## How can appropriate search terms and queries guide targeted searches?

◆　　◆　　◆

Advanced Lesson 1:

## What is the difference between firm and soft search terms? Why I do need to utilize the deep web? How do I find deep web resources? How can context terms help me target my search for what I need?

**LESSON OVERVIEW:** This lesson contains a series of demonstrations where teachers present information and use the Internet to augment what is presented. Students then practice what they learn on their own or in partnerships. The content asks students to build on previous skills at selecting key words with an eye towards building more sophisticated and scholarly searches. Within this exercise, students learn to distinguish between "firm" and "soft" terms, and practice generalizing with context terms in order to locate databases and other special-subject collections. It culminates with students identifying whole collections of information on their topics.

**STANDARDS:**

- *K-12 College and Career Readiness (CCR) Anchor Standards for Reading 4:* Interpret **words and phrases as they are used in a text,** including determining **technical, connotative**, and figurative meanings, and analyze how specific word choices shape meaning or tone.
- *K-12 College and Career Readiness (CCR) Anchor Standards for Writing 7:* Conduct short as well as more sustained **research projects** based on focused questions, demonstrating understanding of the subject under investigation.
- *K-12 College and Career Readiness (CCR) Anchor Standards for Writing 8:* **Gather relevant information** from multiple print and **digital sources**, assess the credibility and accuracy of each source, and integrate the information while avoiding plagiarism.
- *K-12 College and Career Readiness (CCR) Anchor Standards for Language 6:* Acquire and use accurately a range of general **academic and domain-specific words and phrases** sufficient for reading, writing, speaking, and listening at the college and career readiness level; demonstrate independence in gathering vocabulary knowledge when encountering an unknown term important to comprehension or expression.

**RESOURCES/MATERIALS:**

- Advanced 'Picking the right search terms' presentation (individual slides linked to proper locations below)
- Internet access (strongly recommended)

**ESTIMATED TIMING:**

- approximately one 50-minute lesson

**NOTE TO TEACHER:** Throughout this lesson, there are opportunities to click links for search results if you have live Internet access. These links are expressed within most query brackets (e.g., [constitution]). If you do not have Internet access in your classroom, click the links for slides (e.g., Slide 7) prior to class and print out screen shots to show on a document camera to your students, on an overhead, or through hard copies. In some lesson points, there are slides only as an instructional tool. Also please note that the actual square brackets around each query should not be typed into Google. They represent what words and phrases should be actually typed into the search

box.

**LESSON DETAILS:**

1. **Set the stage.** Explain that as the class moves into the Advanced level lessons, an early goal is to develop a more sophisticated sense of selecting search terms. Depending on students' past exposure to Matt Cutts' 3 minute "How Search Works" video, consider playing it again to remind students how Google handles queries and ranks results. Even if they saw the video in earlier years, it is a time-efficient refresher.

2. **Explore "firm" and "soft" terms.**
   - Introduce to students the notion that often, in addition to the main topic of a search task, a query needs to include some language to positively narrow the results, as explored in Intermediate Lesson #4. For example, *"If you were trying to learn how to play the game Cricket, you might get results that were too broad if you just searched for [cricket] (Slide 3). You might chose to focus your search by adding the kinds of things you need to learn about cricket, such as [cricket rules] (Slide 4) or [cricket batter strategies] (Slide 5)."*

   - This example searching for [cricket] is fairly straightforward, but sometimes it is difficult to know what terms will clearly define what is needed so that a search tool recalls the correct sources for to fill that need. Share the following question, that came over a social network (Slide 6):
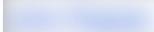
> ▢ - Feb 2, 2012 - Public (locked)
>
> Someone needs to help me find this book. It contained Native Americans as the primary protagonists, included witchcraft, gun-running and I believe a supporting cast of hippies. Big myth sub-text.
>
> Please help.
>
> [ +1 ]

Shared with permission

Challenge students to suggest methods for deciding which terms they would use in their query to identify the book in question. If students have Internet access, allow them a few minutes to try out queries before discussing. They can keep and eye out for a probably answer, but don't need a firm response, since they have no way of knowing when they have found the correct book (Slide 7).

Have students share potential terms, or whole queries, and explain why they would chose one word or another. Record choices and reasoning on the whiteboard.

   - Tell students that the class will look at searchers' responses in a minute. They will see the searcher describing assessing "firm" and "soft" search terms.
     - **Firm** search terms: (1) Are likely to be on most pages that discuss your topic, and/or (2) are described in a limited number of ways; and
     - **Soft** search terms: (2) Are less likely to appear consistently on pages about your topic, and/or (2) can be expressed in so many different ways that it is too difficult to come up with a good term or set of terms that are likely to find what you need.

Given the need to find something that is highly undefined for the searcher, she had to choose what words were most likely to be associated with her answer, for example:

   - [Native Americans] is risky as a search term, because there are so many different ways to express that concept, depending on the author's background, social context, and the time during which it was

written. Possible related terms include [Indian], [Native Peoples], [First Nations], [tribe], and even a specific tribe name, like [Navajo], which—besides being undefined by the asker, could have one of many spellings. So, [Native Americans] is here a soft search term, as likely derived from the asker's context as from the words of the book's author.

- Ask students what other terms the asker used might make good search terms. Try and categorize them as firm or soft.

- Share the searcher's explanation of the search:

> _____ - At a total guess, could it be **The Almanac of the Dead** by Leslie Marmon Silko? Totally just ran a Google Book search, but the description sounds like a possible match.

Shared with permission

> _____ - I just really love search, books, and Google Book Search. Query, actually in Google Web Search [witchcraft gun-runners novel] because I did not know if [hippies] and [native americans] would be described as such, but I know Google looks for related terms for [witchcraft] (i.e., [witch], and someone would probably be talking about it being a novel (which was actually a risk, but it seems to have worked). Sorry, can't help but talk about...well, search.
>
> Feb 2, 2012   **+4**

- Quickly discuss the searcher's assessment of firm and soft terms:
  - Firm: [gun-runner] and [witchcraft]
  - Soft: [hippies] (Note that the asker said he "believe[d there were] a supporting cast of hippies," which indicated to the searcher that [hippies] might be a value judgment of the reader, rather than the words of the author.)

- Note that the designation of firm and soft terms is to some extent contextual. [Native American] is always a hard term to search for, but [hippie] might be firmer if, for example, a student was looking for certain social analyses of the 1960s.

3. **Challenge with firm and soft search terms.**
   - Offer students the following search challenge. Warn them that it is going to require them to use a number of tactics that they have learned over the years to find the answer:

     *The highly sensational story of the Navy's Ghost Blimp is all over the web. But is the story just based upon a rumor? Can you find anywhere on the Navy website that talks about the story as if it really took place?*

   - Give students about ten minutes to work on this challenge. Request that they record their steps in a document or on a piece of paper. One possible process for solving this challenge is shown in Slide 8-16.

4. **Delving into the deep web.**
   - Acknowledge that it was no small feat to dig into the Navy.com website to find the information for the last challenge. Ask if any students tried using the site search on the Navy.com website. Some students may note that *they did, but got no results*. It appears that the part of the site containing the reporting on the Ghost Blimp is not covered by the site search.

- Sometimes, however, websites have content that Google and other search engines do not cover. There are several reason this may happen:

  - **Pages on the site** (or the whole site) include a special file (called a robots.txt) that asks search engines not to record that portion of the web. These are pages that the creator does not want others to be able to find with a search engine.
  - **Pages are password protected**, perhaps to protect privacy (e.g., a school website that posts information about or work by students), or because they are paid content (e.g., a subscription-based magazine). Note that some subscription services still allow Google to craw their content, but when it appears in search results users still need to pay to access the page.
  - **Pages are dynamically created**. One example of a dynamically created page is the results to a search done on Amazon.com, in the school's library catalog, or even on Google.com. Those pages are not stored and made searchable by search engines. Another example is some types of database, which have a static front page to enter and search the database, which is findable in Google, but the statistics which users access when visiting the database are pulled up when a search is run—the tables a visitor sees on the site do not exist until the moment his search creates them. Therefore, they cannot be found and indexed by Google.

- An example of this type of database is the [Bureau of Labor Statistics database](#). Although students can search [[Bureau of Labor Statistics database](#)] and find the entry-point to access statistics, they cannot use Google itself to search the contents of the database. Those searches must be run from within the BLS site.

- Thus, the **visible web** is that content on the web which search tools, including Google, can find, index, and offer to users as the result of a search. The **deep web**, also called the **dark web**, **invisible web**, or several other related names, is that part of the web which cannot be found directly by searching popular search tools. It includes many digital collections of primary sources, statistics, and scholarly works that students may want to access for research purposes, including any article databases to which the school library subscribes.

- However, much of the deep web is findable for students who thoughtfully engage context terms and take a strategy of generalization. In other words, students may not be able to use Google to search inside some of these deep and broad collections of information, but they can use Google to discover what collections exist, and how to enter them in order to search for what they need.

5. **Practice drawing from the deep web.**
   - For articles that appear in paid databases (such as JSTOR, CQ Researcher, Opposing Viewpoints), check with the school librarian and also Advanced Lesson 3, which covers Google Scholar.

   - You might tell students: *"When you run a search in Google, you are typically retrieving a single page that has the terms and other attributes you requested in your query. Sometimes, if you are looking carefully, you may see that you have stumbled upon a larger collection that is relevant to your research. Another tactic is to actually go looking for a collection of information on your topic, and then work within that collection to find what it has to offer."*

   - Review the notion of context terms, or words that would appear in the name or description of a source (i.e., actually appear written on the page) that describe what kind of source it is. As a class, brainstorm

context terms that students have encountered that mean collections of a kind of information. Those words may include, but not be limited to:

- Database
- Links
- Archives
- Collection

- Similarly useful context terms might include:

- Reference
- Pathfinder
- Bibliography
- Research
- Works Cited
- Museum
- Exhibit
- Encyclopedia

- Suggest students look up definitions for any of the terms they do not know.

- Using terms like these can uncover large collections of primary and scholarly resources in one place, which, from a student perspective, might cut down on the time it takes to complete the research process.

- You might display some examples of the kinds of results a student might find when searching for collections. Examples include:

- [Chinese ceramics database OR archive OR collection OR links] Slide 17.
- [minerals database OR archive OR collection OR links OR reference] Slide 18.
- [primates database OR archive OR pathfinder] Slide 19.
- [political database OR archive OR collection] Slide 20.

- Share with the students these tips to remember when using context terms to find large collections:

- Do not use all the context words you can think of at once. Pick several. Otherwise, results may be unmanageable.
- Remember to use OR between options, e.g.: [database OR links OR archive].
- Carefully select the right term to describe your topic. Remember to match the level of formality that your ideal author would us, e.g., [pottery] vs. [ceramics]. Google brings back both, but stack the deck in favor of the better term by using it.
- Certain terms might have a lot of consumer or shopping results, e.g. [samurai sword database OR archive OR collection OR links] (Slide 21)—Diagnose the difficulty and problem solve by eliminating

some terms with the minus sign, or just erasing problematic terms from your query.
- ○ If you find a great collection, try using [site:] to search across it for pages on your topic. You may find it is not deep web; if it is, you can use site search and browsing.
- ○ Check out the Internet Public Library ([IPL.org](IPL.org)) to find large collections on a single topic.
6. **Apply knowledge to research.** Have students select a few, simple terms to describe their topics, and search for collections of sources that relate to them. They should record any successful queries they use, and add any specific sources they find to their working works cited list. In the next lesson, students will use some of the results pages to practice analyzing the quality of their searches in order to determine next steps.

**ASSESSMENTS:**

- Participation in activities
- Additions to works cited list and queries that located them