Evaluating Information Found on the Internet

Introduction

The World Wide Web offers information and data from all over the world. Because so much information is available, and because that information can appear to be fairly "anonymous", it is necessary to develop skills to evaluate what you find. When you use a research or academic library, the books, journals and other resources have already been evaluated by scholars, publishers and librarians. Every resource you find has been evaluated in one way or another before you ever see it. When you are using the World Wide Web, none of this applies. There are no filters. Because anyone can write a Web page, documents of the widest range of quality, written by authors of the widest range of authority, are available on an even playing field. Excellent resources reside along side the most dubious. The Internet epitomizes the concept of Caveat lector: Let the reader beware.

This guide discusses the criteria by which scholars in most fields evaluate print information, and shows how the same criteria can be used to assess information found on the Internet. Use the tabs on this guide to further explore and consider how to effectively evaluate online information.

(With gratitude to Elizabeth E. Kirk for compiling the information about evaluating information sources from the Internet.)

Items to Consider

Authorship

This is perhaps the major criterion used in evaluating information. Who wrote this? When we look for information with some type of critical value, we want to know the basis of the authority with which the author speaks. Here are some possible filters:

- In your own field of study, the author is a well-known and well-regarded name you recognize.
- When you find an author you do not recognize:
  - the author is mentioned in a positive fashion by another author or another person you trust as an authority;
  - you found or linked to the author's Web/Internet document from another document you trust;
  - the Web/Internet document you are reading gives biographical information, including the author's position, institutional affiliation and address;
  - biographical information is available by linking to another document; this enables you to judge whether the author's credentials allow him/her to speak with authority on a given topic;
  - if none of the above, there is an address and telephone number as well as an e-mail address for the author in order to request further information on his or her work and professional background. An e-mail address alone gives you no more information than you already have.

Accuracy and Verifiability of Details

Accuracy or verifiability of details is an important part of the evaluation process, especially when you are reading the work of an unfamiliar author presented by an unfamiliar organization, or presented in a non-traditional way. Criteria for evaluating accuracy include:

- For a research document, the data that was gathered and an explanation of the research method(s) used to gather and interpret it are included.
- The methodology outlined in the document is appropriate to the topic and allows the study to be duplicated for purposes of verification.
- The document relies on other sources that are listed in a bibliography or includes links to the documents themselves.
Steps for evaluating point of view are based on authorship or affiliation:

- Read wants to prove his point, and will use the data and information that assists him in doing so. When evaluating information found on the Internet, it is important to examine Point of view or bias.

- The publishing body also helps evaluate any kind of document you may be reading. In the print universe, this generally means that the author's manuscript has undergone screening in order to verify that it meets the standards or aims of the organization that serves as publisher. This may include peer review. On the Internet, ask the following questions to assess the role and authority of the "publisher", which in this case means the server (computer) where the document lives:
  - Is the name of any organization given on the document you are reading? Are there headers, footers, or a distinctive watermark that show the document to be part of an official academic or scholarly Web site? Can you contact the site Webmaster from this document?
  - If not, can you link to a page where such information is listed? Can you tell that it's on the same server and in the same directory (by looking at the URL)?
  - Is this organization recognized in the field in which you are studying?
  - Is this organization suitable to address the topic at hand?
  - Can you ascertain the relationship of the author and the publisher/server? Was the document that you are viewing prepared as part of the author's professional duties (and, by extension, within his/her area of expertise)? Or is the relationship of a casual or for-fee nature, telling you nothing about the author's credentials within an institution?
  - Can you verify the identity of the server where the document resides? Internet programs such as dnslookup and whois will be of help.
  - Does this Web page actually reside in an individual's personal Internet account, rather than being part of an official Web site? This type of information resource should be approached with the greatest caution. Hints on identifying personal pages are available in Understanding and Decoding URLs.

- The document includes the date(s) at which the information was gathered (e.g., US Census data).
- The document refers to clearly dated information (e.g., "Based on 1990 US Census data.").
- Where there is a need to add data or update it on a constant basis, the document includes information on the regularity of updates.
- The document includes a publication date or a "last updated" date.
- The document includes a date of copyright.
- If no date is given in an electronic document, you can view the directory in which it resides and read the date of latest modification.

If you found information using one of the search engines or any of the services that rate World Wide Web pages, you need to know:

- How the search engine decides the order in which it returns information requested. Some Internet search engines "sell" top space to advertisers who pay them to do so. Read Pay for Placement from Searchenginewatch.com.
- That Internet search engines aren't like the databases found in libraries. Library databases include subject headings, abstracts, and other evaluative information created by information professionals to make searching more accurate. In addition, library databases index more permanent and reliable information.
- How that search engine looks for information, and how often their information is updated. An excellent source for search engine information is Search Engine Showdown, written by Greg R. Notess. All information, whether in print or by byte, needs to be evaluated by readers for authority, appropriateness, and other personal criteria for value. If you find information that is "too good to be true", it probably is. Never use information that you cannot verify. Establishing and learning criteria to filter information you find on the Internet is a good beginning for becoming a critical consumer of information in all forms. "Cast a cold eye" (as Yeats wrote) on everything you read. Question it. Look for other sources that can authenticate or corroborate what you find. Learn to be skeptical and then learn to trust your instincts.

- Understanding and Decoding URLs

Point of View Bias

Point of view or bias reminds us that information is rarely neutral. Because data is used in selective ways to form information, it generally represents a point of view. Every writer wants to prove his point, and will use the data and information that assists him in doing so. When evaluating information found on the Internet, it is important to examine who is providing the "information" you are viewing, and what might be their point of view or bias. The popularity of the Internet makes it the perfect venue for commercial and sociopolitical publishing. These areas in particular are open to highly "interpretative" uses of data.

Read Information and its Counterfeits: Propaganda, Misinformation and Disinformation for learn more about "interpretational views" that exceed the facts.

Steps for evaluating point of view are based on authorship or affiliation:

- First, note the URL of the document. Does this document reside on the Web server of an organization that has a clear stake in the issue at hand?
  - If you are looking at a corporate Web site, assume that the information on the corporation will present it in the most positive light.
  - If you are looking at products produced and sold by that corporation, remember: you are looking at an advertisement.
If you are reading about a political figure at the Web site of another political party, you are reading the opposition.

- Does this document reside on the Web server of an organization that has a political or philosophical agenda?

- If you are looking for scientific information on human genetics, would you trust a political organization to provide it?

- Never assume that extremist points of view are always easy to detect. Some sites promoting these views may look educational. To learn more, read "Rising Tide: Sites Born of Hate", New York Times, March 16, 1999. (This link will take you to the online edition of the Times; you must register, free of charge, to view the article).

Many areas of research and inquiry deal with controversial questions, and often the more controversial an issue is, the more interesting it is. When looking for information, it is always critical to remember that everyone has an opinion. Because the structure of the Internet allows for easy self-publication, the variety of points of view and bias will be the widest possible.

**Referral to and/or Knowledge of the Literature**

**Referral to and/or knowledge of the literature** refers to the context in which the author situates his or her work. This reveals what the author knows about his or her discipline and its practices. This allows you to evaluate the author's scholarship or knowledge of trends in the area under discussion. The following criteria serve as a filter for all formats of information:

- The document includes a bibliography.
- The author alludes to or displays knowledge of related sources, with proper attribution.
- The author displays knowledge of theories, schools of thought, or techniques usually considered appropriate in the treatment of his or her subject.
- If the author is using a new theory or technique as a basis for research, he or she discusses the value and/or limitations of this new approach.
- If the author's treatment of the subject is controversial, he or she knows and acknowledges this.

**Distinguishing Propaganda and Misinformation**

**Information and Its Counterfeits: Propaganda, Misinformation and Disinformation**

If counterfeiters put pictures of their family members on their handiwork, nobody would be fooled. *What constitutes a good fake is how well it resembles the real thing.* This document will help you be able to distinguish real information from its three lookalikes, or counterfeits: propaganda, misinformation and disinformation. Understanding the counterfeits will enable you to become a much more critical consumer of information.

Information, **noun**

"Knowledge communicated concerning some particular fact, subject, or event; that of which one is apprised or told; intelligence, news. spec. contrasted with data." (from Oxford English Dictionary, 2nd ed., 1989)

8,000,000 persons, a growth of 9% since 1990" is indeed information. Adding that information to other information and data on the funding of and expansion in public healthcare services in New York City would help city officials to develop knowledge of the stresses related to delivering healthcare services.

- Read "**Freedom of Expression on the Internet**, a report issued by Human Rights Watch.

You probably found one of these documents more difficult to understand than the other. Both are information. To what extent is the "understandable" or "contextual" nature of information dependent upon our prior knowledge or familiarity with a given topic? Is the value of information intrinsic, or dependent upon the user?

Information should always be accurate and either free of bias or making note of its own bias. Information also needs to be useful for a given purpose to have value.

Propaganda

Propaganda is a commonly misused term. Because of its historical use, such as in the name of the infamous "National Ministry for Public Enlightenment and Propaganda" run by Joseph Goebbels for the Nazi government of Germany, many people associate propaganda with inflammatory speech or writing that has no basis in fact. In reality, propaganda may easily be based in fact, but facts represented in such a way as to provoke a desired response.

Propaganda, **noun**
Evaluating Information Found on the Internet - Library Guides at Johns Hopkins University

"The systematic propagation of information or ideas by an interested party, esp. in a tendentious way in order to encourage or instil a particular attitude or response. Also, the ideas, doctrines, etc., disseminated thus; the vehicle of such propagation." (from Oxford English Dictionary, 2nd ed., 1989)

Political campaign speeches and party political statements are often, in reality, a form of propaganda. They fit this definition when they present the opposing point of view in an unfavorable light. All political organizations do this on a variety of issues.

- "President Bush has named a one-sided, misguided commission that has only one objective: to privatize Social Security. In so doing, he is ultimately risking the future of the program on which millions of Americans rely for retirement security, widow benefits, and disability payments. In fact, the only security about which this commission seems concerned is the security of the financial industries and special interests who stand to make millions if Social Security is privatized." -- from a press release dated May 3, 2001 from the Democratic National Committee (Read the entire release)

- "‘The American people would be better served if the misguided leadership of the Democratic Party were to lower the destructive rhetoric that drives people apart and engage more constructively in the process,’ said Ann Wagner, Co-Chair of the Republican National Committee." -- from a press release dated April 26, 2001 from the Republican National Committee

Identify the terms used to color the information reported in each press release above. Read the releases again, deleting those terms. How do the statements sound different?

When you read documents or listen to audio or video files that characterize opinions or positions in terms of their integrity or moral content, you may well be in the presence of propaganda. Remember, the purpose of propaganda is to 'instil a particular attitude': to encourage you to think a particular way. Think for yourself: base your opinion on the facts, not the hype.

Misinformation

"Nobody's perfect" is an excellent rule of thumb in most cases but a bad omen when you're looking for information on the no-editorial-control Internet.

Misinformation, noun
1. The action of misinforming or condition of being misinformed.

Misinformation differs from propaganda in that it always refers to something which is not true. It differs from disinformation in that it is "intention neutral": it isn't deliberate, it's just wrong or mistaken.

"It's going to require numerous IRA agents." -- George W. Bush commenting on Al Gore's tax plan, which he felt would lead to a larger Internal Revenue Service (IRS), and probably not a larger Irish Republican Army (IRA), in a campaign speech given at Greensboro, N.C., Oct. 10, 2000 (Read the Salon archive of "Bushisms")

One of the most popular forms of misinformation on the Internet, especially e-mail, is the passing along of urban legends. Urban legends are fabricated or untrue stories that are passed along by sincere people who believe them...and feel the need to "inform" others.

"If you’re driving after dark and see an oncoming car with its headlights turned off, DO NOT flash your lights at them. This apparently is a new common gang member initiation "game" that goes like this: the new member being initiated drives along with no headlights on and the first car to flash their headlights at him is now his "target." He is now required to chase that car and shoot at or into the car in order to complete his initiation requirement." -- a recurring legend that has appeared in print, fax campaign, and through e-mail

Misinformation is perhaps the most difficult information lookalike to diagnose. Why? What strategies could you develop to determine whether what you are reading constitutes information or misinformation?

Urban legends, unlike Mr. Bush's acronym problem, sometimes begin in malice. They become misinformation when they are repeated by sincerely misguided people. Everybody makes mistakes...check the validity of everything you read before you put your belief in it and use it.

Disinformation

You have now reached the lowest of the low. Never underestimate the evil intentions of some individuals or institutions to say or write whatever suits a particular purpose, even when it requires deliberate fabrication.

Disinformation, noun
"The dissemination of deliberately false information, esp. when supplied by a government or its agent to a foreign power or to the media, with the intention of influencing the policies or opinions of those who receive it; false information so supplied." (from Oxford English Dictionary, 2nd ed., 1989)

One of the most notorious uses of disinformation was the dissemination of anti-Semitic speeches and writing by the Nazi party in Germany during the 1930s and 1940s. Unfortunately, disinformation didn't end with World War II. In fact, the Internet is an excellent vehicle for disinformation.

- Read Statement by His Excellency U Mya Than, Permanent Representative and Leader of the Myanmar Observer Delegation to the 57th session of the Commission on Human Rights at the United Nations website
- Read 2000 Human Rights Watch report on Burma [Myanmar] at the Human Rights Watch Website. Human Rights Watch is an independent, nongovernmental organization that issues country reports on the human rights situations in countries around the world on a regular basis.
These two reports represent very different viewpoints on human rights in Tunisia. One of them may well represent disinformation. One good starting point in determining whether or not a document may constitute disinformation is to find out who owns the document or domain and then find out what that individual or group's mission or beliefs are. Ask yourself what the document owner has to gain by circulating the document.

Always validate or confirm information on individuals, institutions or groups, and countries that you find on the Internet. If you don't know who wrote what you read or why they wrote it, you don't know if it's trustworthy.

---

Evaluating Internet Resources

Introduction

This document will give you a variety of ways to look for each kind of information. Always remember that there are other, nonelectronic, methods of getting much of the information discussed in this document. Visit the library and ask a librarian for help. Of the five evaluative criteria listed in Evaluating Information found on the Internet, three may be investigated by electronic means:

- Authorship
- Publishing body
- Currency (of the document itself)


Authorship

What do we need to know about the author? When the author is someone unknown to you, ask the following questions:

1. Is the document signed?
2. Can I get more information on the author by linking from this page to other documents?
3. Was there information about the author on the page from which I linked to this one?

If you can answer "Yes" to the second or third question, it's possible that you will have enough information for evaluative purposes, or at least enough information to help you find the author's telephone number or e-mail address so that you may contact him or her with questions.

If you can answer "Yes" to the first question only, you may need to find further information on the author. There are a number of ways in which you might do this:

1. Go to the home page of the web site where the document lives and search for the author's name using any available internal search engine or directory (works best for academic web sites). This may help establish affiliation.
2. Try searching the author's name, enclosed in quotation marks, in Altavista. This may lead to other information on or pages by the same author.
3. Try using an e-mail address finder. A compendium of finding aids and their comparative value may be found at FAQ: How to find people's E-mail addresses by David Alex Lamb, Department of Computing and Information Science, Queen's University, Ontario, Canada. This is what football fans refer to as a Hail Mary: you're tossing a search into the unknown in the hope that it finds someone.

These are not ideal means for finding someone's credentials, even though they may help reveal someone's identity. Evaluating information usually consists of weighing a number of criteria together, so you will need to assess how important authorship is on a case-by-case basis.

If no information on the author can be found, or there is no signature or attribution on the page itself, go directly to...

Publishing body

OK, so where are we, in the geography of cyberspace? Look at the web page you are trying to evaluate. Does it include any of the following?

1. A header or footer that shows its affiliation as part of a larger web site.
2. A watermark or wallpaper that fulfills the same function.
3. A link at the bottom or top of the page that allows you to go to the home page of the web site where the document lives.
4. A link that allows you to send a message to the site Webmaster.

These features help you judge the "official character" of a web page. They act as an assurance that the page you are evaluating functions within some type of institutional setting. Judging the official nature of a web page is extremely important if the page is not signed. Some web sites do not include attributions to individual authors, so you will have to rely on your ability to evaluate the institution, or domain, where the page lives. Caveat: while the Web is looking better, especially official sites maintained at educational institutions or by scholarly societies, not everyone has caught up with the importance of consistent graphics or "return" links. So you may be looking at a perfectly good page that hasn't got any visual clues to its affiliation. Move on to the next step if this is the case...

If your page gives no clues as to its identity, you will need to focus on the URL, or address.

1. Can you find the web site's home page by deleting all the information in the URL after the server name?
2. Can you tell if the page is actually part of someone's personal account, as opposed to being part of an official site? Click for help in evaluating this kind of URL.

3. If all else fails, can you find information on the server or domain? Try using BetterWhois to get the name of the owner. Type the name of the server (e.g., milton.mse.jhu.edu) in the box and press Enter. BetterWhois should find the identity of a server, as well as contact names and telephone numbers.

Once you find the name of the organization owning the server, you may have enough information to judge its reputation as an information source. Remember, this is only of value for official pages from a web site. If the page you are evaluating comes from someone's personal account, you really have no idea what their place is within the organization, or if they are in a position to represent the organization. If you are not familiar with the organization, try one of the following:

1. If it is an association of some kind, look for it at The Scholarly Societies Project. Is it represented?
2. For all others, search the name of the organization, enclosed in quotation marks, in Altavista. Does anyone else have information on it?

Why might an author or publishing body want to remain anonymous? Read Information and its Counterfeits: Propaganda, Misinformation and Disinformation to learn why...and to understand why you need to know where "information" comes from.

If you cannot ascertain either the author or publisher of the page you are trying to evaluate, you are looking at information that is as anonymous as a page torn out of a book. You cannot evaluate what you cannot verify. It is unwise to use information of this nature. Look for another source.

Currency of the Document

Even when you can find information on the author and/or publisher of a web page, you should still consider how "fresh" or "dusty" the document is. This is especially critical if the document discusses time-sensitive information, such as census information or other statistics. Look for internal confirmation of the information:

1. Does it use a caption such as "Based on 1990 US Census data"?
2. Does it include information within the document such as "Closing stock prices, September 30, 1996"?
3. Is the statistical source listed in a bibliography to the page?

If you cannot ascertain where the statistical information comes from, or what its age is, you are once again looking at anonymous information.

It is also valuable to know when a page was last updated. Has it been "pruned" or "dusted" lately, or has it been sitting on the shelf?

1. Look at the bottom of the page. Does it have a "last updated" date?
2. Use the "Document info" feature in the "View" menu on Netscape. Does it tell you the date?
3. Change the URL by backing up to the last slash (/) in the address. This may allow you to see the details of the directory or subdirectory of the server including your page. Usually the last modification date is included.

It's valuable to know the age and updatedness of your page because you may be looking at an orphaned or superseded document that has been replaced by other information.

Always remember that the best counterfeit looks the most like the real thing. How genuine and trustworthy is your information?

Evaluating Social Media

Accuracy Checklist

Social media can provide instant news faster than traditional news outlets or sources and can be a great wealth of information, but there is also an ever increasing need to verify and determine accuracy of this information. Here are some items to consider that can help determine authenticity:

- **Location of the source** - are they in the place they are tweeting or posting about?
- **Network** - who is in their network and who follows them? Do I know this account?
- **Content** - Can the information be corroborated from other sources?
- **Contextual updates** - Do they usually post or tweet on this topic? If so, what did past or updated posts say? Do they fill in more details?
- **Age** - What is the age of the account in question? Be wary of recently created accounts.
- **Reliability** - Is the source of information reliable?

Tips for Citing Social Media

Check the manuals for your style guide for the most up-to-date information. If your style guide doesn't cover it in print, then look to their online website to see if they've included some information there on proper citation for these social media formats. Also try our citing guide for more information or email your librarian.
APA Suggestions for citing twitter on Face Book

MLA Style for Citing Tweets
http://www.mla.org/style/handbook_faq/cite_a_tweet

Chicago Style for tweets

Introduction
According to Twitter, the site has more than 140 million active users, with over 340 million Tweets per day. This means that one billion tweets are sent every three days. More and more people get news and information from news media, but it’s important to remember that fast does not always mean accurate.

How to identify credible information on social media can be challenging. Rumors and misinformation can spread quickly through social media outlets such as Twitter or Facebook. Some of the criteria used to evaluate Internet sources, such as being skeptical, asking questions, looking at the quality of the source of the information, still apply in social media. At the same time, a new and quicker way to exchange information, without some of the clues of authorship provided by more traditional online sources, means we can add some new techniques and approaches to evaluating information.

Storyful recently added four case studies that show how they verified online videos and other content for clients. From their blog:

“At Storyful, we interrogate content shared on the social web in a style not dissimilar Barber’s grilling of dignitaries of media and politics. We adopt a natural skepticism to every item of content we discover. Verification is a cornerstone of our work and it has to be. Information and content often spreads across social media in 'Chinese whispers' fashion. Videos and images are spliced, diced and re-posted. Context and details change, agendas compete. Falsehoods and fabrications are deliberately issued.

Within an hour of the Pacific tsunami alert being issued on April 11, 2012, Twitter and YouTube abounded with videos purporting to show monster waves striking the coast of Sumatra and Aceh in Indonesia. However, these were versions of the devastating 2004 tsunami and other dramatic videos, reissued with the 2012 date. In December 2011, when a police officer was killed at Virginia Tech in the US, a picture of the 2007 massacre was widely circulated as the 2011 event. In both instances, Storyful was quickly able to debunk this content as false.”

Verifying information is important regardless of the type of social media outlet. Twitter is not immune to fake accounts. Recently, a manager from Google created a bot that attracted followers on Twitter. When the bot posted about a running injury, followers expressed concern over her injury. Wired initially reported on this story and it seems that the bot account has since been suspended. In an experiment conducted late last year, researchers created nine Twitter bots that were able to attract, on average, 62 followers each over a three-week period. This rise in the use of bots raises questions about how easy it might be for someone to try to influence news via social media outlets.

If you are unsure whether a social media post is from a person or bot, do not use the source.

Verifying Information from Social Media
Paul Bradshaw, a leading digital media expert and teacher in Europe, writes the Online Journalism Blog. In 2011, he wrote a post that provides a variety of basic guidelines about online verification with a section related to social media:

How long has the account existed? If it’s only existed since a relevant story broke (e.g. Jan Moir’s column; an earthquake where someone claims to be a witness) then it’s likely to be opportunistic.

Who did the person first ‘follow’ or ‘friend’? These should be personal contacts, or fit the type of person you’re dealing with. If their first follow is ReadWriteWeb, then it may be that you’re not actually dealing with a Daily Mail columnist.

Who first followed them? Likewise, it should be their friends and colleagues.

Who has spoken to them online? Ditto.

Who has spoken about them? Here you may find friends and colleagues, but also people who have rumbled them. But don’t take anyone else’s word for their existence unless you can verify them too.

Can you correlate this account with others? The Firefox extension Identify is a useful tool here: it suggests related social network accounts which you can then try to cross-reference. For companies the Chrome extension Polaris Insights does something similar for companies.