

Các kỹ năng cơ bản tìm kiếm trên internet

Nguyễn Tuấn Anh – 45K

2007

HÀ NỘI

Các kỹ năng cơ bản tìm kiếm trên internet

Nguyễn Tuấn Anh – 45K

Đại học Thủy Lợi

Nội dung

1. Các thông tin phổ biến trên internet.....	2
2. Các công cụ tìm kiếm thông tin trên mạng:.....	3
1.1. Bộ máy tìm kiếm của Việt Nam:.....	3
1.2. Bộ máy tìm kiếm của nước ngoài:	3
3. Sử dụng công cụ tìm kiếm một cách phù hợp.....	4
3.1. Phân tích yêu cầu tìm	4
3.2. Diễn đạt lệnh tìm kiếm - giới thiệu về cú pháp của lệnh tìm	5
3.2.1. Các phép toán của lệnh tìm	5
3.2.2. Toán tử Boolean.....	6
3.2.3. Một số cách thu hẹp phạm vi tìm kiếm khác	6
3.3. Phân nhóm yêu cầu thông tin	8
3.4. Chọn công cụ tìm kiếm phù hợp	8
3.5. Tìm lời khuyên từ một người!	13
3.6. Nếu bước đầu bạn chưa thành công –,, hãy thử lại! ”	13
3.7. Đánh giá kết quả tìm	13
3.7.1. „Hãy suy nghĩ trước khi nhấn chuột”	13
3.7.2. Một địa chỉ web (URL) gồm có những yếu tố nào?	14
Tài liệu tham khảo:	15

1. Các thông tin phổ biến trên internet

Những thông tin gì bạn có thể tìm thấy trên Internet? Không thể liệt kê tất cả các loại nguồn tin có trên Internet, đặc biệt khi Internet thay đổi rất thường xuyên và nhanh chóng. Tuy nhiên, nhìn chung Internet hữu ích nhất khi tìm kiếm những thông tin trong các nhóm sau đây (chú ý phần lớn những thông tin này là bằng tiếng Anh)¹:

- Các sự kiện đang diễn ra, ví dụ: tin tức ngày hôm nay, hay những xu hướng mới nhất
- Thông tin kinh tế, ví dụ: thông tin cáo báo chí của một công ty, chỉ số chứng khoán, thông tin về sản phẩm
- Thông tin của chính phủ, ví dụ: các chính sách hiện hành, luật pháp, các cuộc thảo luận tại quốc hội, quyết định của tòa án, thông tin cáo báo chí .
- Văn hóa đại chúng, ví dụ: phim, nhạc, truyền hình, thể thao chuyên nghiệp
- Thông tin về máy tính và Internet, ví dụ: thông tin các khoá học.

Nhiều trang web mới và các “cựu binh” về tìm kiếm trên mạng đều đưa ra những dịch vụ chuyên biệt giúp bạn tìm được thông tin nhanh hơn. Nếu bạn chỉ dựa vào duy nhất một công cụ tìm kiếm thì đôi khi phải mất rất nhiều thời gian. Do đó, những “thấu kính” chuyên dụng để tìm thông tin cần thiết nhanh chóng và chính xác hơn đã ra đời.

Theo một chuyên gia của Search Engine Watch (www.searchenginewatch.com) „kỳ nguyên của những guồng máy tìm kiếm đang cố sức bao trùm tất cả mọi thứ, một phần cũng vì bây giờ

¹ Vũ Thị Nha, *Tìm kiếm thông tin trên internet*, Trung tâm Thông tin Phát triển Việt Nam

Nguyễn Tuấn Anh – 45K

July 28, 2007



chúng ta muốn kết quả tìm kiếm cụ thể hơn, và rõ ràng công cụ tìm kiếm chuyên biệt đang là xu hướng tất yếu”.

2. Các công cụ tìm kiếm thông tin trên mạng:

1.1. Bộ máy tìm kiếm của Việt Nam:

Hiện nay ở Việt Nam có 2 bộ máy tìm kiếm tạm được bạn có thể tham khảo như:

- [Vinaseek](#)
- [Panvietnam](#)

1.2. Bộ máy tìm kiếm của nước ngoài:

Bộ máy tìm kiếm ở nước ngoài rất đa dạng và phong phú, vấn đề là bạn phải biết cách tìm kiếm như thế nào để có được thông tin mình cần tìm.

- | | | |
|---------------------------------------|--|---------------------------------------|
| - Yahoo Search | - AboutSearch | - MSN Internet Search |
| - Altavista | - Asiaco | - Inference Find |
| - Infoseek | - Metacrawler | - Euro Seek |
| - Go.com Search | - Webcrawler | - Search Port |
| - Excite | - American Online Search | - Find What Search |
| - Lycos | - Simple Search | - Search Engine Guide |
| - Hotbot Lycos Search | - Businesseek | - Ask Jeeves Search |
| - Snap Search | - CNET Search | - Virtual Library |
| - Google | - Company Site Locator | - InfoHighway |
| - Lookmart | - Netscape Search | - Copernic |
| - Aj Search | - Direct Hit | - Ask |

Nhưng phổ biến và tiêu biểu có bốn bộ máy sau thường dùng để tìm kiếm thông tin, có thể nói đây là những bộ máy thông dụng và cho chúng ta những tài liệu cần tìm kiếm một cách nhanh chóng :

- [Google](#)
- [AltaVista](#)
- [Alltheweb](#) (trước đây là "Fast Search")
- [Ask](#)

Trong các máy tìm kiếm này thì có thể nói Google là lão làng và sử dụng nhiều nhất. Với các trang tìm kiếm nhiều phần thông dụng.

Những trang tìm kiếm chuyên biệt của Google²

<http://images.google.com> ([Google Image Search](#))

Tìm hơn 500 triệu hình ảnh được truy cập công khai do Google thu thập từ các site, tìm theo tên, mô tả và từ khóa

<http://labs.google.com/glossary> ([Google Glossary](#))

Tìm định nghĩa từ không chỉ ở các bộ từ điển mà còn tham khảo ở các Website khoa học, y tế, pháp luật,...

<http://froogle.google.com>

Tìm, duyệt và phân loại các sản phẩm trên các site bán sỉ hay bán đấu giá.

<http://www.google.com/options/universities.html> ([Google University Search](#))

Thông tin thu thập từ site của các trường đại học, gồm tin tức, lịch học,...

² Nguyễn Duy Quốc, haveplus@yahoo.com



<http://directory.google.com> (Google Directory)

Duyệt Web theo những chủ đề phân loại nhất định.

<http://answers.google.com> (Google Answers)

Hỏi và nhận câu trả lời từ những chuyên gia về những vấn đề khó. Trả phí 2.5 \$/ mỗi câu hỏi.

http://www.google.com/advanced_search (Google Advanced Search)

Tìm từ đây nếu bạn muốn thiết lập nhiều điều kiện tìm kiếm.

<http://groups.google.com> (Google Groups)

Duyệt kho lưu trữ của các nhóm tin Usenet từ cuối thập niên 1970, hoặc gửi tin mới tới các nhóm thảo luận.

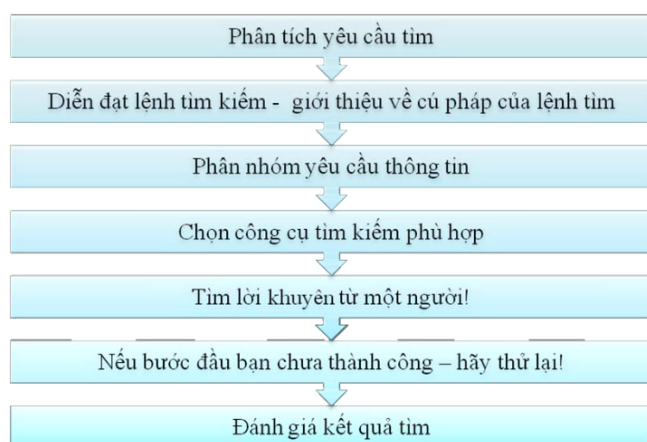
<http://news.google.com>

Tin tức chọn lọc từ các site tin tức trên khắp thế giới, cập nhật mỗi 5 phút

Hiện nay các thông tin trên mạng rất đa dạng bạn có thể tìm thấy thông tin trên mạng cần thiết mà không mất tiền. Hiện nay kho tàng thông tin vô tận của nhân loại thì bạn phải biết cách tìm đến thông tin mà bạn cần tìm.

3. Sử dụng công cụ tìm kiếm một cách phù hợp

Chiến lược tìm kiếm thông tin gồm 7 bước:



3.1 Phân tích yêu cầu tìm

Tự đặt các câu hỏi để làm rõ yêu cầu tin của mình. Ví dụ: nếu bạn quan tâm đến thông tin về buôn bán trẻ em, hãy tự đặt cho mình các câu hỏi phù hợp, ví dụ:

- o Tôi muốn biết thông tin cụ thể về nạn buôn bán trẻ em ở nước nào, vùng nào?
- o Tôi cần những thông tin về các vấn đề nóng hổi hiện nay liên quan đến buôn bán trẻ em hay là thông tin mang tính lịch sử?

Biến yêu cầu của mình thành một câu hoàn chỉnh, ví dụ

- o Nạn buôn bán trẻ em từ Việt Nam sang Trung Quốc.

Phân chia yêu cầu thành những khái niệm nhỏ:

Nguyễn Tuấn Anh – 45K

July 28, 2007



Khái niệm 1	Khái niệm 2	Khái niệm 3
Buôn bán trẻ em	Việt Nam	Trung Quốc

Chú ý quan trọng: Tìm xem có những từ ngữ nào khác cùng thể hiện chủ đề mà bạn quan tâm hay không, ví dụ cách viết khác nhau, từ đồng nghĩa. Từ đó bạn có thể xây dựng được một tập hợp các thuật ngữ khác nhau có thể dùng trong quá trình tìm kiếm.

Khái niệm 1	Khái niệm 2	Khái niệm 3
Buôn bán trẻ em	Việt Nam	Trung Quốc
Bắt cóc trẻ em	Vietnam	China
Trafficking	Viet Nam	
Child	Vietnamese	
Children		
Kidnap		

3.2 Diễn đạt lệnh tìm kiếm - giới thiệu về cú pháp của lệnh tìm

Cú pháp của lệnh tìm là cách thức chúng ta sử dụng để liên kết các khái niệm một cách phù hợp cho lệnh tìm của bạn. Các máy tìm kiếm có những cách thức khác nhau trong việc liên kết các thuật ngữ tìm.

Nguyên tắc cơ bản nhất cho hầu hết các máy tìm kiếm là tương tự như nhau, tuy nhiên có một vài điểm khác biệt nhỏ về cách diễn đạt lệnh tìm. Nếu chưa hiểu rõ, bạn nên tham khảo phần “help” để biết thêm thông tin.

Chú ý:

- ✓ Phần lớn các máy tìm kiếm không phân biệt chữ hoa và chữ thường
- ✓ Khi nhập từ tìm kiếm vào cửa sổ tìm, cần đưa thuật ngữ mà bạn cho là quan trọng nhất lên đầu lệnh tìm.
- ✓ Không cần nhập cả một câu đầy đủ vào lệnh tìm kiếm, ví dụ như: “Hiện trạng của nạn buôn bán trẻ em tại từ Việt Nam sang Trung Quốc”. Thay vào đó bạn có thể nhập các từ “buôn bán trẻ em” “Việt Nam” “Trung Quốc”. Không giống như ngôn ngữ tự nhiên, các máy tìm kiếm không quan tâm đến sự chính xác về ngữ pháp của thuật ngữ tìm.
- ✓ Nhiều máy tìm kiếm thường bỏ qua các từ thông thường trong tiếng Anh như “the” “and” “in”, v.v...
- ✓ Nếu bạn nhập nhiều từ tìm kiếm thì phạm vi tìm kiếm sẽ được thu hẹp

3.2.1 Các phép toán của lệnh tìm

Phần này giới thiệu các ký hiệu toán học được sử dụng khi diễn đạt lệnh tìm kiếm:

Dùng dấu cộng + phía trước những từ mà bạn muốn nó **phải** xuất hiện trong kết quả. Ví dụ: nếu bạn muốn tìm thông tin về giá cả phê ảnh hưởng như thế nào đến thu nhập của người thu hái cà phê (tiếng Anh).

+coffee +pickers +price +wages

sẽ cho kết quả bao gồm **tất cả** các từ: *price, coffee, pickers wages*.

Dùng dấu trừ - phía trước các từ mà bạn muốn **không xuất hiện** trong kết quả. Ví dụ: bạn chỉ muốn tìm thông tin về cà phê mà không có từ tách (cup) hoặc bông (cotton):

+coffee -cup -cotton



Dùng dấu ngoặc kép " " đối với một tập hợp các từ mà bạn muốn xuất hiện trong kết quả chính xác như là một **cụm từ**. Ví dụ,

"coffee pickers"

"somali refugees"

"emperor penguins"

"nạn vi phạm bản quyền phần mềm"

"buôn bán trẻ em"

Chức năng tìm kiếm theo cụm từ kể trên rất hữu ích, nhất là khi bạn muốn tìm kiếm tên người (Võ Nguyên Giáp), tên cơ quan tổ chức (Bộ tài nguyên và môi trường), lời bài hát, một cụm từ hoặc câu nói nổi tiếng (to be or not to be).

Một số máy tìm kiếm cho phép bạn sử dụng các ký tự thay thế (ví dụ ký hiệu * thường dùng) để thay thế các ký tự. Ví dụ, nếu bạn muốn tìm các từ education, educators và educate, v.v..., bạn có thể nhập educ*

3.2.2 Toán tử Boolean

Một số máy tìm kiếm cho phép sử dụng các toán tử "**AND**", "**OR**" và "**NOT**" để diễn đạt những lệnh tìm phức tạp. Thường các toán tử này phải viết bằng chữ hoa.

Sử dụng **AND** nếu bạn muốn nhiều thuật ngữ phải xuất hiện trong kết quả tìm, ví dụ nếu bạn nhập như sau:

"lao động" AND "trẻ em"

thì kết quả sẽ có các tài liệu có cả 2 thuật ngữ **lao động** và **trẻ em**.

Sử dụng **OR** nếu bạn muốn một trong các thuật ngữ xuất hiện trong kết quả tìm (hoặc là cả 2 hay nhiều thuật ngữ).

Toán tử OR có thể sử dụng một cách hữu ích nếu bạn muốn tìm các từ đồng nghĩa, các cách viết khác nhau của một từ, ví dụ:

organization **OR** organisation

vietnam **OR** "viet nam"

quy chế **OR** qui chế

Sử dụng **NOT** để loại bỏ những tài liệu bao gồm những từ bạn không cần.

Ví dụ: nếu bạn đang tìm thông tin về "vàng" (kim loại quý) chứ không thông tin về "vàng" (màu sắc), bạn có thể diễn đạt như sau:

vàng **NOT** màu

3.2.3 Một số cách thu hẹp phạm vi tìm kiếm khác

- Giới hạn theo định dạng file (.pdf, .doc, .mp3, ...). Ví dụ: muốn tìm các tài liệu học tập về đồ họa bạn có thể tìm kiếm ở các dạng word, Adobe Reader

filetype: doc "learn photoshop" --> tìm tài liệu file word

filetype: pdf "learn photoshop" --> tìm tài liệu file Adobe Reader

filetype: pdf "graphic design" --> tìm tài liệu liên quan đến graphic design

Muốn tìm các file mp3, trong ô Search của google gõ vào đoạn mã sau:

+("index of") + ("/mp3"/"mp3s"/"music") +(mp3|zip|rar) +apache

Bằng cách biên tập mã này, bạn có thể tìm kiếm bất kỳ định dạng truyền thông(media) nào cả ood, wav, pdf, v.v... Cần bảo đảm phải biên tập cả tên thư mục cũng như phần mở rộng của file.



- Giới hạn theo ngôn ngữ (tiếng Anh, tiếng Pháp, ...)
- Giới hạn theo từng loại địa chỉ web (ví dụ chỉ tìm các tài liệu từ các trang web có đuôi .gov.vn, hoặc .edu)
- Giới hạn theo địa điểm xuất hiện của từ tìm kiếm (ở tên tài liệu hoặc trong nội dung)
- Tìm tranh ảnh hoặc bản đồ: Với Google, bạn hãy thử vào Google Labs (<http://labs.google.com>) để cụ thể hóa những gì cần tìm. Chẳng hạn khi gõ địa chỉ bắt đầu và kết thúc trong Google Maps (<http://maps.google.com>), bạn sẽ nhận được một bản đồ 2D cùng với đường đi và trong nhiều trường hợp là hình ảnh vệ tinh có thể phóng to về đoạn đường đó. Bạn có thể dùng hình vệ tinh này như tấm bản đồ thông thường. Tuy nhiên, Google Maps chỉ bó hẹp thông tin tại Mỹ và dịch vụ này vẫn còn trong giai đoạn thử nghiệm. Trong nước, bạn có thể tham khảo bản đồ Việt Nam tại www.basao.com.vn với bản vẽ 2D khá chi tiết.
- Đi tìm CD-Key đã mất: Một chiều đẹp trời phần mềm trong máy hết hạn xài chùa và yêu cầu bạn đứng trước hai con đường trả tiền, nạp Cd-key hoặc khôi xài nữa. Gỡ bỏ thì tiếc còn tiếc,... Ồ, không sao Google sẽ giúp bạn: 94FBR "ADOBE PHOTOSHOP "94FBR " Corel Draw 11". Nếu kết quả tìm kiếm không như ý bạn có thể thay cụm từ đầu bằng những cụm từ tiếp theo sau: GC6J3, GTQ62, FP876, D3DX8. Và thay thế bằng tên phần mềm cần tìm, bạn sẽ có tất cả trong tay.

Sau khi đã phân tích nhu cầu thông tin, bạn có thể tiếp tục quá trình tìm kiếm sử dụng bảng sau đây:

3.3 Phân nhóm yêu cầu thông tin		3.4 Chọn công cụ tìm kiếm phù hợp		
		❖ Chọn công cụ tìm kiếm phù hợp với thông tin mà bạn cần. ❖ Xem xét cách thức làm việc của từng công cụ tìm và điều chỉnh lại lệnh tìm để có thể khai thác tối đa các chức năng của công cụ tìm đó. ❖ Cố gắng thực hiện việc tìm kiếm trên nhiều công cụ tìm kiếm khác nhau. ❖ Xem các kết quả tìm và sử dụng các thuật ngữ được sử dụng trong các văn bản tìm được để tìm kiếm lại.		
Loại yêu cầu tìm	Ví dụ	Máy tìm kiếm (Google, AltaVista...)	Danh mục theo chủ đề	Web ẩn (invisible web)
Gồm các từ/cụm từ rõ ràng, dễ hiểu (gần như không thể bị hiểu nhầm sang nghĩa khác).	- o “buôn bán trẻ em” - o “đầu tư nước ngoài” - o “Robusta coffee” - o “Fair trade” +chocolate	Đề các từ, cụm từ trong dấu ngoặc kép “”, để máy tìm kiếm có thể tìm chính xác cụm từ theo trật tự đó.	Xem xét các khả năng xuất hiện của chủ đề bạn đang tìm trong các mục khác nhau vì một chủ đề có thể xuất hiện trong nhiều nhóm.	Nếu bạn tìm: - o Dữ liệu - o Số liệu thống kê - o Bản đồ v.v...
Gồm các từ phổ biến hoặc khái quát có thể có kết quả là quá nhiều kết quả không phù hợp	- o Weekend break - o Quản trị kinh doanh - o Du học	Tìm những thuật ngữ cụ thể hơn để phân ánh nhu cầu tin của mình.	Tìm các danh mục theo chủ đề để hiểu rõ hơn về các chủ đề nhỏ có trong nhu cầu tin của mình.	khi đó các cơ sở dữ liệu chuyên ngành có thể phù hợp với bất kỳ yêu cầu nào. Dùng một danh mục theo chủ đề hoặc một máy tìm kiếm để bắt



Tìm thông tin khái quát về một chủ đề.	- o Quyền sở hữu trí tuệ / Intellectual property rights - o Rain forest ecology - o Văn học Anh	TRONG TRƯỜNG HỢP NÀY MÁY TÌM KIẾM KHÔNG PHẢI LÀ CÔNG CỤ TÌM KIẾM TỐI ƯU.	Bạn sẽ có được kết quả tốt nhất nếu tìm được một danh mục theo chủ đề mà trong đó có liệt kê về chủ đề mà bạn quan tâm.	đầu tìm kiếm cơ sở dữ liệu phù hợp. Ví dụ: - muốn tìm cơ sở dữ liệu về dân số, nhập vào Google: population database - tìm số liệu thống kê về giáo dục: education statistics - tìm các cổng thông tin: education portal/gateway
Tìm kiếm thông tin theo một chủ đề hẹp.	- o Thống kê về nạn buôn bán trẻ em ở Việt Nam - o Migration patterns of Emperor penguins in the Antarctic - o Các dự án về giao thông nông thôn ở Việt Nam	Diễn đạt lệnh tìm bằng cách sử dụng các từ nối thu hẹp hoặc mở rộng phạm vi tìm như AND, OR, NOT.	Tìm một danh mục theo chủ đề và xem lướt theo các thư mục của nó để tìm chủ đề nhỏ mà bạn quan tâm.	
Có thể diễn đạt bằng nhiều cách khác nhau.	- o “Falkland Islands” OR Malvinas - o “New Zealand” OR Aotearoa - o Cassava OR manioc - o “Freedom of speech” OR “press freedom” OR “anti-censorship” OR “1st amendment” - o “nhà đất” OR “bất động sản”	Sử dụng toán tử OR để diễn đạt lệnh tìm.	Các danh mục theo chủ đề có thể sử dụng nhiều thuật ngữ khác nhau, tuy nhiên thuật ngữ thông dụng nhất thường được sử dụng.	

Bảng so sánh nhu cầu tin của bạn với các tính năng của máy tìm kiếm

Việc phân tích nhu cầu thông tin của bạn trước khi tiến hành tìm kiếm là để quyết định xem thuật ngữ nào sẽ được sử dụng và nên sử dụng tính năng nào của máy tìm kiếm cho phù hợp. Bảng sau đây liệt kê những đặc điểm của nhu cầu thông tin ở cột bên trái. Cột bên phải liệt kê những tính năng của máy tìm kiếm phù hợp với đặc điểm của từng loại nhu cầu thông tin của bạn.

Đặc điểm của nhu cầu thông tin	Các tính năng tương ứng/phù hợp của máy tìm kiếm
Bạn đang tìm một tên người, tên tài liệu cụ thể hoặc một thuật ngữ/cụm từ rõ ràng? - Tên của một tổ chức, hiệp hội hoặc phong trào - Tên của một cá nhân, một sự việc, sự vật - Một thuật ngữ/cụm từ xác định rõ ràng	TÌM THEO CỤM TỪ là tính năng bạn có thể sử dụng trong mọi công cụ tìm kiếm. Chọn tính năng yêu cầu kết quả phải có cụm từ theo đúng trật tự bạn đánh vào. Đặt cụm từ vào trong dấu ngoặc kép " " Ví dụ: "bất động sản" "world health organization" "Nguyễn Minh Châu"
Một hay vài thuật ngữ của bạn sẽ có nhiều nghĩa trong nhiều ngữ cảnh khác nhau? - Ảnh hưởng của bạo lực trên truyền hình đối với trẻ em (Children in conjunction with television and also violence) Sự kiểm duyệt như là một khía cạnh của đạo đức trong báo chí Censorship as an aspect of ethics in journalism	Toán tử AND hữu ích trong trường hợp này: “trẻ em” AND “truyền hình” AND “bạo lực” (children AND television AND violence) “báo chí” AND “đạo đức” AND “kiểm duyệt” (journalism AND ethics AND censorship) Các máy tìm kiếm Google và AllTheWeb và phần lớn các công cụ tìm kiếm khác đều sử dụng toán tử AND một cách ngầm định: children television violence journalism ethics censorship
Bạn đoán rằng trong danh sách kết quả sẽ có rất nhiều trang web có những từ mà bạn không muốn? Ví dụ: bạn tìm kiếm tài liệu về “biomedical engineering” và cancer. Tuy nhiên trong trang kết quả có rất nhiều trang web	Toán tử AND NOT hoặc NOT hữu ích trong trường hợp này: "biomedical engineering" AND cancer AND NOT "Department of" AND NOT "School of" hoặc dùng dấu trừ - ở phía trước từ muốn loại bỏ:

Nguyễn Tuấn Anh – 45K

