

Google the Search Engine LIS 598

what is Google?

what is a search engine?

searching for information:
the last 50 years

ask someone (the preferred method)

manual

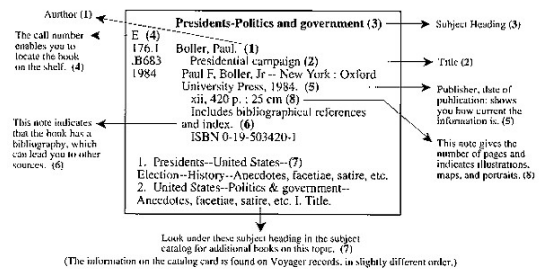
digital

Web

card catalog



catalog cards



index (bibliographic)



index (bibliographic)

Example: from *Psychological Abstracts* subject index.

Subject → **Violence [See Also Family Violence, patient Violence]**
drug use & violence & victimization, male vs female 6th-12th grade Mexican American vs White vs American Indian Dropouts vs students with academic problems vs in good academic standing, 32335.

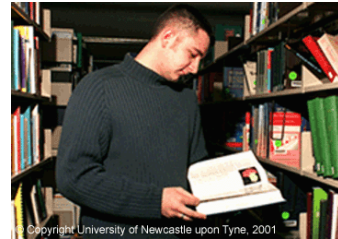
Example: from *Psychological Abstracts* abstract listing.

Author → 32335
Journal Title → Beavrais, Frederick et al "Drug use, violence and victimization among White American, Mexican American and American Indian Dropouts, Students with academic problems, and students in good academic standing."
Pages → Journal of Counseling Psychology, 1996 (Jul), Vol. 43(3), 292-299
Abstract → Problem-prone behaviors of White American, Mexican American, and American Indian high school dropouts, students in good academic standing, and students in poor academic standing were surveyed. Generally, dropouts...
Abstract Number → 32335
Article Title → Beavrais, Frederick et al "Drug use, violence and victimization among White American, Mexican American and American Indian Dropouts, Students with academic problems, and students in good academic standing."
Volume & Issue → Journal of Counseling Psychology, 1996 (Jul), Vol. 43(3)

within a resource

indexes, tables of contents, directories, etc

browsing



manual searching

by metadata

title, author, subject indicator (incl words, names)

possibly controlled, possibly not

within structure

by 1st character

incl browsing to correct, reinforce spelling

digital

online catalogs

online databases

CD-ROM databases

digital searching

by metadata

all previous, plus full text where available

free text/keyword searching

still within structure

expanded notion of what that structure is

document representation may incorporate document itself, plus metadata

loss of browsing

or at least different notion

Internet searching

the very very early days: nothing; word of mouth

the very early days: Archie (FTP) and Veronica (Gopher)

the early days: Web searching

simple keyword searching

no Boolean searching

no phrase searching

no relevance ranking

no truncation, wildcards

and directories

Web searching matures

adds those important features, esp ranking

based on term frequency, placement (TITLE, META, Hx, high in document)

meta-search engines

subject specific search

lots of entries: AltaVista, HotBot, Lycos, Inktomi, Magellan, Excite, InfoSeek, LookSmart, Northern Light, WebCrawler, GoTo, etc etc etc

and then...Google

slowly and surely has become the predominant search engine

why?

technology & performance?

business plan?

cultural impact?

ease of search?

let's find out....

Google searching stuff

search strategy

syntax

auxiliary stuff

features

many of the following from *Google Hacks*(O'Reilly 2003)

search strategy

OR does work (capitalized), also pipe |

* as wild card for word in phrase search

word order

repetition

syntax

file type
intitle:
inurl: site:
date range

auxiliary stuff

cached results
advertising
look @ results metadata
specialized vocabulary

features

spelling, dictionary
phonebook
articles
directories
search by number (UPS tracking, patent)
travel conditions
translation
calculators
see also the Labs

so what is Google?

search engine?
something more?
what else could it be??