1

Digital Information Fluency in an Age of Information Consumption

Searching is becoming easier than thinking. Enter a query in a search engine, and the searcher is instantly flooded with results. Information has never been easier to retrieve and consume. At the same time, determining the quality of the results remains a daunting task. Despite the attempts to make search tools "brain dead easy"[1] to use, searching that reduces the need to think invites problems. Machines cannot reliably predict what each individual is hunting for, machines cannot determine what is credible, yet that is the direction search engine development is headed.

One sign that searching takes less thought is seen in changes to a browser's address bar, now known as the omnibox. A common mistake ten years ago was to type a query in the address bar of a browser rather than the search box of a search engine. Back in the Netscape and Alta Vista heydays, unless the query was also a URL, typing a query in the address bar would yield a "404 page not found" error. This result became an opportunity for a little trial and error learning—why didn't that work? Today one can type a query in the omnibox and specify which search engine results to display for that query. This is a clear accommodation to general patterns of user behavior (and a preference that may be edited). Typing in the wrong box is a thing of the past, a problem overcome by making the address bar more inclusive. As a result, searching is easier; there is no longer a need to think about where to type a query.

Another consumer accommodation involves the query itself. Predictive services offered by Google, Bing, Yahoo!, and others display alternate queries based on user input. For example, if *assessment* is typed in the address bar, many suggestions appear: assessment test, assessment for learning, assessment and taxation, a Wikipedia article on assessment, and a Web page on the characteristics of twenty-first century learners, among others. These suggested queries may or may not be relevant to the user's need, but there's a good chance the user did not have these specific alternatives in mind

when embarking on the search. These automated suggestions could take the search in directions the user did not initially have in mind.

It is certain that search technology will continue to evolve, requiring less effort from the searcher to locate information. However, this does not diminish the need for competent, thoughtful searching and careful consumption, which requires training and deliberation. Instant and predictive results do not diminish the need for strategic searching. Interestingly, solo trial-and-error searching does not result in effective strategies.[2] If the purpose of improved tools is to reduce error, it also diminishes the possibility of learning from one's mistakes through trial-and-error practice. Determining whether instant and predictive results are accurate or objective is optimally a skill learned by practice, not an automated process. Tools may become easier to use, but there is no substitute for instruction and collaboration where searching is involved. Skilled searching is not brain-dead easy.

The short definition of digital information fluency (DIF) is the "ability to find, evaluate and use digital information effectively, efficiently and ethically."[3] Most definitions of information literacy cover the same ground. Take, for example, the competencies described by the American Library Association (ALA):

- Determine the extent of information needed
- Access the needed information effectively and efficiently
- Evaluate information and its sources critically
- Incorporate selected information into one's knowledge base
- Use information effectively to accomplish a specific purpose
- Understand the economic, legal, and social issues surrounding the use of information, and access and use information ethically and legally[4]

The real difference between information literacy and fluency lies not in what searchers or consumers do but how they do it. Fluent individuals have more experience, know more about the "language" they use, and can accomplish what needs to be done more easily using multiple pathways. The emphasis with information fluency is to equip users with sufficient strategies and methods to overcome most retrieval challenges. A fluent searcher knows how digital and print information differs, quickly learns specialized tools for finding digital information, and engages effectively in a digital information environment. The sections that follow provide a framework for understanding these competencies and the dispositions necessary for developing fluency.

SPECULATIVE AND INVESTIGATIVE SEARCHING

Viewed from the highest perspective, information fluency is part of a research process that encompasses two types of searching: speculative and investigative searching. Together, these are part of a larger process that starts with speculation and extends

beyond investigation. One view of the larger picture is the Big Six. The process outlined by the Big Six starts with (1) task definition, followed by (2) information-seeking strategies, (3) location and access, (4) use of information, and (5) synthesis, and ends with (6) evaluation of the product and process.[5] In information fluency terms, the first three steps in the Big Six involve speculation, and the fourth step (use of information) is where investigation occurs when students extract information that is relevant and of quality. The process does not have to be linear, though in practice it tends to start with (1) task definition and end with (6) product and process evaluation. Taken together, mastery of the elements equals information literacy.

Speculation and investigation together do not result in a finished research paper. Each is a different type of research that involves different competencies. The ultimate outcome could be a research paper, but it could also be locating a good-quality TV for the best price, fact checking claims made during an election, or hiring a reputable contractor for a household project. No matter the outcome, the combined objective of speculation and investigation is to locate information that can be trusted. To this end, speculation relies on prediction and guesswork; investigation does not. Understanding the differences between them—as well as the applications of each—is crucial to effective digital research.

Speculative searching is far more common than its investigative counterpart. Everyone who looks for information online is familiar with speculation. Speculative searching is how the search process begins. For many, it is synonymous with querying and browsing. Speculation starts with a question, proceeds to a results page, and moves on to a clicked link. Faced with an inexhaustible supply of information and a limited amount of time, searchers must make calculated guesses about which word(s) to query, which search engine(s) to use, and which result(s) to follow in order to find answers not already known.

Sometimes guessing is unnecessary, for example, using a mobile device to find simple facts like the current temperature or the spices that go into a curry. But even the easiest searches may require guesswork the first time. First-time searches require thinking about which words to search with and where to use them. Consider this example: "Find the top speed of earth's fastest animal." The words to use in such a search include one or more of those that appear in quotation marks as well as additional terms (and operators) yet to be discovered in the process. Many words and combinations are possible. Will the query *top speed of earth's fastest animal* work? Or is *fastest animal* better? Does it make a difference?

It's hard to answer these questions without entering keywords into a search engine and checking results. Before jumping to a conclusion, which search engine should be used? Google's search engine is by far the most common default choice.[6] But is Google guaranteed to work in this case? Will a different search engine yield a different answer? To be sure, the same query must be tried in multiple search engines. Google is a great general-purpose search engine, but it is not the only one. Other choices include Yahoo!, Bing, Ask, AOL, MyWebSearch, Dogpile, WebCrawler, and Hakia. There are also many choices that retrieve content stored in databases unavailable to Google: EBSCO,

JSTOR, Grolier, PubMed, Whois, WayBackMachine, and so on. Selecting Google is a speculative choice; it's one option among many.

The query *fastest animal* entered into Google produces around eight hundred results. The results page states there are over five million results, but the accessible list ends well before that. (For Google, efficiency is more important than depth; it would take more than a few seconds to load up several million records.) A mere ten results on the first page satisfies an average searcher's needs, so trimming the results to only 0.016 percent of what's available to Google for this query remains quite reasonable.

In fact, there is a surplus of information in the top ten results. These animals are mentioned or pictured in the context of *fastest animal*:

- Peregrine falcon (2)
- cheetah (3)
- sailfish (2)
- pronghorn antelope
- wildebeest
- lion
- Thompson's gazelle
- quarter horse
- man

Which is the correct answer? If the top answer (from Wikipedia) is accepted, the answer is the Peregrine falcon. Or should the most frequent answer, cheetah, be accepted? The situation doesn't improve if *top speed of earth's fastest animal* is substituted. Now the top ten results include:

- Peregrine falcon (2)
- man
- greyhound
- cheetah
- cow dropped out of a helicopter

Determining the best answer takes more than looking at the results page snippets (also known as abstracts). Opening and scanning each of the top ten results shows that three sites claim the fastest animal is the Peregrine falcon, while three other sites identify the cheetah. This is a classic student dilemma: Which source is right? Upon closer reading, the results describe the falcon as the fastest *bird*, the cheetah as the fastest *land animal*, and the sailfish as the fastest *fish*. Since the original question merely asked for *animal*, the falcon ought to win hands down. But is Peregrine falcon the best answer? If the falcon beats the cheetah by virtue of gravity, a cow dropped from a sufficient height could break the sound barrier and plummet past the falcon in a flash. A 1960s-space-program chimpanzee or a present-day astronaut orbiting the earth might actually be the record holder, but was that the intent of the question?

Knowing what question to answer makes a crucial difference in recognizing an appropriate answer. Interpreting the question becomes part of the speculative pro-

cess. What exactly is being sought? Whether the question is given as an assignment or arises out of a personal need, speculation may be involved. Searches fail when information is sought to answer the wrong question. The first question one searches for is often not the right question. Clearly articulating what information is being sought is the first step in the information fluency model and helps eliminate a lot of aimless speculation.

Digital searching consists of five sets of competencies.[7] Three of these apply to speculative searching:

- The first, as indicated, is being able to understand the question in order to translate it into a query.
- The second is to know which database to search—and how to find and use an unfamiliar database.
- The third is to figure out how to home in on information through iterative searching, browsing, skimming, and scanning.

Students experience problems with each of these competencies. A fourth problem area—evaluating the results—encompasses the domain of investigative searching, discussed later. The fifth area students find difficult is how to use or consume the findings, which in the context of research means avoiding plagiarism and not violating guidelines of fair use, both matters of ethical use. These five competency sets, along with dispositions, make up the Information Fluency Model.

INVESTIGATIVE SEARCHING

Investigation involves evaluating results by searching with finely honed keywords and specific search engines. Investigative searching requires a skeptical mindset and the perseverance to look well beyond the first page of results. The chief purpose of investigative searching is fact checking: Do the facts or claims found by speculation stand up to scrutiny? Is the content accurate, fresh, and credible? Is the author sufficiently knowledgeable and experienced? Is the publisher reputable and unbiased?

The persistent problem encountered with investigative searching is the failure to do it. Most students—adults, too—accept the face validity of initial results and typically forego investigation,[8] even though in many respects it is easier than speculation. Virtually no guesswork is involved with investigative searching. Using the *fastest animal* example, an investigative search uses *given* keywords and *known* databases to fact check either the source, the content, or both.

Starting with content, a good query is *Peregrine falcon* along with *speed OR fastest animal.* All are keywords found in the assignment and initial results. The purpose of fact checking is to establish how fast the falcon travels. At least two additional sources should be located. Triangulation, comparing three or more sources when fact checking, is an essential tactic for investigation. If several credible, unrelated sites agree, the information appears to be more reliable. If several unrelated sites disagree, then it is wise to remain skeptical and continue the investigation.

Here are the top three results for the query *Peregrine falcon fastest speed:*

- wikipedia.com: 322 km/h (200 mph)
- topspeed.com: easily eclipsing 200 mph
- extremescience.com: 160–440 km/h (99–273 mph)

This information may be found by clicking search results for three different sites, then locating the word *speed* (with the "Find" command, Ctrl+F) on the respective pages. The most conservative speed is faster than any known cheetah. Several sources claim the falcon, thanks to gravity and aerodynamics, can attain such speeds.

So what's to prevent a cow dropped from a helicopter or an orbiting astronaut from being an acceptable answer? It immediately becomes apparent that attempting to answer one's own question by speculative and investigative searching reveals ambiguities or incompleteness in the wording of questions. For educators it's always a good idea to test initial questions by searching online to discover what students are likely to find. Doing so leads to better questions that lead to better results (and fewer sarcastic answers). Poor searching starts with ill-defined questions that fail to be improved by iterative querying. This is why most models, the Big Six and ALA included, start with determining the information needed.

The differences between speculative and investigative searching frame how the search may be understood. The types of questions that guide research activity and the training received will help learners grasp where they are in the process and how to proceed fluently. Speculation and investigation form the two halves of the Digital Information Fluency Model (see figure 1.1). One without the other is an incomplete search. The following five sections introduce the competencies that comprise the whole of information fluency.

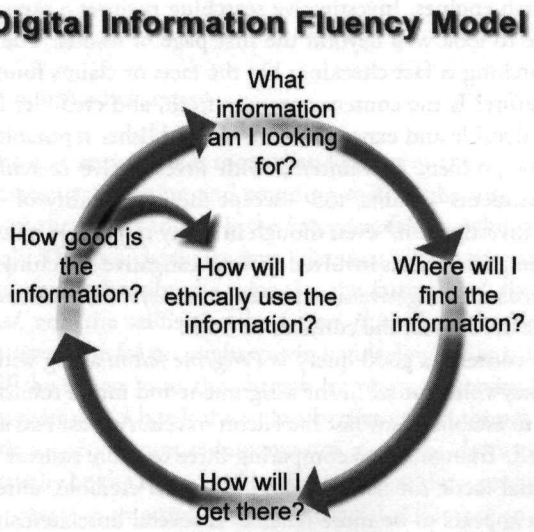

Figure 1.1. Digital Information Fluency Model

WHAT AM I LOOKING FOR?

The Information Fluency Model is based on five questions that occur throughout the search process. What am I looking for? Where will I find the information? How will I get there? Is the information any good? How will I use the information ethically? Like the Big Six, these tend to start with the first question and end with the last, but in between the order is not necessarily linear. In practice, the order is often iterative. As information is retrieved, the speculative question is refined, different sources are consulted, and better keywords are substituted until relevant information is found. As a result of investigation, a decision to revise or accept the information is made. If revised—known as the revision decision—the process may restart with a better speculative question. Without exception, a search is conceived when an individual acknowledges a need for information. In the context of digital searching, no computer is needed at this moment. The "I need information step" happens in the head of the searcher.

How the question arises may alter only slightly what happens next. If the need for information arises from immediate need or, more formally, from a teacher, the process is limited to words in the user's memory, the environment, or the assignment. The words available to the searcher are what matter most. A person with an average vocabulary may be at disadvantage compared to one with an extensive knowledge of words. Why? Because the words taken from memory or the environment may not be the words necessary to locate the information needed. The more words one knows, the better one's chances of matching words used by an online author. Nonetheless, a subject matter expert may use professional vocabulary that is unfamiliar or seldom used by a general audience. Predictive search suggestions may level the playing field somewhat. The words used in assignments may also help. From observation, it is uncommon for students to use keywords in a search other than those given in an assignment. Knowing this, a teacher may introduce words that serve as useful clues. Carefully worded assignments can take a lot of guesswork out of speculative searching. But not all assignments are carefully worded, and self-initiated searches generally start with words—and questions—that need to be improved.

It doesn't take much digging to unearth a complex set of competencies around the question "What am I looking for?" Some of these are bedrock competencies, like memory and the ability to read, both of which are beyond the scope of this book. Others are more application oriented or operational. In this regard the ability to turn a question into a query is the core competency. Knowing how to transform a question into a query involves understanding parts of speech, knowing the meaning of words, and recognizing which words make powerful search terms. Choosing the best words at hand is an effective way to distill the essence of what is being sought.

"Keyword Challenge"[9] on the 21st Century Information Fluency site is designed to help users think about words, their meanings, and how to create effective queries. Users are directed to sort the words in a question into four categories: words that are effective "as is," words that are important but probably not effective as is, words that are unimportant, and "stop" words. Sifting a sentence to identify the good as-is terms and the terms that could be better often distills the meaning of the question in as few

words as possible. Discarding terms that add nothing results in a more powerful query. Students tend to adopt the words given them, so taking this approach makes sense.

To illustrate, take the question "Who first claimed that China's Great Wall could be identified from space?" Sort all the words into one of these categories:

- Good as is: Great, Wall, space
- Important, but there's probably a better word: claimed, identified
- Unimportant (and stop words): who, first, that, China's, can, be, from

As-is words are usually nouns. Objects, in particular proper nouns and numbers, tend to make better search terms than verbs. A verb may be important, but there is often a better word for it. Unimportant words are small parts of speech: pronouns, prepositions, and modifiers. The proper noun *China* is unimportant here because it is redundant. Because the Great Wall is in China, *China* adds nothing.

Keeping the first query to the fewest terms possible, the user must remember one important detail from the question now missing from the query: who. The answer to the question is going to be a person's name. Competence depends on more than choosing what goes in the first query. An effective query retrieves only a collection of relevant information. From this the information consumer must make a selection using terms that may not have been part of the first query. This is an important point. Trying to include everything necessary in a query is not how to get the most value out of a search engine. The user who filters the first round of returns using what could be called an invisible query finds the best results. The searcher scans for terms never entered in the search box but that are important nonetheless. This is another example of how searching is not brain-dead easy. The search engine rarely receives all the speculative information, and if it does, the overloaded search suffers. More about search-box strategies is found in chapters 3 and 4.

WHERE WILL I FIND THE INFORMATION?

The first core competency is coupled with a second: knowing where to search. Once again, this step is speculative and has no active role for a computer.

Where does a search happen? There are chiefly two places to start: a search engine interface or a Web page. A third option, infrequently used, starts with a URL and involves truncating that URL to access directories on a Web site. The user must start somewhere. A competent user bases the decision on a prediction: Which starting point and path will be most effective? The informed user is already thinking ahead to step 3: How will I get there?

Experience, training, or both inform the decision. The experienced user knows that browsing may involve a hard path to follow. Compared to querying or truncation, browsing requires luck and a high degree of sensitivity to hyperlinked words and contextual clues. Browsing can quickly devolve into a hollow game of clicking links. Hoping for the best, the searcher is quickly sidetracked or lost. This approach can be

pleasurable, such as surfing the Web, but trying to do research this way produces frustration. Curiously, browsing seems to be students' default search method. It doesn't take much thought to click a link. It takes a high degree of discipline and persistence to click the right one (and to click the back button when the search goes awry).

A more straightforward approach may be to ask someone knowledgeable. While going to a teacher or librarian may be the shortest path to information, it depends on who is available at the time. Needless to say, this option must be weighed alongside the digital alternatives. So, how does one decide whether to consult a person or turn to a digital resource?

Returning to the Great Wall of China example, a strategic step-2 question to ask is "Who would know this information?" Maybe the searcher knows someone who has been talking about this topic. If not, this is going to involve a lot of guesswork and hunches. Deductive reasoning and imagination may be involved in this competency. Because the topic pertains to space, someone from an organization like NASA might know about it. But who could that be? Someone who's been in space, like an astronaut, might know if the wall is visible or not, but would they know who was the first to claim the Great Wall could be seen from space? That's the real question. Maybe a historian would know best.

Without specific knowledge of this subject matter, deciding whom the expert is—let alone trying to find such a person—is a guessing game. The simple fact that it's hard to imagine someone specific who might know the answer excludes all but one starting point. Starting on a Web page or truncating a URL doesn't make sense because there's no Web page or URL to start with. The appropriate starting place is a standard query.

The Information Fluency Model defines this set of competencies as "learners effectively and efficiently select digital collections based on their characteristics."[10] If the best starting point is a digital collection, competent searchers should understand how digital information is organized and that different collections have different characteristics. A digital collection is a database. Digital information is saved in databases that can be retrieved by searching. Google searches its database when a query is sent through its engine. Yahoo!'s database is separate from Google's, although much of the same information may be found in both, aggregated from public pages on a myriad of sites. NASA's database is different than Google's and Yahoo!'s and contains, potentially, a vastly different collection. Some of NASA's content is proprietary and off limits to other search engines. This is what makes NASA's collection part of the Deep Web: It's only available by going to NASA's site. No other search engine can retrieve it—although Google, Yahoo!, and others index some parts of it. Using the reasoning outlined earlier, a search might start by querying NASA rather than one of the large aggregated databases.

HOW WILL I GET THERE?

Once a question comes to mind and a place to start is determined, it's customary to go to a computer. Numerous competencies cluster around the process of homing

in on information. If a direct query is the preferred method, a search engine page is brought up. If a different starting point is identified, other skills come into play. Each of these and the associated competencies is introduced in turn.

Two sets of competencies go hand in hand to "get there": querying and browsing. For querying, tools consist of a search engine, all the features that come with it, and a results page. Browsing mainly involves scanning Web pages and trying hyperlinks.

Starting with query, the software behind the search box provides powerful functionality, part of which is standard no matter what search engine is used. Other specialized tools are specific to a particular engine. Search algorithms may be proprietary and secret, giving a company a competitive advantage over its rivals. It's not possible to know everything about the tool at the code level, and that's not the searcher's goal anyway. However, a key competency is the ability to use a search engine effectively, which means the fluent user must be able to *learn* to use an unfamiliar engine. This implies a strategic approach to search-engine use on several levels: first-time use, basic search, and advanced search. Information on teaching these strategies is addressed in chapter 3.

As search algorithms advance, the search engine will learn by feeding on queries and provide automated query suggestions, giving the fluent searcher more to think about. If queries are being assessed in order to provide the user with more personalized results, other information must be filtered out. The specter of the filter bubble[11] signals a level of control that may blind the searcher and eventual consumer of information to some results based on incremental data submitted by other searchers pursuing similar results. Implementing ways to overcome the filter bubble is an emerging competency, discussed in chapters 5 and 9.

Competent searchers know that the number and order of terms in a query makes a difference. The more keywords combined, the fewer the results. This may be a good strategy for the end of a search but not its beginning. The most effective first attempt is to use two to five keywords and see what happens. Homing in on information is likely to involve finding better keywords that the user did not start with. To locate these words, it is best to start by searching with just a few terms in the query.

Fluent searchers also understand concepts of linguistic granularity, known as hyponyms and hypernyms. While the terms may sound technical, the ideas are not: Many words have synonyms that are more general (hypernyms) and other synonyms that are more specific (hyponyms). The ability to adjust the specificity of a term can be crucial to querying. For the keyword *automobile*, *vehicle* is a hypernym and *Chrysler* a hyponym. The importance of choosing the right grain-size term depends on the number of results retrieved by a query. Faced with too few results, a hypernym will broaden the search as well as the results. A hyponym will narrow the range of results.

Search operators are additional tools the competent searcher can use. Improvements to search-engine algorithms have reduced the need for operators but not removed them entirely. Many search engines built on older code still require Boolean operators, so it is in the fluent searcher's best interests to know about operators. In brief, operators may be classified as simple or advanced. Simple operators are AND

(+), OR (|), NOT (-), and "ALL," where "ALL" is replaced by a search string or phrase. Most search engines have eliminated the need for AND (+) by making it the space-bar function. The purpose and application of these and advanced operators, such as "link:" and "site:," are examined later in this book (chapter 5).

Returning to the earlier example, if the query *great wall space* with no operators is entered into Google, a page of snippets is displayed filled with relevant-looking information and even a possible answer. Something that happens just before the results are delivered should not be overlooked. Google suggests adding the term *myth*. Apparently other searchers associate the term *myth* with this query. This is an immediate red flag and implies something may be amiss with the claim that the wall can be seen from space.

As soon as the query comes back, homing in switches from querying to browsing. Browsing includes skimming or reading results snippets, scanning for specific terms, clicking on links, and reading the URL. Skimming, the more general reading skill, involves quickly reading the content of a page to get a sense of its meaning. Scanning is more targeted: looking for specific terms, names, or images on a page. It is possible for homing in to stop once the first snippets are obtained, however this is not a good decision. The name *Jake Garn* is found by skimming the contents of the first ten results without having to click through to any other content. This is a speculative answer (and incorrect). A user who fails to investigate if astronaut Jake Garn is the first person to make this claim has performed only half a search.

Typically, the process of homing in is more complicated. For example, in this case the first snippet is from an article on NASA's site. Because NASA could be a credible source on the topic, this article seems worth browsing (skimming the content or reading more closely). The article claims the wall is difficult, if not impossible, to detect with the naked eye from close-earth orbit. Halfway down the page is this statement: "The theory that the wall could be seen from the Moon dates back to at least 1938. It was repeated and grew until astronauts landed on the lunar surface."[12] While it doesn't mention who theorized the visibility of the wall, the article contains an important clue: 1938. This date is a very specific hyponym for *when*. The competent searcher now has an improved query: *great wall space 1938*. This second query is very likely to retrieve the same NASA article in addition to other articles that may actually name the "who" in question.

Browsing the results, the first snippet for the improved query now contains a new name: Richard Halliburton. The second snippet, snopes.com, also names the same person, found by reading the article. Using the "Find" command to locate *1938* on the page, Snopes states that while the original source is unknown, Halliburton's book *Second Book of Marvels: The Orient* attributes the claim to astronomers.[13] Several other snippets lead to information naming Halliburton as at least a catalyst for spreading the idea. Jake Garn, whose mission did not launch until forty-two years later, certainly was not the first person to make this claim.

The same Snopes article refers to Henry Norman, whose 1904 book *The People and Politics of the Far East* talks about the wall being visible from the Moon. This

predates Halliburton and the space program and thus far is the earliest known reference to the theory. There may be no definitive answer to the question; in fact, the question could now be improved. A little searching has provided two names and two dates for information published long before anyone could substantiate it.

Homing in, as this illustrates, may involve alternating between querying and browsing. Browsing often leads to more specific keywords and stronger queries in an iterative process. But there are other ways that rely predominantly on an alternation of hyperlink clicking and page skimming. Like surfing the Web, this involves clicking through an indeterminate number of pages until the user discovers something interesting or useful. Hyperlinks and their surroundings are the critical tools at the searcher's disposal. A great deal of information is located this way and is exactly how most searches end. Like public transportation, a query is good at getting one close to a destination. Browsing is like having to walk the remaining distance. Here's where the analogy differs: Unlike walking the rest of the way home, the location of home may or may not be close, and there is no map.

Even when "home" is only two clicks away, it is possible to get hopelessly lost when searching for digital information. As part of an assessment of the search skills program developed by the 21st Century Information Fluency Project for the Center for Talent Development at Northwestern University, participants were directed to find information about members of the Actroid robotics team on a company website.[14] This turned out to be one of the hardest challenges in the assessment, despite the information being just two clicks away. A considerable number of students got frustrated and consulted their parents. Their parents got lost.

The task is completed by using a search engine to get close and browsing the rest of the way. A search for *Actroid* on the company site results in two links. (A query is unable to find the required page.) With hindsight and a map, it doesn't seem overly complicated:

- First click: "Intelligent Systems Research Institute"—This snippet includes a literal reference to Actroid Group Members.
- Second click: "Group HP"
- Information located

But these are not the only links that could be clicked. Thinking of it as a path, the first junction is a fork with two choices: "Intelligent Systems Research Institute" or a link/snippet written in Japanese. Taking the first choice leads to a page of information and another junction: "Group Members" and "Group HP." The first link seems more likely but doesn't match the specific information in the directions. "Group HP" does that, but to get to that link, the user might need to back up. Besides being sensitive to terms (What is the closest match?) and context (Am I getting closer?), the user must be able to retrace steps (Where was I before this when things made more sense?).

Browsing without a map is like playing the children's game "Hot and Cold": The closer one is to the information, the warmer one becomes, except here the user is

never explicitly told what to do. The context must be interpreted to determine if the search is headed in the right direction. Even when the user has a one-in-four chance of clicking on the right links, browsing can be a challenge.

Related both to querying and browsing is truncation, another way to home in. As it applies to keywords, truncation involves removing the end of a word to make it less specific, converting it to a hypernym. For example, *instructor* or *instruction* may be truncated to *instruct** (the * is a wildcard operator), telling the search engine to find words that start with *instruct* but may end with anything else. This may yield a wider range of results, a sound strategy to expand a search.

Truncating a URL removes tailing sections of a Web address divided by slashes. For example: https://21cif.com/tutorials is a truncated version of https://21cif.com/tutorials/micro. This is a quick way to navigate to directories or other pages of a site. Sometimes access is forbidden; many times it is not, and the user gets to see the contents of a directory or an index page. Truncation may be useful to home in on information, however it is more commonly used as an investigative technique to find information about an author, a publication date, or other content.

In order to locate information, a fluent searcher must understand effective ways to construct a query, then use browsing either to locate information or, more frequently, find better keywords from which to build more powerful queries. A fluent searcher:

- depends heavily on language skills,
- has the ability to learn how to use a search engine,
- knows how to overcome the filtering effects of a search engine,
- edits queries to broaden or narrow their focus,
- knows how to use operators,
- finds better keywords in snippets and articles,
- navigates through hyperlinks,
- truncates keywords and URLs, and
- strategically knows when to use the methods that make the most sense.

EVALUATION

Most people recognize "If it's on the Internet, it must be true" is meant as a joke. The phrase has been popularized in posters, cartoons, and television commercials. Despite this, it's not uncommon to fall victim to erroneous information online. Among information's many qualities are accuracy, objectivity, and freshness. To the fluent evaluator/consumer, each one of these qualities is important. The ability to detect inaccuracies, bias, and outdated information is essential. Otherwise, the outcome is hardly different than believing everything online is true.

Information online can be inaccurate due to a variety of causes. Misinformation can occur because of unintended mistakes, like typos, or more deliberate attempts to

deceive, like scams. Some misinformation is harmless, even fun; some is malignant and may properly be labeled malinformation.[15] The fluent evaluator aims to find credible, consumable information for research and seeks neither to be misinformed nor malinformed.

This is accomplished through investigative searching. Like speculative searching, this involves querying, browsing, reading, and thinking. The main difference is that the searcher starts with information from "homing in" to be evaluated. Evaluation encompasses the first three questions but with far less speculation. For example, instead of *great wall china*, an investigative query might read *jake garn great wall.* Here the evaluator is gathering evidence about Mr. Garn: Who is Jake Garn? Did he say this? When? Why? Was he the first? If the user stopped after just one query and turned to investigate a single clue (Jake Garn), it would take no time to see that, while this material is accurate, it fails to answer the question, and needless consumption of the wrong information would be avoided.

Knowing what clues to investigate is half the competency. The other half is actually doing it. The difficulty with information fluency is having the dispositions to practice strategies and methods that are not difficult to use. Developing confidence, persistence, curiosity, open-mindedness, and flexibility is, arguably, more challenging than learning how to identify a good keyword to investigate. Yet lacking sufficient dispositions, all but the easiest searches become exercises in incompetence. Getting students to *use* sound search strategies is harder than teaching them how to do it.

Unfortunately, the Internet provides plenty of opportunities for information consumers to be fooled. It takes constant vigilance to avoid Internet hoaxes and scams. The way to avoid being fooled is through investigation. Investigating a Web site or an author's claim can be a rewarding experience on several levels. First, there's the joy of discovering that something that sounds good is not what it seems (or vice versa). There's also the intrinsic enjoyment of the problem-solving process, which can be very motivating. For students, the process is stalking clues like a detective. Investigation can become gamelike. There is almost always something to be discovered.

Evaluation starts by examining a search result to determine its usefulness. Does it address the natural language question? If not, how can the search be revised to produce a more accurate result? This is the revision decision that leads to selecting other search results or changing the query based on the results. If information passes the initial "sniff test," it is time to evaluate either the reliability of the source or the accuracy of the content.

Whether source or content is selected first depends on what looks more suspicious. If a fact or claim seems implausible, then triangulating that information is a good first step. If something about a Web site or author (e.g., of a blog) seems unusual, then investigating the source is a good place to start. In either case, the fluent searcher must be able to detect red flags in information that looks suspicious, out of place, tampered with, or just not right.

What makes digital information appear "not right" is hard to define and varies from one person to the next, but a skeptical searcher knows it when he or she sees it.

Red flags may include typos, misplaced or misused words, graphics that don't look right or are out of place, claims that seem outrageous or too good to be true, facts that are missing, advertisements, popups, personalizations, other evidence of profit seeking, and viruses that are downloaded. The ability to detect red flags includes having a gut feeling something is just not right without being able to put one's finger on it. Masahiro Mori coined the term *uncanny valley* to describe the creepy feeling people get when confronted with robotics that are too lifelike.[16] People are increasingly put off the more a robot looks human. Something in what they see raises a red flag. Even though the robot's nonhuman characteristics are minimal, on some level of awareness, they are perceptible. Information that may be just a little off can have the same unsettling effect. When that happens, it's usually a good idea to follow one's instincts and investigate the content, the source, or both.

Content evaluation takes two forms: internal accuracy and external validation. It's more of a task to determine the accuracy of contents than to check external sources for validation. Internal accuracy revolves around what the author wrote and how and when the author said it. In some cases content may be inaccurate because it doesn't coincide with or support other facts in the text. Crucial details may be omitted or extraneous information may be included to cover up facts that are missing. At other times, information will be inaccurate simply because it was mistyped or written so long ago that it is no longer current. Is the information fresh or outdated? Determining the date information was written or published is an essential competency.

Author bias is also a factor that may contribute to inaccuracy. Bias, or subjective viewpoint, potentially skews content and communicates information about the objectivity of the author. Information passed through the lens of personal perspective may become distorted, exaggerated, or cloaked. This happens to authors and readers alike. Detecting internal inaccuracies and bias depends on careful, objective, even slow reading, something hurried searchers tend to gloss over. The connection between careful reading and evaluative searching should not be taken lightly. Without sufficient language skills, searchers will be blind to internal inconsistencies or information that should be there and isn't. Without self-awareness, searchers will be blind to the bias they bring to their reading. This has direct implications for reading levels and when the time is right to teach information fluency in schools (chapter 2).

External validation is, in many respects, easier to determine. It requires plugging a found fact like *peregrine falcon* or *jake garn* into a query along with one or two given details. Competent searchers can extract a fact from a page. Oftentimes these facts are proper nouns or numbers, as previously discussed. Other times, the fact may be a string, in which case it may be good to use it with the quotes operator. For example, to check the external validity of the phrase, "If it's on the Internet, it must be true," the entire phrase could be pasted in a query as is. Results will be all the instances that match the phrase exactly (if quotes are used) or come close (if quotes are not used). Google reports over 800,000 instances of the exact phrase in its database of digital information. Obviously, it's a very popular phrase and not one the present author can claim to have originated. If that were the claim, it would be easy to determine its

external reliability by including the name *Carl Heine* with the string. There is only one relevant result for that query, and it is the author's blog.[17] No site supports the claim. However, if the claim to be evaluated is whether this phrase occurs in posters, cartoons, and television commercials, external results unequivocally support that claim.

Source evaluation takes into account the credentials and point of view of the author and/or publisher. Where did the information come from? Who is responsible? Identifying the author and publisher is one aspect of this competency. Another is fact checking their credentials. What qualifies them as an expert source? What education have they had? What is their relevant experience? What else have they written or published? On what basis do they select information to be published? Questions that are answered by fact checking, supported by careful reading, matter greatly in evaluating source credibility.

Content that passes the red-flag test reduces the need to evaluate source credibility. There's no reason to check one's credentials unless the content is important and the author's reputation is unknown. Credential fact checking makes sense before acting on an unknown or alleged source's recommendation (e.g., to avoid scams). It also makes sense before including an unknown source in a research paper (e.g., no one has ever cited this person before). The fluent searcher knows when to check the source.

Thinking back on the Great Wall example, would Snopes need to be fact checked? Not really. The site is a popular source for checking urban legends and other claims. However, if the user had never heard of Snopes before, it would be wise to see who links to Snopes. If no one links to the snopes.com site, then the user has a reason to be skeptical, at least initially. To discover who links to the site, the searcher must know how to use the "link:" operator or perform a domain name check or both. The result of a "link:" query for snopes.com provides a long list of sites that bolster the face validity of Snopes: "Directories, Calendars, Research Guides, Encyclopedias and Hoaxes,"[18] the English Department of St. Columba's College,[19] Pasadena City College,[20] and many more. More about external link checking—also known as backlinks—and domain ownership is found in chapter 5, "Investigative Searching."

Many factors can go under the microscope in order to evaluate source credibility. In addition to what may be normally found on an author's resume, how the author says something can reveal something personal about beliefs and values. This can supplement what is provided in a resume. Therefore, bias detection is not only a valuable part of content evaluation; it offers potential insights about an information source. Many books have been written on bias,[21] so this one is not going to add anything new to understanding this topic. Most of the time, competent bias detection will correctly recognize propaganda techniques, stereotypes, language, and images that exaggerate (hyperbole), as well as underemphasize and omit details. Bias detection, often challenging to teach, should be done in the context of content and source evaluation. With so many authors uploading digital information, there has never been a greater opportunity—and need—to evaluate bias.

ETHICAL USE

Ethical use is the ultimate destination of the search process. The question "How will I ethically use the information?" lies at the heart of the matter. Once the decision is made that the information is good enough to be used, the task becomes how to incorporate digital resources as part of one's work. In an ideal world, researchers and writers always cite the source of the information they use. However, students do not always practice ethical use: copying and pasting, remixing and recycling other authors' insights and facts without citation.[22] Acknowledging the origin of ideas while contributing unique insights is what thinking, learning, communication, and education is all about in an Internet-connected world.

Educators know that students fail to cite digital materials, and when they do, the citations can be so weak that it is difficult or impossible to verify their sources. This is a persistent problem rooted in student inclinations to accept the first results returned by a search and to copy those results blindly without citation into their research work. According to a survey of educators by turnitin.com, copying and pasting is the most common form of plagiarism, followed by paraphrasing someone else's work.[23] When confronted about plagiarism, students may plead ignorance, but some openly admit that this type of cheating is a shortcut that saves them time.[24] Teachers who think that students should arrive in their classrooms understanding citation and plagiarism believe it is someone else's job to teach these skills. Assuming plagiarism instruction is someone else's responsibility is not a good instructional strategy, nor does it help students become ethical researchers or consumers of information. Teaching the nuances of proper citation and paraphrasing takes preparation. Ways to teach plagiarism avoidance and proper citation are detailed in chapter 7, "Ethical Consumption."

Ethical use begins when a searcher decides to integrate digital information into the research task. Fluent searchers know proper ways to use quotations or paraphrase and then accurately cite the source. The use/not use decision depends on the quality of the information and its relationship to the search task. Information that is deemed useful will likely find its way into the research project. How that information is managed is critically important. For example, information gathered during the search may need to be bookmarked and perhaps marked up with annotations and notes. This helps greatly when it comes time to quote and cite the materials. This can be accomplished with specialized tools like Diigo[25] or Evernote,[26] which may need to be learned on one's own.

Deciding how to attribute text and media resources is based on knowledge of copyright law and fair-use guidelines. In casual blog writing, a hyperlink may function as an informal citation because it can take the reader directly to the source. If a writer is working from a blog or curated platform, hyperlinking an author's name back to a source and proper use of quotations is sufficient citation to insure ethical use of the information. Most writing does not require a formal citation described by the American Psychological Association (APA), Modern Language Association

(MLA), or Chicago Manual of Style (CMS), to name a few. Proper citation (and style) is most important in academic writing and journalism. In these circumstances, citation alone is not enough. The writer must accurately cite the source even when the source has been paraphrased. For technical and academic work, footnotes may be required for both direct quotation and paraphrased writing. The literary ability to quote, paraphrase, and cite sources properly is essential for ethical use.

Developing a formal citation is a demanding task when working with online resources. Unlike print materials, online resources do not always present the author, date, and copyright information in a manner that is easy to find. This means the searcher must use investigative search techniques described earlier in this chapter to ferret out the information required to create a formal citation. Tracking down details for a citation is much the same process involved in finding the author and publisher when investigating credibility and bias. With this in mind, citation may be seen as one aspect of the detective game of investigative searching. Indeed, investigative searching is necessary before a writer selects a source in the first place. Investigative searching combined with the intent to cite and the ability to organize online sources using social bookmarking or cloud-based research tools defines a fluent researcher's skill set.

Students may think of citation as a highly demanding formatting experience where the order of author's names or the proper placement of a comma or use of italics are the most important elements of the process. This focus on form over function diverts attention from critical thinking and the investigative nature of the process. Online citation generation systems like the 21CIF Citation Wizard[27] provide templates that insure proper academic formatting so searchers can concentrate on investigation.

An understanding of fair use is also part of a fluent searcher's competencies. Using materials ethically means knowing how much of copyrighted work to quote or what percentage of a video to show. The Four Factors test helps navigate the gray areas of ethical use. Skilled searchers will understand the importance of the purpose and character of the use, the nature of the copyrighted work, the amount and substance of the portion taken, and the effect of the use upon the potential market.

When it comes to actual use of online writing, video, or audio, there are limitations as to how much of the materials can be used. These limitations vary based on the purpose of the writing and the copyright restrictions of the materials. Fair use is addressed more fully in chapter 7. To be an ethical user, the fluent searcher must have a deep understanding of:

- intellectual property,
- paraphrasing versus copying,
- educational fair use,
- citation of sources, and
- writing conventions of research.

Ultimately, the disposition of the fluent searcher is to use technology to organize the digital information found in a way that ensures citation and attribution. When

researching a topic, the fluent searcher aims for transparency, creating an archive of bookmarks, often including annotations and notes stored in a cloud-based repository. When the information task is complete, the writer creates a document that allows the (online) reader easily to click through to the original sources or, in the case of print, turn to footnotes, endnotes, or a bibliographic reference. Using information-fluency skills in this way honors copyright and safeguards the fair use of intellectual property.

TEACHING INFORMATION FLUENCY

The model of information fluency presented in this book includes many practical applications for teaching and learning. While this book is not explicitly a curriculum, it contains many curricular resources, starting with discreet competencies, such as converting natural-language questions into queries, and culminating in research strategies. For example, chapter 6 offers four mini-lessons in investigative searching. Chapter 8 demonstrates three ways to embed information-fluency instruction into existing lessons. Finally, chapter 9 presents a vision for engaging students as curators of information that synthesizes all the steps in the Information Fluency Model. The appendix conveniently groups the model lessons into one table.

There is no national curriculum for information fluency, nor is one likely to emerge any time soon. Not many schools allow time for courses or even units of instruction dedicated to this discipline. Consequently, educators who wish to help students become fluent must be inventive with the model resources provided.

The reader is advised not to treat the examples as ready-to-use lesson plans but as a vision for what is possible. The recommended approach is to embed learning opportunities into existing courses where students are expected to conduct digital research: language arts, social studies, and science. Ascertain students' strengths and weaknesses by giving them a challenge and watching them search. Then provide specific instruction and practice. While this work would best be coordinated at a district or building level, the efforts of individual librarians and teachers working alone nevertheless makes a difference in helping students achieve fluency.

SUMMARY

Although tools continue to improve, there is no indication that searching will ever become brain-dead easy. Finding information is not enough. Without evaluation there is no information fluency; having to investigate results makes thinking imperative. The highly skilled searcher outlined in this chapter is the exemplar of fluency. The searcher is in control of the tools, not the other way around. The filter bubble is no hindrance; the fluent searcher is not limited to what personalized search engines are programmed to retrieve. From an array of strategies and methods, optimal approaches are selected

to meet the demands of each search challenge. If one approach doesn't work, another one is available. The fluent searcher locates information efficiently, evaluates it effectively, and uses it ethically.

To become fluent takes training and practice. There are some individuals the authors have observed who have all the right natural inclinations. But this is not the norm. More common, even among gifted populations, is the student who is effective about half the time. Being right only 50 percent of the time is unacceptable performance in most educational activities. With just a few hours training, the same students improve to being right two out of three times; many exceed that mark. With repeated exposure to good strategies and practice, students continue to improve, as the next chapter illustrates.

NOTES

1. Matt Richtel, "A Silicon Valley School That Doesn't Compute," *New York Times*, last modified October 22, 2011, http://www.nytimes.com/2011/10/23/technology/at-waldorf-school-in-silicon-valley-technology-can-wait.html.

2. Carl Heine and Dennis O'Connor, "Final Report for the Center for Talent Development" (unpublished report, Chicago, IL, 2010), 2.

3. David Barr, Bob Houston, Carl Heine, Dennis O'Connor, and Dan Balzer, "Digital Information Fluency Model," *21st Century Information Fluency*, last modified January 1, 2013, https://21cif.com/resources/difcore/index.html.

4. "Information Literacy Competency Standards for Higher Education," *American Library Association*, accessed December 15, 2012, http://www.ala.org/acrl/standards/informationliteracycompetency.

5. Mike Eisenberg and Bob Berkowitz, "Introducing the Big6," *The Big6*, last modified January 28, 2012, http://big6.com/pages/lessons/presentations/big6-overview-2011.php.

6. Kristen Purcell, Lee Rainie, Alan Heaps, Judy Buchanan, Linda Friedrich, Amanda Jacklin, Clara Chen, and Kathryn Zickuhr, *How Teens Do Research in the Digital World* (Washington, DC: Pew Research Center, November 1, 2012), http://www.pewinternet.org/~/media/Files/Reports/2012/PIP_TeacherSurveyReportWithMethodology110112.pdf.

7. Carl Heine, "Five Things Today's Digital Generation Cannot Do," *21st Century Information Fluency*, last modified January 1, 2013, https://21cif.com/resources/features/fivethings_directory.html.

8. Eszter Hargittai, Lindsay Fullerton, Ericka Menchen-Trevino, and Kristin Yates Thomas, "Trust Online: Young Adults' Evaluation of Web Content," *International Journal of Communication*, *4* (2010): 468–94.

9. Carl Heine, "Keyword Challenge," *21st Century Information Fluency*, accessed January 20, 2013, https://21cif.com/rkitp/challenge/v1n3/Keyword_Challenge/KeywordChallenge.swf.

10. See note 3.

11. Eli Pariser, "Filter Bubbles, Meet Upworthy," *The Filter Bubble*, March 26, 2012, http://www.thefilterbubble.com.

12. "China's Wall Less Great in View from Space," *NASA*, last modified November 30, 2007, http://www.nasa.gov/vision/space/workinginspace/great_wall.html.

13. "Great Walls of Liar," *Snopes.com*, last updated September 6, 2010, http://www.snopes.com/science/greatwall.asp.

14. Carl Heine and Dennis O'Connor, "Final Report for the Center for Talent Development" (unpublished report, Chicago, IL, 2012).

15. SchemeHater, "Malinformation," *Urban Dictionary*, last modified April 20, 2009, http://www.urbandictionary.com/define.php?term=Malinformation.

16. Masahiro Mori, "The Uncanny Valley," *Energy*, *7*, no. 4 (1970): 33–35. Translated by Karl F. MacDorman and Takashi Minato, *Android Science*, last modified September 19, 2005, http://www.androidscience.com/theuncannyvalley/proceedings2005/uncannyvalley.html.

17. Carl Heine, "Tips from Google: What's Missing?" *Internet Search Challenge*, last modified July 3, 2012, http://internetsearchchallenge.blogspot.com/2012/07/tips-from-google-whats-missing.html.

18. "Directories, Calendars, Research Guides, Encyclopedias and Hoaxes," *Online Public Relations*, last modified December 22, 2012, http://www.online-pr.com/directory.htm.

19. "SCC English," *The English Department of St. Columba's College*, accessed January 20, 2013, http://www.sccenglish.ie.

20. "Useful Links," *Pasadena City College*, accessed January 20, 2013, http://www.pasadena.edu/studentservices/lac/links.cfm.

21. For example, Thomas Kida, *Don't Believe Everything You Think: The 6 Basic Mistakes We Make in Thinking* (Amherst, NY: Prometheus, 2006).

22. "White Paper: The Plagiarism Spectrum: Instructor Insights into the 10 Types of Plagiarism," *Turnitin.com*, accessed January 21, 2013, http://pages.turnitin.com/rs/iparadigms/images/Turnitin_WhitePaper_PlagiarismSpectrum.pdf.

23. See note 22.

24. Kathy Slobogin, "Survey: Many Students Say Cheating's OK," *CNN.com*, last modified April 5, 2002, http://articles.cnn.com/2002-04-05/us/highschool.cheating_1_plagiarism-cheating-students.

25. https://www.diigo.com.

26. https://evernote.com.

27. Carl Heine, "Citation Wizard," *21st Century Information Fluency*, last modified February 12, 2012, https://21cif.com/tools/cite.