hơn và tiết kiệm năng lượng “Tall boy” car.

Lý thuyết chuyển đổi tri thức (Nonaka)

Qua giao tiếp xã hội hoặc chia sẻ kinh nghiệm trong các thành viên. Thí dụ: qua học nghề.

Chuyển tri thức riêng (ngầm) thành tri thức hiện. Thí dụ, do hệ thống, diễn giải các kinh nghiệm, các bài học, …



Chuyển tri thức chung (hiện) thành tri thức cá nhân. Thí dụ, do học và hiểu từ việc đọc và thảo luận.

Tạo tri thức hiện mới bằng ghép nối, phân loại, hợp nhất, tổng hợp các tri thức hiện đã có.

Vài đặc điểm về Nhật bản

- Tiếp thu và đồng hóa những tư tưởng và kỹ thuật từ nước ngoài
- Lập kế hoạch, tổ chức và thực hiện các dự án lớn
- Khả năng về nghệ thuật cũng như kỹ thuật phát triển ở một trình độ rất cao
- Giáo dục: coi trọng tinh thần yêu nước cũng như sự quy củ, khuôn phép
- Ý thức rất cao về bản sắc văn hóa dân tộc
- Mang tính cách Á đông, nhưng cũng chịu nhiều ảnh hưởng của phương Tây (theo phong cách riêng của Nhật bản!)
- Khao khát và ước mong giảm trong giới trẻ!

Nhật bản và quản lý tri thức

- Thế kỷ 16: học hỏi (chế tạo súng)
- Từ thế kỷ 17 đến giữa thế kỷ 19th (học hỏi từ Hà Lan)
- Giai đoạn Meiji: **"Cần phải tìm kiếm tri thức từ mọi nơi trên thế giới"**.
- Thế kỷ 19-20 Nhật bản tiếp thu tinh hoa của thế giới nhằm đẩy mạnh vị thế quốc gia (thông qua chiến tranh và hòa bình)

Đặc điểm chính về phong cách quản lý Nhật Bản

- Trọng người tài, trọng quan hệ, trọng chữ tín, và giữ truyền thống cùng lúc với các thay đổi.
- Làm việc suốt đời cho công ty (thay đổi)
- Chế độ đãi ngộ dựa trên thâm niên công tác (?)
- Các **quyết định được đưa ra dựa trên sự nhất trí cao** (đồng thuận)

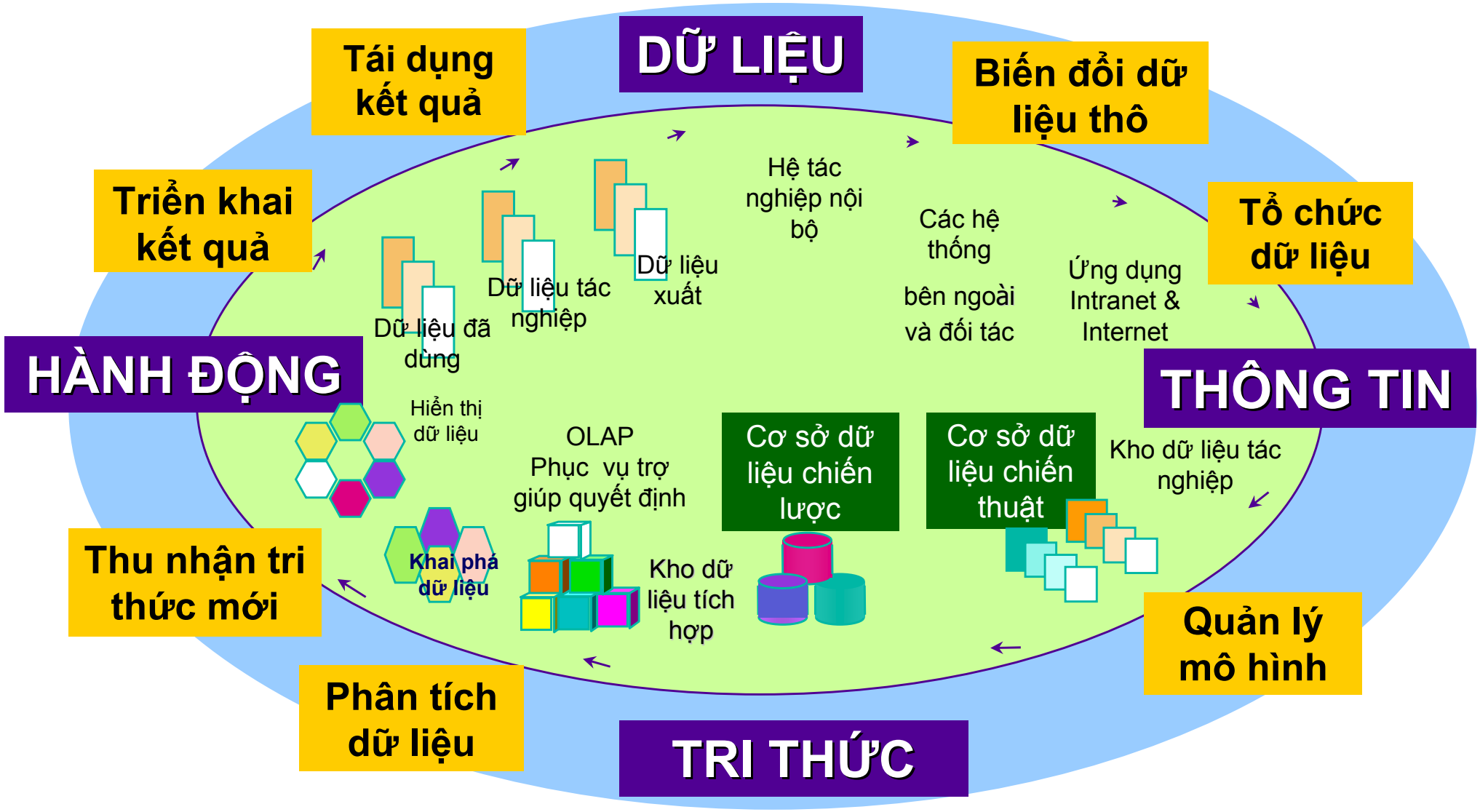
Nội dung

Quản lý tri thức (KM)

Quản lý tri thức ở Nhật bản

CNTT và quản lý tri thức

Vòng kín của quản lý tri thức và tác động của CNTT

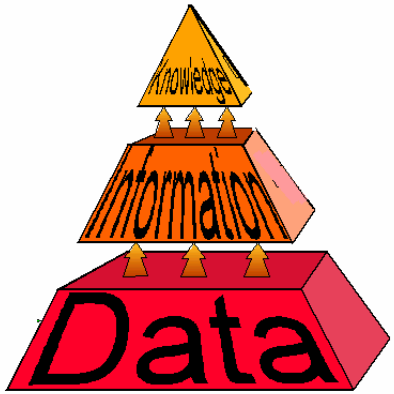


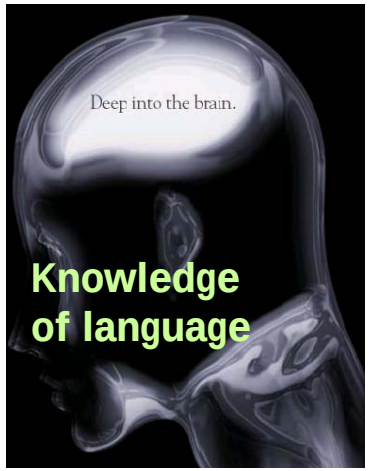
Web là nguồn tri thức khổng lồ

- Khoảng 80% dữ liệu trên đời ở dạng không cấu trúc: văn bản (unstructured formats, source: Oracle Corporation).
- Đặc điểm của Web: hypertext và hyperlink, nội dung trong nhiều ngôn ngữ khác nhau.
- Web thay đổi và ảnh hưởng sâu sắc đến cách học và thu nhận tri thức.
- Làm sao có được tri thức trên Web?

Xử lý tiếng Việt?

- Công nghệ thông tin: xử lý dữ liệu để tìm thông tin, tri thức.
- Xử lý ngôn ngữ và tiếng nói: xử lý dữ liệu thu được từ tiếng nói và văn bản của con người.
- Thí dụ:
  - Nhận dạng tiếng nói (tiếng nói → văn bản), tổng hợp tiếng nói (văn bản → tiếng nói)
  - Dịch Anh-Việt, Việt-Anh (văn bản → văn bản)
  - Tìm kiếm thông tin trên Web bằng Google (cơ sở dữ liệu văn bản khổng lồ, “dioxin”, “law”)
  - Trích chọn thông tin cần thiết từ Web/thư viện điện tử (hỏi: “những điều luật nào trong luật của Mỹ cản trở ta trong vụ kiện chất độc màu da cam”?)
- Thiết yếu: máy tính phải có tri thức về ngôn ngữ

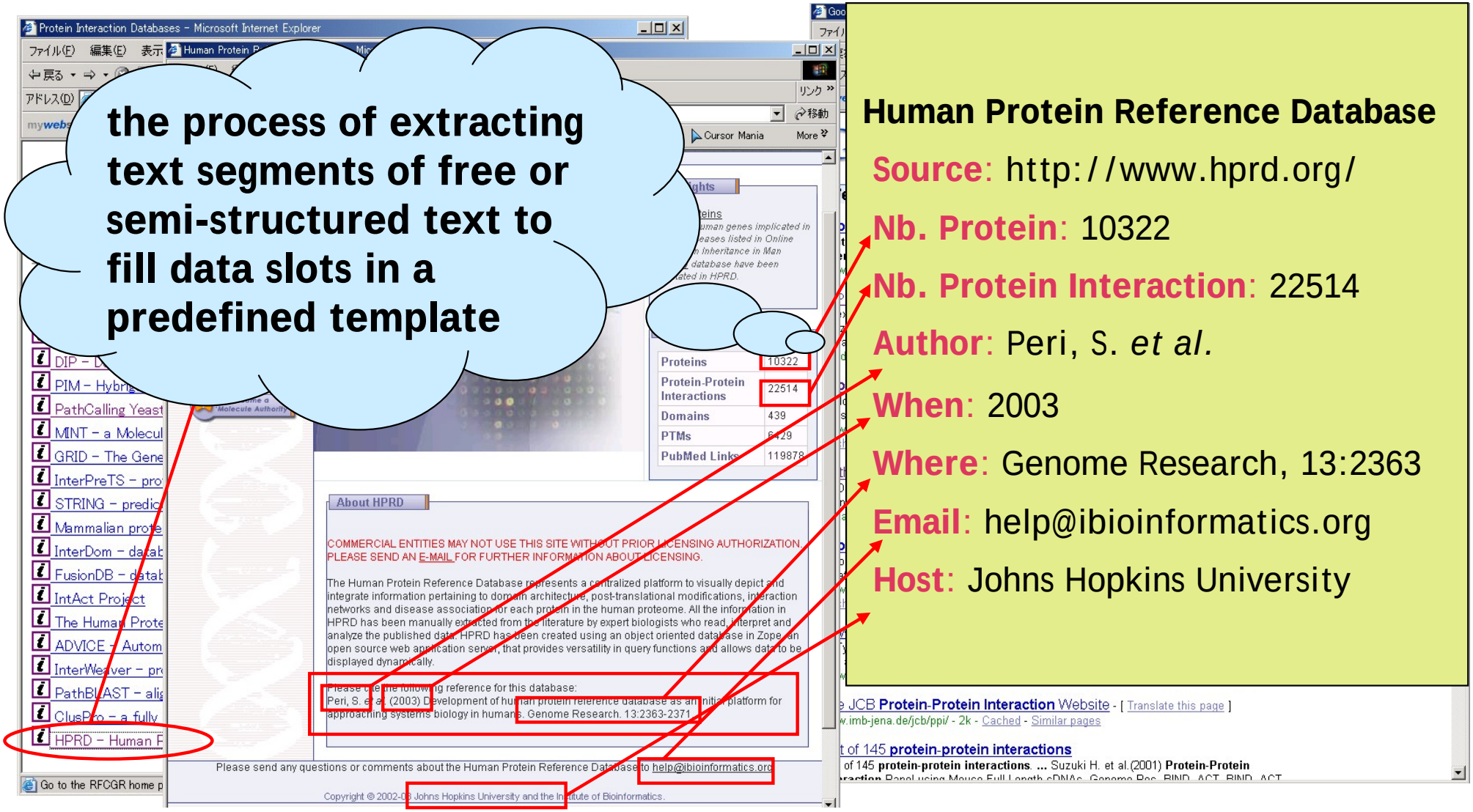




Khó khăn: nghĩa của từ và câu

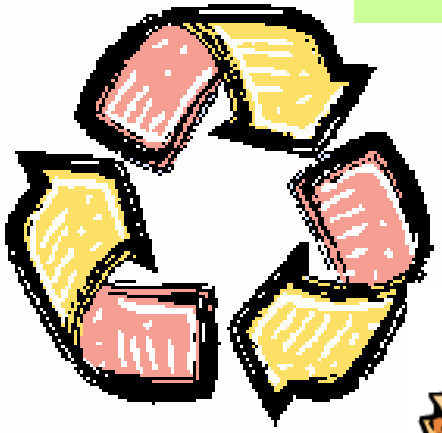
- "Xét một từ, chẳng hạn "sợi dây". Ngày nay không một máy tính nào có thể hiểu nghĩa từ này như con người. Ta có thể kéo một vật bằng một sợi dây, nhưng không thể đẩy một vật bằng sợi dây, nhưng không thể ăn sợi dây. Ta có thể gói một gói hàng hoặc thả diều bằng một sợi dây, nhưng không thể ăn sợi dây này. Trong vài phút, một đứa trẻ nhỏ có thể chỉ ra hàng trăm cách dùng, hoặc không dùng, một sợi dây, nhưng không máy tính nào có thể làm việc này." (Marvin Minsky, 1992)
- "Ông già đi nhanh quá" có thể hiểu nhiều cách khác nhau.

Information extraction vs. Information retrieval: Finding "things" but not "pages"



Khai phá văn bản & thông tin trên Web

Ngữ dụng và khai phá dữ liệu văn bản & Web



Phương pháp xử lý ngôn ngữ và tiếng nói tiếng Việt

Công cụ, dữ liệu, tài nguyên, phương tiện

- Nhận dạng và tổng hợp tiếng nói
- Dịch tự động
- Tóm tắt văn bản
- Tìm kiếm thông tin
- Trích chọn thông tin
- Quản trị cơ sở dữ liệu văn bản
- Khai phá dữ liệu văn bản và Web
- Ứng dụng khác

Nhu cầu tìm thông tin trên Web

- Sự tăng nhanh lượng dữ liệu text và phi cấu trúc trên mạng Internet
- Nhu cầu tìm kiếm thông tin của người dùng cuối
  - Tìm kiếm thông tin giải trí
  - Tìm kiếm thông tin phục vụ chuyên ngành
  - Tìm kiếm thông tin phục vụ cho báo chí, xuất bản
  - Tìm kiếm thông tin phục vụ cho quản lý
- Một số tiêu chí khi tìm kiếm thông tin
  - Thông tin “nóng” nhất
  - Thông tin chất lượng nhất (theo nghĩa nào đó)
  - Thông tin ngắn gọn nhất
  - Thông tin đầy đủ nhất

Nhu cầu tìm thông tin trên Web

Một vài thí dụ:

- Một sinh viên muốn tìm thông tin về "tiến trình gia nhập WTO của Việt Nam" (Google đưa ra 28000 trang web, chưa kể tài liệu tiếng Anh)
- Một người dân Nam Bộ tìm kiếm thông tin về "tình hình xuất khẩu cá Tra và cá Basa" (Google đưa ra 2540 trang web tiếng Việt, 1860 trang web tiếng Anh)
- Một nhà báo muốn duyệt nhanh các bài bình luận của các hãng tin (CNN, AP, WashingtonPost, etc.) về sự kiện "Việt Nam tổ chức kỷ niệm 30 năm ngày đất nước thống nhất"
- Một người kinh doanh muốn tổng hợp và tóm tắt "biến động của giá cà phê"

IREST: Information Retrieval, Extraction, Summarization and Translation

Khó khăn:

- Một người có thể mất hàng tiếng đồng hồ để ngồi lục lọi trong kết quả trả về của Google nhằm đọc, trích chọn, và tóm lược lại những thông tin cần thiết
- Một người dân bình thường không thạo ngoại ngữ sẽ chỉ tìm kiếm được thông tin từ những trang tiếng Việt
- Kiểm soát thông tin bằng mắt, bằng tay là công việc khó khăn và mất thời gian

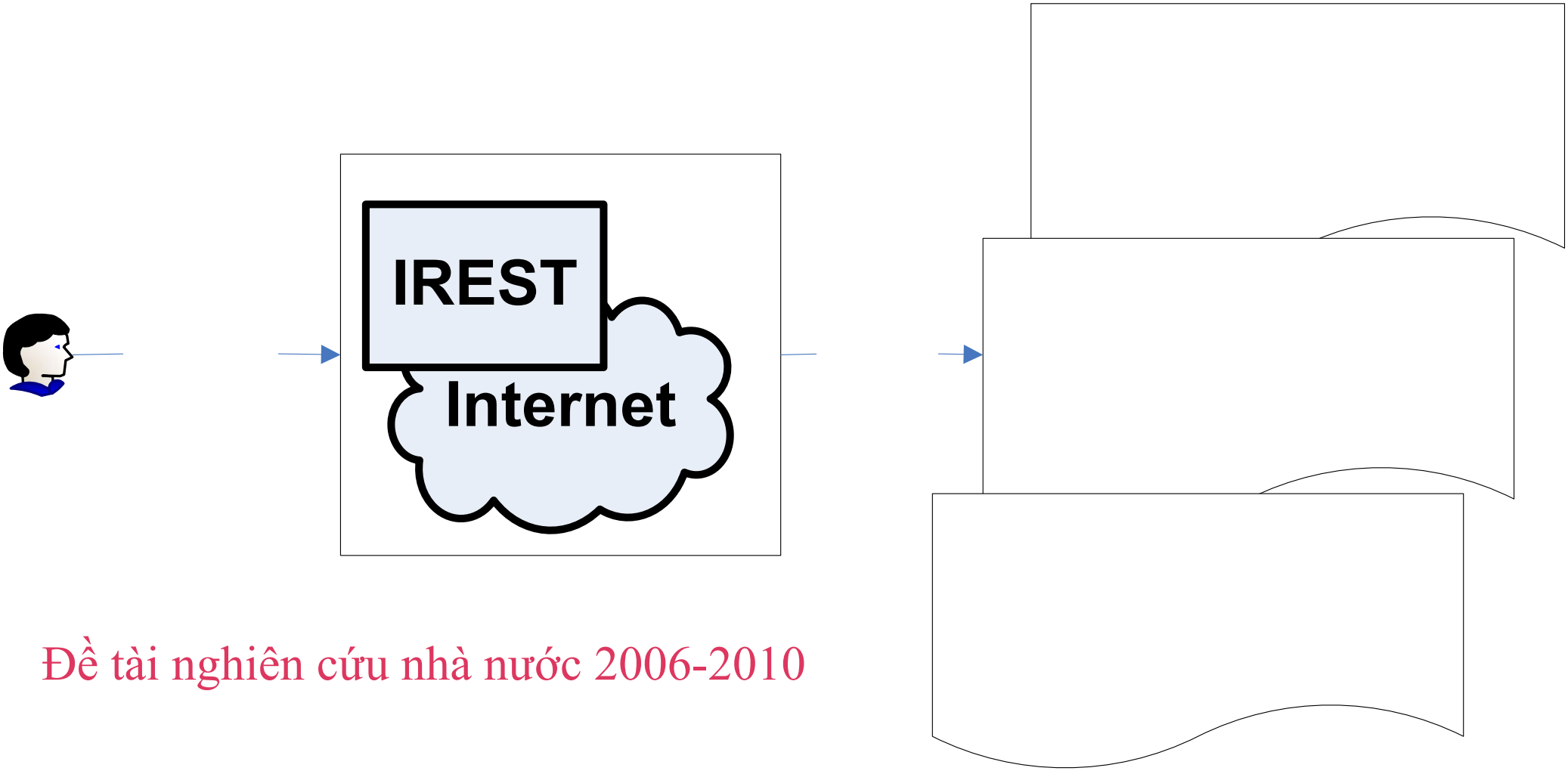
Một nỗ lực cho giải pháp (IREST)

- Một hệ thống cho phép tìm kiếm, trích chọn, và tóm tắt nhằm trả về cho người dùng những thông tin cô đọng nhất
- Có thể tìm kiếm và tóm tắt song hành tiếng Việt lẫn tiếng Anh nhằm tránh bỏ sót thông tin
- Hệ tìm kiếm có thể thực hiện offline các giao dịch (user query, transaction) của người dùng nhằm tiết kiệm thời gian
- Và một tiêu chí quan trọng: "Thông tin tự tìm đến người cần nó"

Yêu cầu của hệ thống

- Yêu cầu của người dùng bằng tiếng Việt và có thể ở nhiều dạng và mức độ phức tạp khác nhau
  - Yêu cầu là các từ khóa (keywords) => phục vụ các giao dịch đơn giản
  - Yêu cầu tuân theo một khuôn mẫu (template) được tổ chức theo một cấu trúc ngữ nghĩa nào đó => chủ yếu phục vụ cho các giao dịch phức tạp, hướng tới những dạng thông tin chuyên ngành
  - Yêu cầu là một câu tiếng Việt (ở mức độ đơn giản) => nhằm đơn giản hóa và tăng tính thân thiện của hệ thống đối với người dùng
- Hệ thống sẽ tìm kiếm, lọc, trích chọn, tóm tắt các tài liệu liên quan (cả tiếng Anh lẫn tiếng Việt) và trả lại cho người dùng các mẫu tóm tắt ngắn gọn bằng tiếng Việt

Yêu cầu của hệ thống (thí dụ minh họa)



Đề tài nghiên cứu nhà nước 2006-2010

Vài lời kết

- Tri thức (hiểu biết) là cốt tử cho sự phát triển.
- Quản lý tri thức liên quan với CNTT, và cần được tìm hiểu, thực hiện ở nước ta.
- Web là nguồn tri thức khổng lồ nhưng cần chuẩn bị để học/tìm được tri thức từ Web.

Chỉ dùng Unicode để soạn văn bản tiếng Việt!