

Using Search Engine Relevancy Ranking to Teach Internet Evaluation

JOSEPH YUE

Reference/Instruction Librarian
UNIVERSITY OF COLORADO AT BOULDER
joseph.yue@colorado.edu

Circumstances of the Assignment:

Many search engines now provide search result rankings based on different criteria or computerized evaluation processes. How ranking really works in each search engine remains a mystery. While one may be able to find marketing literature on how well search engine results answer queries, the underlying mechanism or algorithm is proprietary information. Thus, it is important for users to understand the importance of evaluating sources based on their information needs rather than external structure.

Search engines that organize results in a logical or systematic way based on a programmed relevancy evaluation can be deceptive. Users need to be aware that what a search engine considers to be highly relevant to a query does not necessarily equate to users' expectations. This is different from users' experience in searching library catalogs or bibliographic databases. Instead of trying to force a value judgement on the results by displaying results according to computerized criteria, search results from library catalogs and bibliographic databases are usually organized alphabetically or chronologically. Users have to decide whether the search results are relevant to their topic.

This assignment tries to teach students that they should not rely on search engine relevancy rankings, but instead use their own evaluation criteria to critically scan the search results. Through looking at search results critically, student may also learn from this assignment that they do not only need to evaluate the top few results presented. They need to go a step further and review the first two to three screens to locate relevant Websites that are considered low relevancy by inflexible computer evaluation criteria.

This assignment is most effective as a take-home assignment, with discussion of the results occurring at the next class meeting. To facilitate critical thinking, students can work in groups and discuss their search results and findings. However, if the assignment is used as an independent exercise, detailed written analysis can substitute for the class discussion.

The preliminary exercise can be 10-15 minutes long, depending on class setup. The actual search exercise may take 15-20 minutes for each search engine. Post-exercise time will depend on how in-depth the discussion is.

Objectives of the Assignment:

At the conclusion of this assignment, successful students should:

- have a deeper understanding of the use of search engines,
- recognize the difference between computer ranking and human evaluation,
- cultivate critical thinking and skepticism to information presented as logical and systematic, and
- integrate evaluation skills into their day-to-day Internet searches.

Components of the Assignment:

This assignment consists of four parts: Parts 1 and 2 are optional and should be conducted in class if they are used. Part 3 can be used as a take-home assignment and Part 4 is a discussion activity to be used after the search/evaluation exercise is completed. The entire assignment can also be completed during a 1¹/₂ to 2-hour long session.

Part 1: Evaluation Criteria

The librarian may choose to make some comments or conduct some warm-up exercises on evaluation. A pre-defined search topic is presented to students. The librarian will prepare a list of basic evaluation criteria in short question format. Alternatively, the librarian can lead the discussion to generate a list of evaluation criteria based on a predefined topic.

Part 2: Ranking of Results

A list of Website addresses or printouts of Websites on the same pre-defined topic in Part 1 are provided to students. Students, individually, in groups, or in the class as a whole, will rank the list of Websites that best meet each evaluation criteria. This can be accomplished using a point system, assigning the worst Website a point of one and the best one the total number of Websites used in this exercise. In other words, if five Websites are used, the best one will be assigned five points. The emphasis is on how they come to the decision and the rationale behind it. This exercise can be repeated as many times as desired with different criteria until the librarian feels students have an understanding of how evaluation can be applied to Websites. The librarian concludes this warm-up exercise by adding all the points each Website received and assigning rank order. The Website with the highest score will be assigned the rank order of one.

Part 3: In-Class or Take Home Search/Evaluation Exercise

Individual students or groups of students are then asked to conduct a search on the topic using the same keywords in one search engine that provides relevancy ranking. They need to make note of the ranking and reassign final rank orders according to the evaluation criteria provided in class. They are required to write down the rationale behind such rankings.

Variation 1 for Part 3

In a large classroom environment, students can be divided into groups, with each group assigned a different topic or search engine. This will help maintain a comparable framework for later discussion. It is important that students in all the groups use the same search words.

Variation 2 for Part 3

To make the assignment more challenging for advanced students—those who like to find out how things work or what lies behind the scenes, and who have a good understanding of Web page creations—encourage them to find out how search engines execute their evaluation process. The goal in this case is to demonstrate how computer predefined evaluation is rigid and context-independent and is no match for human, context-based evaluation of Websites.

Students are asked to look at what information is included in the META and TITLE tag in the first five to ten Websites in addition to the search/evaluation exercise outlined above. Their task is to identify how this kind of information, which normal users do not pay attention to, will affect the search engine results' relevancy ranking. Alternatively, students can be asked to review the help file online to find out how ranking is accomplished by specific search engines (context, word frequency, back links and others as new technologies are developed).

Note

For a shorter format in a small classroom setting, Parts 1-3 can be combined in the following manner:

A sample search is conducted before class with a pre-selected set of search engines to identify Websites that can illustrate certain aspects of evaluation criteria. In class, the same search is conducted and projected on a screen. The relevancy rankings received from each search engine are then analyzed. The librarian guides students to a specific evaluation criterion through focused questions or small-group discussions. The ranked order of Websites are then gathered and analyzed briefly.

Part 4: Group or Class Discussion

If the search exercise is conducted in groups in class, students within the group will need to reach a consensus on why they ranked the results in a particular order. The class can then discuss any differences and similarities of their search results and analysis. The librarian can provide guidance in highlighting the effectiveness of each search engine and providing a framework for future Website evaluation.

General conclusions the librarian may want to cover in Part 4:

- Predefined criteria of search engines are context-independent to user needs, as opposed to human evaluation.
- Ineffective result rankings may hide relevant information in later screen displays, so students need to evaluate the results actively and go beyond the first 10-20 search engine "top" results.
- Websites can mislead computers through various means as they compute relevancy rankings. Plausible examples, including word occurrences in title or meta tag, mean that authors of a Website can include a term multiple times to increase its ranking, popularity through hit counts, backlinks (how many Websites link to a particular site), etc. Human beings are harder to fool.

Evaluation:

Success of the assignment is demonstrated through students' analysis of their search results, differences in their own rank order compared with that of search engines, and students' articulation of the limited relevancy ranking capabilities of search engines.

Supplementary Materials:

1. A sample worksheet for the activity is included. If search engines are used that do not display relevancy percentages, students can assume that the order of the results represents the search engine's ranking.
2. Partial listing of search engines that provide relevancy ranking:

Percentile of relevancy listed in the result screen:

<i>Excite</i>	www.excite.com
<i>HotBot</i>	www.hotbot.com
<i>Infoseek</i>	www.infoseek.com
<i>Magellan</i>	magellan.excite.com
<i>Direct Hit</i>	www.directhit.com
<i>Inference Find</i>	www.infind.com

Search Engine Relevancy Ranking and Internet Evaluation Worksheet

Name: _____ Class: _____ Date: _____

Searching

Write down your assigned topic, search engine, and keywords.

Topic:

Search Engine:

Keywords:

Evaluation

From the results screen, please choose five Websites that are highly relevant to your topic. For each Website, answer the questions listed.

Website 1

Website name:

Website address:

Search engine indicates that it is ____ % relevant to your search.

A. Does the Website provide the information you are looking for?

B. How current is the information?

C. Is the information presented objectively?

D. Is the information based on facts or on personal opinion?

Website 2

Website name:

Website address:

Search engine indicates that it is ____ % relevant to your search.

A. Does the Website provide the information you are looking for?

B. How current is the information?

C. Is the information presented objectively?

D. Is the information based on facts or on personal opinion?

Website 3

Website name:

Website address:

Search engine indicates that it is ____ % relevant to your search.

A. Does the Website provide the information you are looking for?

B. How current is the information?

C. Is the information presented objectively?

D. Is the information based on facts or on personal opinion?

Website 4

Website name:

Website address:

Search engine indicates that it is ____ % relevant to your search.

A. Does the Website provide the information you are looking for?

B. How current is the information?

C. Is the information presented objectively?

D. Is the information based on facts or on personal opinion?

Website 5

Website name:

Website address:

Search engine indicates that it is ____ % relevant to your search.

A. Does the Website provide the information you are looking for?

B. How current is the information?

C. Is the information presented objectively?

D. Is the information based on facts or on personal opinion?

Ranking

- i. Compare your answers to Question A for all five Websites. Give the best Website 5 points and the worst 1 point. For example, if based on Question A, Website 1 is the second best site, write "4" in the box for Website 1 under column A.
- ii. Repeat the same process outlined above for Question B-D.
- iii. Add up all the points for each Website across columns A-D and enter the total into "Total Points" column.
- iv. Rank the Websites from 1 through 5 according to the total points, with the highest scored Website receiving the rank of 1.

	Question				Total Points	Your Ranking	% of Relevancy according to Search Engine	Search Engine's Ranking
	A	B	C	D				
Website 1								
Website 2								
Website 3								
Website 4								
Website 5								

- v. For each Website, under the column "% of Relevancy according to Search Engine," enter the percentage you wrote down in the first part of this worksheet.
- vi. In the column "Search Engine Rank Order," rank the Website from 1 through 5 according to search engine relevancy percentage, with highest percentage receiving the rank of 1.
- vii. Compare your ranking with the search engine's ranking. Are they the same? Why or why not?