

SEARCH STRATEGIES

“SEARCHING THE INTERNET IS LIKE SHOOTING INTO A CROWDED ROOM. YOU MAY BRING DOWN YOUR TARGET, BUT YOU ALSO GET A LOT OF INNOCENT BYSTANDERS”

Precision vs. Recall - skip this if you want!

Searching for online information presents us with two mutually exclusive goals: we can “shoot” for precision or “shoot” for recall. What does this mean?

Precision is the ratio of relevant items retrieved to all items retrieved. For example, if you retrieve 10 items in a search and nine of them are relevant you have high precision.

Recall is the ratio of relevant items retrieved to all the relevant items that exist; the goal is a comprehensive search. For example if you retrieve 100 items in a search and 50 of them are relevant and there are no other relevant items in the database you have high recall.

YOU ALSO HAVE A LOT OF GARBAGE!

MOST OF THE TIME, MOST PEOPLE WANT A HIGH PRECISION SEARCH . IN OTHER WORDS, YOU WANT A FEW HIGHLY RELEVANT ITEMS.

HINT: find a search engine you like (such as Altavista): stick with it and learn it well. This is, for most people, more effective than trying to learn the nuances of many different search engines.

The following chart lists some strategies for increasing your precision by narrowing your search. The search features pertain to an *altavista standard search*, but some of the other search engines work this way too:

SEARCH KEYS	HOW IT SEARCHES	RESULTS
journalism careers	<i>either or both</i> words anywhere they appear	best recall – worst precision
“journalism careers”	phrase searched in text of web pages	narrower recall – better precision
title: “journalism careers”	title of web pages searched	still narrower recall – even better precision

OVER OVER OVER OVER OVER OVER OVER OVER OVER

The following searches assume you are using Altavista's Advanced search function - Altavista's basic search does not do "and" searches nor does it do "near" searches.

SEARCH KEYS	HOW IT SEARCHES	EFFECTS
journalism and careers	boolean "and" search – both terms <u>must</u> present	"and" narrows - increases precision
journalism or careers	boolean "or" search - either one or the other or both may be present	"or" broadens - increases recall
journalis*	truncation - allows for plurals, adverbial and adjectival forms such as journalism, journalist, journalistic etc.	broadens search - increases recall
journalism near careers	proximity search - requires search terms appear within ten terms of each other	narrows search - increases precision

OVER OVER OVER OVER OVER OVER OVER OVER OVER