

INTERNET-BASED RESEARCH: THEORY AND APPLICATION

MVZ507

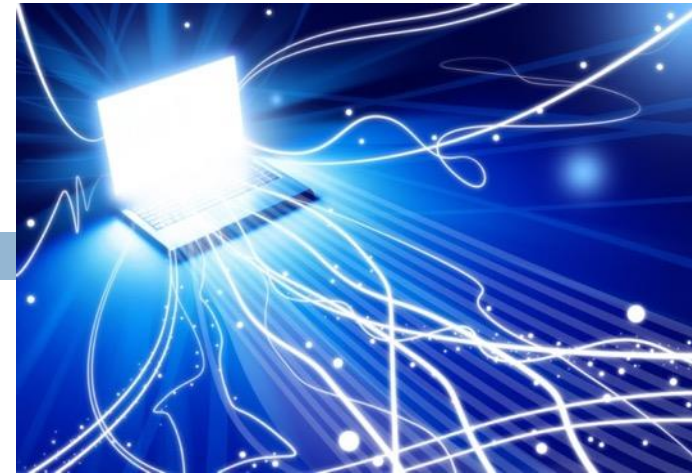
SPRING 2019

Session 2



Maya Hadar

Agenda For Today



- 1. Introduction: Research and the internet**
- 2. Researching using the internet: things you should know**
- 3. The Internet as a source of quality research material**
- 4. Searching, Evaluating and citing web sources**
- 5. The Wikipedia controversy**
- 6. Internet Polling**

A Brief history of Internet mediated research (IMR)

- Pioneers implementing **online data collection methods** around the mid-1990s *Dillman (1991), Hewson (1994), Bordia (1996); Gaiser (1997); Reips (1997)*
- Reports (e.g. validation studies) have since **supported the quality of data** that can be obtained in **IMR** (*Corley & Scheepers (2002); Hewson & Charlton (2005)*)
- **A new era of IMR-** The web as organic, collaborative, interactive, ever-changing (e.g. Wikipedia); the rise of social media (e.g. Facebook, Twitter).

Research and the Internet



- The Internet as a **source of quality material** to be used in your academic writing
- The internet as a **tool for research**
 - Obtrusive Approach (experiments, surveys, interviews)
 - Unobtrusive Approach (observation using traces and archives, e.g. blogs, social networks)
- **'Big Data' & indices** - large data sets
 - Potentially difficult to manage
 - Sometimes hard to read/understand



Research and the Internet



- The **Internet's transformative impact**: makes research easier, but also has some potential dangers
- **Explosion of information** makes it even more necessary to have a good analytical framework to handle it, otherwise simply **information overload**
- Theoretical frameworks become more, rather than less important: need to develop a **critical attitude** towards **information obtained via the web**.

Using the Internet for Academic Research

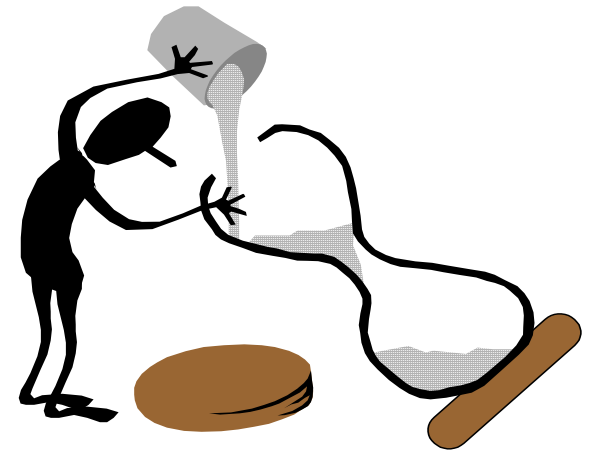
Things you should know...

- Not everything is available on the Internet => massive amount of info, but **subject dependent**
 - What is available may not be as **appropriate** as the information in other sources
 - The Internet is **only** one research tool, provides access to only **some** of the many sources of information available to you
- Information on the Internet is **not stable**; At any time, information may be moved, altered, or deleted.

Using the Internet for Academic Research

Things you should know...

- Research on the Internet will **take time** (all research does)
- **Not everything on the Internet is accurate**, true, current, or reliable



What is a quality online resource?



- **A resource fit for purpose** –in academic settings that might mean peer-reviewed or just credible – e.g. information found on the British Library website
- In some **circumstances** could be **informal sites**, news sites/sites expressing an opinion
- Still, we should take into account some **criteria** which can be **universally adopted to evaluate** whether sites are worth using.

What is a quality online resource?



- Fit for Purpose – things to look for:
 - Original content?
 - Statement of intention/intended audience
 - Title/author/ date – minimum needed for citation
 - Site map- gives an overview- helpful

Using the Internet for Academic Research

Finding Information online:

- There are two major ways to begin a search on a research topic: by **subject** or by **keyword**

Searching by Subject:

- Use an **academic subject directory** => portals organized by librarians/other academics providing a collection of links to sites that are appropriate for academic research



Useful Subject directories

- Political Studies Association of the UK, highly rated gateway site
<http://www.psa.ac.uk/>
- Entry for each country with information on economy and politics
<https://www.cia.gov/cia/publications/factbook/>
- British National Archives
<http://www.nationalarchives.gov.uk/default.htm>
- News outlets (e.g., Thompson, Reuters)



Using the Internet for Academic Research

Finding Information online:

Searching by Keyword:

- A keyword search may be more appropriate for a very **specific topic**
- Use **Internet search engines** to do a keyword search
- There is **no standard** or controlled vocabulary for finding information on the Internet. This means you will have to think of **synonyms, variants in spelling, different word endings, etc.**



Using the Internet for Academic Research

Finding Information online:

- **Google** is one of the best Internet **search engines**; It displays the search term in context and has an good results ranking system
 - Google has "**Basic**" and "**Advanced**" search modes
- **What's wrong with Google?**
 - Too many hits, too unfocussed, sorting results by 'popularity', some types of site e.g. news sites are favoured, not always helpful for academic research
- **Consider also using other search engines** (subject-specific search engines, Bing, etc.).

Using the Internet for Academic Research

“Basic” Internet searches using Google:

- **Phrase searching:** Use quotation marks for words that should be found together in that order (e.g. "electoral reform")
- **Multiple terms:** Boolean "**and**" is stated as the **automatic default**, so entering two or more words should retrieve pages containing all of the terms you enter. (e.g. elections reform)
 - If not all terms are found, results without them will be displayed
- To **force** retrieval of results for all keywords enter each word with a **plus sign (+)** directly before it (e.g. +elections +reform).

Using the Internet for Academic Research

"Basic" Internet searches using Google:

- **Narrowing a search:** Enter more search terms to specify more clearly what you want to find (e.g. elections reform canada)
- **Broadening a search:** Use "**OR**" to search for alternative terms at the same time (e.g. "election reform" OR "electoral reform")
- **Truncation/stemming*** (a search technique which refers to the ability to search just a portion of a word) is not available on Google.

*wom!n = woman, women
colo?r = color, colour



Using the Internet for Academic Research

“Advanced” searches using Google (https://www.google.com/advanced_search)

- See the Google Advanced Search page for more ways to search efficiently (language, file type, etc.)

| Find pages with... | | To do this in the search box. |
|----------------------------|--|---|
| all these words: | <input type="text"/> | Type the important words: tri-colour rat terrier |
| this exact word or phrase: | <input type="text"/> | Put exact words in quotes: "rat terrier" |
| any of these words: | <input type="text"/> | Type OR between all the words you want: miniature OR standard |
| none of these words: | <input type="text"/> | Put a minus sign just before words that you don't want: -rodent, -"Jack Russell" |
| numbers ranging from: | <input type="text"/> to <input type="text"/> | Put two full stops between the numbers and add a unit of measurement: 10..35 kg, £300..£500, 2010..2011 |

Then narrow your results by...

| | | |
|------------------|---|---|
| language: | <input type="text" value="any language"/> | Find pages in the language that you select. |
| region: | <input type="text" value="any region"/> | Find pages published in a particular region. |
| last update: | <input type="text" value="anytime"/> | Find pages updated within the time that you specify. |
| site or domain: | <input type="text"/> | Search one site (like wikipedia.org) or limit your results to a domain like .edu, .org or .gov |
| terms appearing: | <input type="text" value="anywhere in the page"/> | Search for terms in the whole page, page title or web address, or links to the page you're looking for. |
| SafeSearch: | <input type="text" value="Show most relevant results"/> | Tell SafeSearch whether to filter sexually explicit content. |
| file type: | <input type="text" value="any format"/> | Find pages in the format that you prefer. |
| usage rights: | <input type="text" value="not filtered by licence"/> | Find pages that you are free to use yourself. |

Google Scholar

Articles Case law

Stand on the shoulders of giants

Google scholar Pros =>

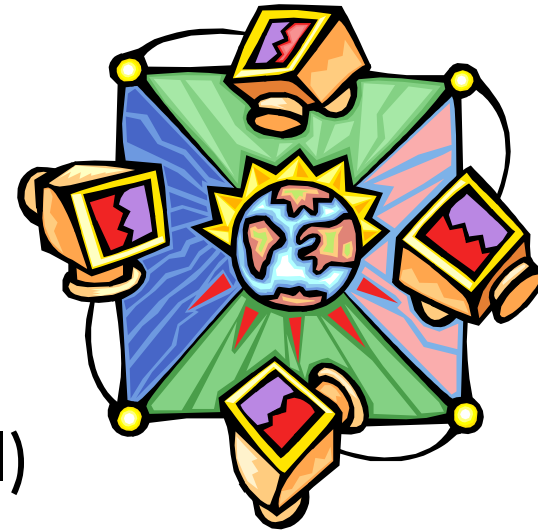
- Useful, focussing on academic websites
- May often pick up book chapters which can be harder to find
- Can download to reference manager software/easy to cite from
- Can often link in to full text

Google Scholar Cons =>

- Can be inaccurate (picks up items from bibliographies)
- Often randomness about results
- Won't go into all institutions/databases/repositories- need subscriptions
- Older items often listed first as more often linked to
- Older items may not be available (just citations)

Specialist Search Engines – Invisible Web

- **Invisible web** – items that are not found by conventional search engines because:
 - Exist within sites that have been blocked by owners
 - Spiders (bots) cannot penetrate them (no/restricted access)
 - Are not linked to by other sites so not found by spider
- **Institutional repositories, databases** (can be commercial)
 - OAlster
 - Solo (only searching Oxford University resources)



Why do we need to evaluate online info?

- It is **always** important to evaluate the information you intend to use for a research paper => even when using printed books and articles found in a library, but even more so for information found on the Internet
- **Quality** in printed resources is often assured by editors and publishers who pay the costs of publishing, and by libraries that select what books to buy
 - Unlike most print sources, **web sources do not have to be professionally accepted** and edited to be published (e.g. Wikipedia);
 - Some web sites have strict editorial policies; some have none at all

Why do we need to evaluate online info?

- **Reliability-** Virtually any person can publish almost **anything** on the Internet at any time
 - Content can seem to be **objective** but is **biased/commercial**
 - Sometimes the actual **purpose** of the web site may **not** be **clearly articulated**
 - **Personal sites** are used to express individual **opinions, not facts**
 - Similar to 'quality'- web sources **do not have to be reliable-** no/little control Wikipedia

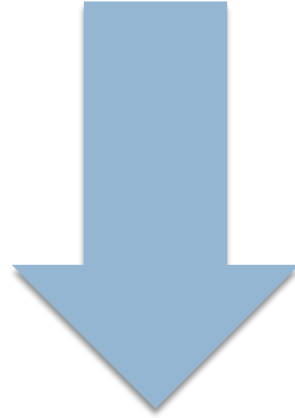


Why do we need to evaluate online info?

- **Appropriateness-** A basic keyword search on a search engine will find various sources (reliable/non reliable/commercial sites, etc.)
- It is, thus, up to you to determine which are appropriate.



Why do we need to evaluate online info?



- Need for a general attitude of **critical skepticism**
- Check the **entire document** you are viewing
- Check all the "**meta-data**" available, ie. all **clues** you can find that put the information in context or provide details about it.

Evaluation Criteria



Things to look for when evaluating information on the Internet:

- **Identify the web site =>**
 - Informative pages
 - Personal web pages
 - Political/interest group pages
 - Marketing-oriented / “infomercial” pages
 - Academic material

Evaluation Criteria



Things to look for when evaluating information on the Internet:

- **Host site** or "Publisher" =>
 - Web addresses often indicate the **country of origin** (.ca = canada, .fr = france), or the **type of organization** hosting the web site (.edu=educational, .com=commercial, .gov=governmental, .org=organization)
 - You may have to **back up to the home page** to find out more about the web site on which a document is found and who is responsible for it
 - If the information at the site is not original, **locate the the original source** and make sure you cite it properly.

Evaluation Criteria



Things to look for when evaluating information on the Internet:

- **Type of Information =>**
 - Many **different kinds of information** resources can be found on the Internet:
 - Peer-reviewed journal articles and books, government documents, professional working papers, student essays
 - Personal letters, fiction, spoofs of serious research
 - In print these are usually easy to **distinguish-** not online.

Evaluating Criteria



Things to look for when evaluating information on the Internet:

- **Authority**=> who wrote it, could be a person or an organization
 - Who are they?
 - What is their background or expertise?
 - Why should they be trusted to know about the field? What are their credentials?
 - Are they affiliated with an institution or university?
 - What is their bias or point of view?
 - Is it peer reviewed? Who else thinks this is good?
 - Other publications/references by the author
 - 'About ' section
 - Institutions' home page

Evaluating Criteria



Things to look for when evaluating information on the Internet:

- **Date** \Rightarrow Most, if not all, **information** is only **relevant** in a **context of time**
 - If **no date** is given: suspicious
 - There may be an **original creation date** or/and a date for when the information was last **modified** (often small print at the bottom)
 - **Test** a few **links**
 - Look for **dates** in the **references (bibliography)**
 - References to **current events?** (main text)

Evaluation Criteria



Things to look for when evaluating information on the Internet:

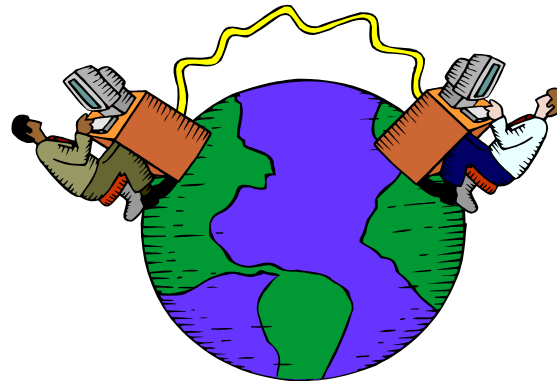
- **Depth and scope of the information presented =>**
 - Does the material **look like an academic research** (references? hyperlinks? footnotes?)
 - Does the author consider **opposing** points of view?
 - How closely does the site **match** the information for which you are searching?
 - **Corroborate information** whenever possible.



Why do we need to evaluate online info?

Things to look for when evaluating information on the Internet:

- **Purpose** => **Commercial uses** of the Internet are growing faster than any other, and much of the "information" on the web is advertising
 - The Internet is also a very effective **propaganda tool**
 - **Be aware of the purpose** of the site and of the document, you are viewing (hidden?)



Citing Online Information



Using and Citing Internet Sources:

- It is a good idea to check the suitable **style guide/s** BEFORE starting your research, so that you know what information to include in your footnotes or bibliography for all sources you retrieve from the Internet
 - The standard citation manuals include instructions on how to cite electronic sources in the body of your paper and in the bibliography
- **Information on the Internet is protected** by copyright unless specifically stated otherwise. Be sure to **cite all information** used for your paper.

The Wikipedia Controversy



- **Strength:** free, up to date, large range of topics, scores highly on Google, various languages
- **Credibility:** there is no peer review process
- Entries may deteriorate
- **Unbalanced** – entry on McDonald's critical but very little about history, problematic concerning “sensitive” issues (e.g., Israeli-Palestinian conflict)
- Gropes towards a **consensus**, but that it's not the same as validated knowledge
- A **rough guide to knowledge**, useful for a first/quick overview of a subject
- Do not use in academic contexts.

New forms of Politics



- Highly individualised => People create *their own political home* by framing their own ideas and channels for political action
- Easy to mobilize/manipulate people using social media
- Sharing information easily, globally => fake news?
- Micheletti argues that **geographical closeness** can be replaced by **Internet interactions**

New forms of Politics



- **Challenge:**
 - Traditional conceptions of politics as a process of **debate**, where our views are **modified** as a result of **argumentation**, **learning** about the views of others
 - **Highly dependent on the information** presented to us/we seek/ we find
 - **Feasible online?**

Internet Polling



You Gov (<https://today.yougov.com/>)

- A community of 6M people around the world who share their political views. The information is being analysed and posted online

A screenshot of the YouGov website homepage. The background is a dark purple color. In the top left corner is the YouGov logo. In the top right corner, there are navigation links: Ratings, Topics, Solutions, Join, and Company. The main heading is "What the world thinks" in large white text. Below it is a sub-heading: "YouGov is a global public opinion and data company". In the center, there is a white search bar with the placeholder text "Search for music artists, politicians, brands and more" and a magnifying glass icon on the right. At the bottom, there is a line of text: "Explore the popularity and fame of anything and everything with [YouGov Ratings](#)".

YouGov Ratings Topics Solutions Join Company

What the world thinks

YouGov is a global public opinion and data company

Search for music artists, politicians, brands and more

Explore the popularity and fame of anything and everything with [YouGov Ratings](#)

Internet Polling



Advantages:

- **Cheap**
- **Quick**, although probably no quicker than telephone polling
- People more **likely to be honest**, especially on sensitive issues
- People may give more **thoughtful answers** than when stressed by an interviewer with time constraints
- Makes it **easier to reach specific** population, e.g., high income – difficult door to door, phone, streets
- **No problem of interviewer bias**
- Respondent **convenience**.

Internet Polling



Drawbacks:

- **Sampling bias** => Not everyone in the population have computers and are connected to the internet. Are those on You Gov database typical?
- More men, relatively **few elderly** (You Gov has to weight elderly x 6)

You Gov's defence:

- Biases are not huge
- Have a very large base of respondents to choose from
- Got the some election results and Pop Idol results right
- Telephone response rates falling.

Internet Polling



Conclusion:

- Internet polling is here to stay
- Good for opinion/consumer polls
- If proper, we may cite in academic context (subject dependent)

In Conclusion



- Before searching the Internet consider...
 - **Strategy** – think through what sort of information you want and where you might expect to find it
 - Make sure you are clear and focussed about your research area – time wasting is very easy on the Internet
 - **Suitability of search engine/source** – Google? Google scholar? Specific website? a bibliographic database?

In Conclusion



- Before searching the Internet consider...
 - **Search engine as a starting point-** once you've found a lead, systematically browse specific journals/authors
 - Use a **range of Internet tools** (gateways and specialist search engines)

In Conclusion



Evaluation Criteria: Key considerations =>

- **Currency** – when was this written? When updated? Are the links live?
- **Reliability/Verifiability** – Are there references to other credible sites/publications? Are statements backed up? Who is allowed to edit/add to the site?

In Conclusion



Evaluation Criteria: Key considerations =>

- **Standpoint** – what's the purpose of the site? Why has the author written it? Are there obviously unbalanced arguments?
- **Fit for purpose** – Is the information of an appropriate breadth and depth for your purposes? Who is the intended audience? Does the information add to information obtained from other sources?

Next Session...



42

- Reading week- no class!



Thank You For Your Attention!

Questions???