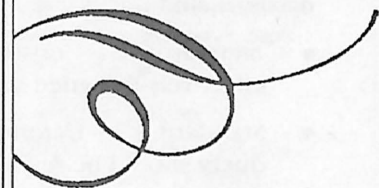

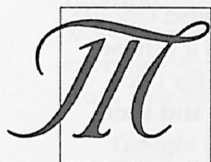


GOOGLEWHACKING:
EXPLOITING GOOGLE
IN AN INSTRUCTION CLASSROOM

by Meg Atwater-Singer



googlewhack! 
The Search for "The One"



Most students already use Google, so why not exploit its teaching and learning potential in a library instruction classroom? University of Evansville Libraries (UEL) introduces concepts such as keyword selection, search construction and techniques, invisible/visible web, Google as a business and more by having students use Google. This article will describe how to incorporate a critical review of Google by employing a *Googlewhacking* exercise into an instruction session and generating questions to get students thinking critically about information.

WHAT IS GOOGLEWHACKING?

The Googlewhack website was created by Gary Stock in 2002 [<http://www.googlewhack.com>]. The site allows anyone to record a "whack," which is a search of any two words in Google's search engine which yields only one result. A few rules pertain to this simple premise: quotation marks are not allowed; words must be between four and thirty characters long; and the search terms must be present in the dictionary Google uses. When one does a search in Google, the search terms are displayed in the results bar. If the terms are clickable, the words are found in the dictionary Google uses. This rule discourages whacks with creative spelling or spelling errors.

Another rule that one must be aware of is, no word lists. When trying to find a whack, it can be frustrating to have Google return one result for your search terms only to discover that the page is a list of words. Googlewhack does not allow one to record a whack found on a word list. No whack for you!

WHY DOES UEL USE GOOGLEWHACK?

The University of Evansville is an independent, liberal arts and sciences university, with selected

professional programs in the southwestern corner of Indiana with an enrollment of approximately 2,500 FTE. All first-year students must enroll in a two-semester sequence called *World Cultures*. In this course, students read literature from all time periods and cultures, discuss the themes in class and hone their writing skills. In the fall semester, most *World Cultures* sections come to the Libraries. During an instruction session with a librarian, students learn how information is produced and organized using Pennsylvania State University Libraries' *Information Cycle* video (2001). Afterwards, students familiarize themselves with UEL's website, which is their main conduit for information delivery, matching specific information types demonstrated in the video with finding tools available on the website.

During the second semester, *World Cultures* students are required to write a research paper. This is often their first college paper, a potentially daunting experience! To help students discover the ease with which information retrieval skills are transferable, the Instruction Department decided to incorporate a Googlewhack exercise into their instruction session. This exercise demonstrates transferability of research skills using a tool that they are comfortable with (Google) and then introduces new tools such as online databases that UEL purchases on their behalf. This exercise also builds upon the students' fall semester library experience, reinforcing what they learned about information generation and the different types of information available.

WHAT IL STANDARDS ARE PRACTICED?

The Association of College and Research Libraries has formulated *Information Literacy Competency Standards for Higher Education*, which serve as a

useful guide in structuring outcomes for instruction sessions (2000). Through the use of the *Googlewhack* exercise, students gain competency in several standards, including:

- Standard 2.2 – Constructing and implementing effectively-designed search strategies (p. 9);
- Standard 3.7 – Determining whether the initial query should be revised (p. 12); and,
- Standard 5.1 – Understanding many of the ethical, legal and socio-economic issues surrounding information and information technology (p. 14).

While trying to find a *Googlewhack* is an artificial assignment, students constantly revise their search terms, using synonyms and related words, in order to locate a whack. This exercise mirrors real-world searching for relevant information that students would cite in their research papers. Trying and failing to find information and then ultimately succeeding is an important hallmark of the information-literate individual. Additionally, learning that there is more information accessible via the Internet than can be found through commercial search engines like Google, and the differences between free and fee-based information sources help students understand the many layers of information access that exist.

THE GOOGLEWHACK EXERCISE

After students have come to the instruction session in the computer lab, introductions are made and goals for the class are outlined. The goals for the session are to determine how the search engine or database searches, how results are displayed, and what types of information are retrieved.

Students are asked to open two browser windows: Google in one and *Googlewhack* in the other. After explaining what a *whack* is, the rules that set parameters for finding a *whack*, and parts of the *Googlewhack* website (especially the *Whack Stack*,

which can be mined for search terms, see Figure 1), students are instructed to find a *whack* in Google. Students are given five minutes to try and find a *whack*. The librarian can wander around the room offering advice on term selection and commiserating when students' find a single result but one of the words is not in the Google dictionary or the *whack* result is a word list. When the allotted time has expired, the librarian leads a discussion of Google using the exercise's goals as a guide, which follow below.

HOW DOES GOOGLE SEARCH?

Google automatically ANDs the search terms together, looking for webpages that include both terms. Additionally, Google searches for keyword variations. For example, when searching for *ballerina antarctica*, Google will also return results for *Antarctic* and *ballerina*.

One of the reasons why quotation marks are not allowed when searching for a whack is because Google interprets search terms inside quotations as a phrase. Students can test this concept by searching for two words within and without quotation marks and then discussing how the results differ.

HOW DOES GOOGLE DISPLAY RESULTS?

Google displays results according to how "relevant" a webpage is to the search terms. Relevancy is determined by several factors including the frequency of the search terms on a page, the proximity of the terms to one another, and the placement of terms in titles, headings, and subheadings (Cutts, 2005, ¶ 11).

Equally important is the popularity a page enjoys within the Google database. As others search Google for information and click on webpages, a page's popularity increases with each visit. With over 300 million Google searches per day, a page's popularity could be positively or negatively impacted by all that traffic (Arnold, 2005, p. 21).

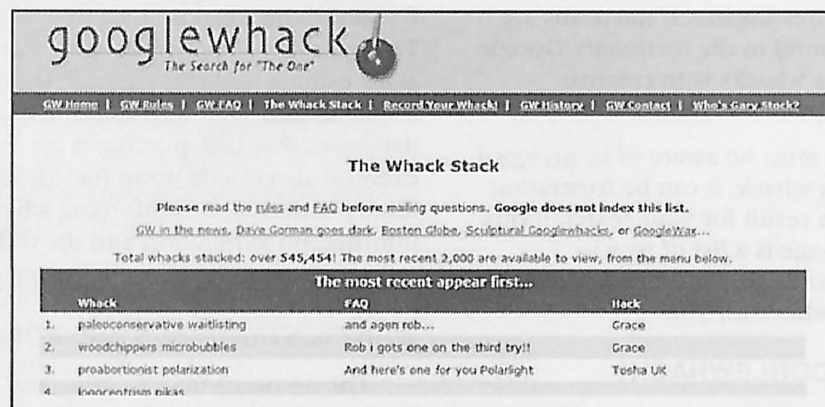


Figure 1

PageRank also plays a significant role in Google results. This is an algorithm developed by Google to assign a ranking to webpages within its database. According to the *Google's Newsletter for Librarians*, "PageRank evaluates two things: how many links there are to a webpage from other pages, and the quality of the linking sites" (Cutts, 2005, ¶ 10). Matt Cutts also writes that links from a "trusted" website, like CNN or the *New York Times* websites, are more highly valued and thus rank higher on page results than links from less reputable websites, like *The Onion* or *Schoolsucks.com* (¶ 10).

Unfortunately, this model can be manipulated by search engine optimization (SEO) and less sophisticated techniques like *Google bombing*. Stephen Arnold describes SEO as a "discipline of crafting publicly accessible web content in order to boost a website's ranking in Google" (2005, p. 19). Hiding text on a webpage, using *link farms* to list your website in multiple places, and incorporating lots of popular keywords into a page's metatags even when there is no correlation with the page's content are a few of the ways that SEO occurs (Arnold, p. 19).

Google bombs are another way to manipulate search results. By having many sites link to a single page using the same words as the hypertext link can result in a bomb. Wikipedia explains Google bombs (or Google washers) and lists many examples ("Google bombs," n.d.). To illustrate this concept, ask students to run a search for "miserable failure" without quotation marks. At the time of this writing, President George W. Bush's official White House biography is the first result. According to Google's PageRank and other criteria for organizing results discussed earlier, this site should be the most relevant. Have students analyze why Bush's biography is listed first and then explain how Google bombs work.

Critical examinations of Google search results are imperative for understanding how information is organized and retrieved. By introducing ideas of relevance, proximity and ranking, students can transfer learned skills to other information resources.

WHAT DOES GOOGLE SEARCH?

When asked to articulate exactly what Google searches, students will usually say, "Everything on the Internet!" Librarians know that this response is false. Google and its competitors can only index information that is visible to their web-crawling robots. Websites whose information is contained in databases that are dynamically created in response to a query cannot be seen by Google. Library catalogs, freely available on the internet, also fall into this category. Password-protected information cannot be seen by search engines either.

Social communities like *MySpace* and *Facebook* are examples of websites accessible by password and thus invisible to Google. Finally, some information is only available for a fee. Libraries buy access to information such as full text journals, e-books, and databases, and let their patrons use these resources for free.

Catalogs, government information, password-protected pages and database-driven websites are just the beginning of the invisible or deep web. In 2001, Michael Bergman of BrightPlanet surveyed the deep web and reported it to be four hundred to five hundred and fifty times larger than the commonly defined World Wide Web or visible web (p. 1). For example, Google announced in 2005 that its database contained over 8.2 billion webpages and 2.1 billion images (Kay, 2005, p. 28). The 500 billion plus English documents in the deep web dwarfs Google's database and illustrates the vastness of the deep web (Kay, p. 28).

Additionally, Google is actually a database of crawled, indexed websites. What this means to the searcher is that results are returned for any query in fractions of a second. If Google had to run its search against the "live" web, results would display much slower and the site would probably not be as popular.

BEYOND GOOGLEWHACK

UEL has successfully used the Googlewhack exercise for the past two years. Searching Google is second-nature to most students, so asking them to tweak their normal search habits to find a whack is not a stretch. Analyzing Google's database, the results it returns and how it searches makes for a lively discussion where students learn key skills that are easily transferable.

After wrapping up the discussion of Google, UE librarians request that students answer the same three questions after searching UE's online catalog and an article database. Again, these questions are, how does the database search, how are results displayed, and what type of information is searched. Comparing and contrasting the different interfaces and search mechanics enable students to learn important information literacy skills that will serve them well during their years at UE and beyond.

REFERENCES

- Association of College and Research Libraries. (2000). *Information literacy competency standards for higher education*. Retrieved June 1, 2006 from <http://www.ala.org/ala/acrl/acrlstandards/standards.pdf>
- Arnold, S. (2005). Relevance and the end of objective hits. *Online*, 29 (5): 16-21. Retrieved March 21, 2006, from Academic Search Premier.

Bergman, M.K. (2001). *The deep web: Surfacing hidden value*. Retrieved June 1, 2006, from <http://www.brightplanet.com/pdf/deepwebwhitepaper.pdf>

Cutts, M. (2005). How does Google collect and rank results? *Google's Newsletter for Librarians*, 1. Retrieved May 31, 2006, from http://www.google.com/librariancenter/articles/0512_01.html

Google bomb. (n.d.). Retrieved June 1, 2006, from http://en.wikipedia.org/wiki/Google_Bomb

Kay, R. (2005). Deep web. *Computerworld*, 39 (51): 28. Retrieved June 1, 2006, from Academic Search Premier.

Pennsylvania State University Libraries. (2001) Information cycle [online video]. Retrieved April 28, 2006, from <http://www.libraries.psu.edu/instruction/infocycle/infocycle.html>

ABOUT THE AUTHOR

Meg Atwater-Singer (ma35@evansville.edu) is the Reference/Instruction Librarian at the University of Evansville. This article is based on a presentation delivered at the 2006 Indiana Library Federation's annual conference.

