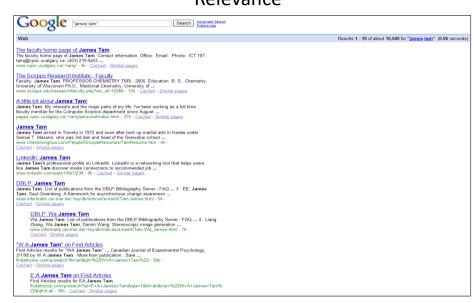
#### The Internet

You would learn how to effectively search for information online as well as how to evaluate the quality of online sources.

# Search Engines Rank Results According To Relevance



#### Making A Site More Noticeable (Higher Rank)

- Search database built via search spiders (e.g., Google):
  - -Add relevant keywords to your page.
  - The frequency and location of keywords may play a role in determining relevance.
  - Links (to your site) also affect ranking
  - —Trying to artificially rank your page higher in search results may result in a page being put at the bottom of the list e.g., adding celebrity 'gossip' to my CPSC page may rank it higher but not relevant info for most visitors.
- Search databases built via human researchers (e.g., Yahoo):
  - Make sure that your site is examined by the people who build the database.

# Searching For Information On The Internet: Google

• Easy? Just type in what you're searching for...



...you may get the information that you were looking for plus much more!



You Will Learn Strategies For Narrowing Your Search Results (For Google But Largely Applies To Others):

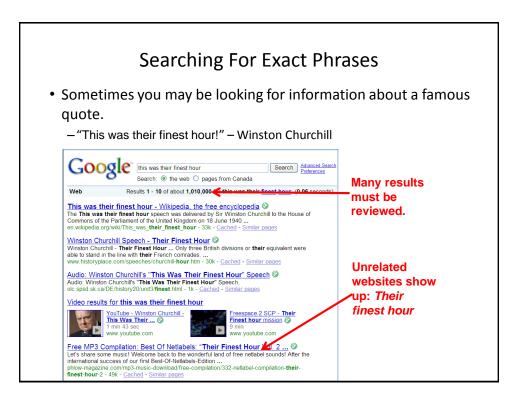
- 1. Reducing the number of unrelated results.
- 2. Explicitly avoiding pages with certain words.
- 3. Searching for information from select pages.

Source (last accessed 2009): http://www.google.com/support/websearch and http://www.googleguide.com

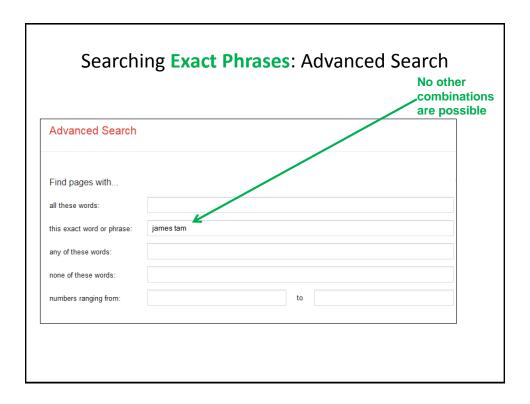
### **Employing These Search Strategies**

- Many (but not all) of the features can be found through an "Advanced Search" subpage of the Google site.
  - -http://www.google.com/advanced\_search
- This leads you to a page with several options (to be covered shortly):









## Words That Are Commonly Ignored By Google

- Stop words are ignored by search engines such as Google:
  - -Common words
  - -Reserved words
- The search engine can be forced to include the stop words:
  - Use quotes
     best places to eat in Calgary
     "best places to eat in Calgary"

2. Use the 'plus' operator.

١/-

. . .

Star Wars +I

Star Wars I

#### What If More Than One Word Can Be Used?

- Some concepts can be represented by using different words.
- The '~' operator includes synonyms in the search.
- Example:
  - ~cheap computers

## The Search Criteria Is Incomplete/Partially Unknown

• For example, searching for information on this person:



Image: Koch/Munich Security Conference



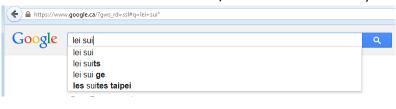
Image: www.cnn.com (2015)

- How is his name spelled?
  - -Arnold Scwartzengger?
  - -Arnold Schwartzenger?
  - -Arnald Scwartzencker?
- · Note: many search engines try to 'guess' what you are thinking but sometimes software guesses wrong (think of "auto correct" on mobiles)
- Also handy for less common words and phrases e.g., technical terms, foreign FYI it's "Arnold Schwarzenegger" languages translated to English



#### Using The Wildcard

- Search engines such as Google have been refined to reduce but not entirely remove the need for wildcards.
- Google may not provide useful alternatives when searching for technical terms or foreign words translated into English
- Example: "Lei Sui, Lung" (English translation of the Chinese name of the famous martial artist/movie star: Bruce Lee)



# **Searching A Range**

- Searching numerical values within a certain min max range
- Range operator .. (multiple dots)
- Example:

Qatar History 2000..2009 computer 2000 QR..5000 QR

Find pages with		
all these words:	qatar history	
this exact word or phrase:		Numerical
any of these words:		range
none of these words:		
numbers ranging from:	2000 to 2009	9

### **Searching Among Alternatives**

- Normally when a word is included in the search box Google will try to find web pages that include all those words.
- Example:

```
–(All words – Logical "AND")cute wallpapers cats dogsVs.cute wallpapers cat OR dogs
```

#### Example:

```
"Bruce Lee" OR "Little Dragon" OR "Lee Siu Lung"
"Wayne Gretzky" OR "The Great One" OR Number 99" Or "Number ninety nine"
```

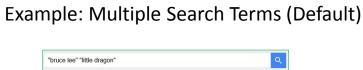
### Searching Among Alternatives (2)

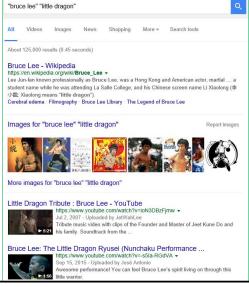
Note: Google is case sensitive in this situation! (OR must be upper case in order to search for alternatives).

To be or not to be

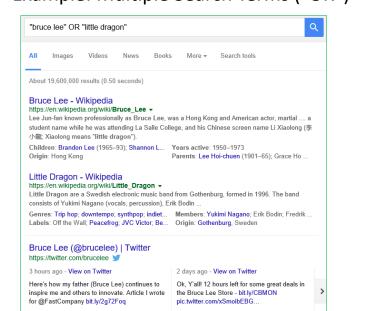
Vs.

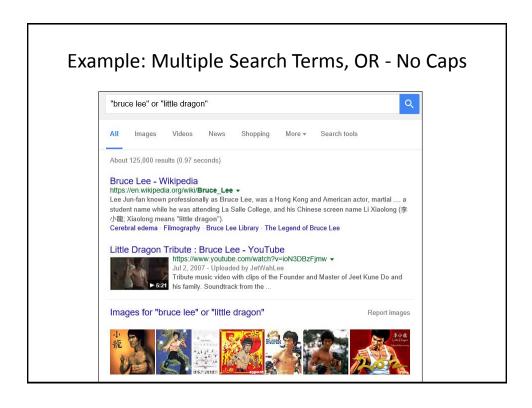
To be **OR** not to be

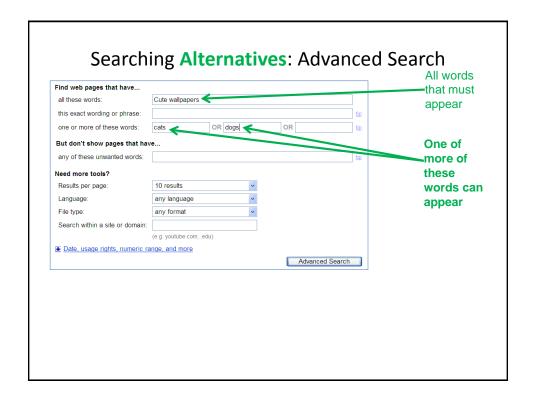




# Example: Multiple Search Terms ("OR")







## Using 'OR': Exceptions<sup>1</sup>

- Google was designed to approximate how people think and behave when searching for information.
- Consequently exceptions to the rules are sometimes made.
- Example<sup>2</sup>:

What you tell the search engine: "For better" **OR** "for worse" What the search engine looks for: "For better **or** for worse"

1 From <a href="http://www.google.com/support/websearch">http://www.google.com/support/websearch</a>

2 Google and other search engines may automatically assume you are looking for the Lynn Johnson comic strip (https://www.fborfw.com/) when you enter this phrase.

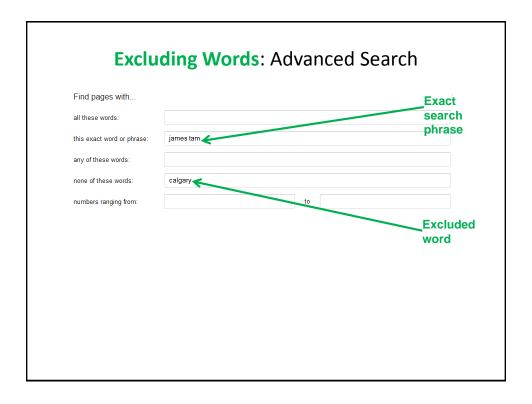
### **Excluding Words**

- There may be times when you want Google to exclude sites with certain words or phrases.
- This can be done with the **subtraction operator** (subtract the words that follow the operator from search results).
- Example:

"James Tam"

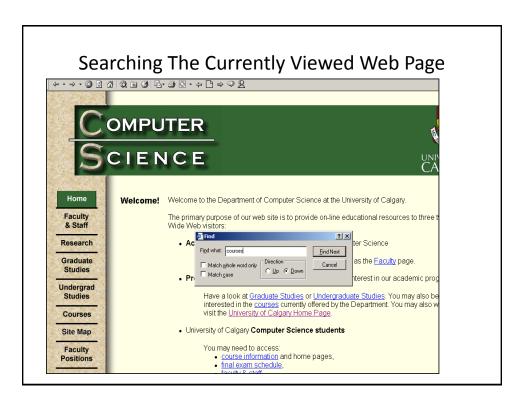
Vs.

"James Tam" -calgary

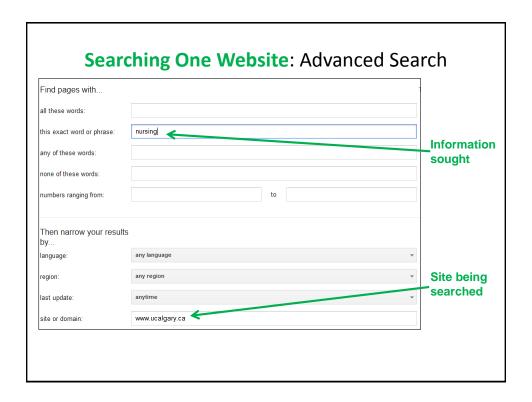


## Site Specific Searches

- Useful when a webpage is large and/or not well organized:
  - -Searching the current webpage
  - -Searching the entire site (and only that site)







### Finding And Evaluating Information

- This section is crafted for students who need to conduct research for an undergraduate course e.g., writing a paper
  - But the principles can be applied when researching information for personal reasons because it may improve the quality of the results (e.g., "Are there technical reasons why a MAC is any more or less secure than a Windows-PC?")
- Finding: Where to look for information
- Evaluating: Determining the quality of information?

#### Finding Information

- When conducting formal research (or even research for undergraduate courses) information sources should be evaluated for quality
  - -Not all sources of information are equal in terms quality.
  - Contrast: a private individual's personal blog vs. a website containing the current research being conducted by the experts in a particular field.
- Where to start?
  - -Check with your course instructor: what level of research is expected?
  - -Journals and conference proceedings
  - -Books
  - -Online: websites

### **Journals And Conference Proceedings**

- Publications that describe the latest theories and research produced by domain experts
- The quality of publications can vary
  - First tier "prestigious" venues only containing the best or higher quality work publications:
  - -Second tier "average" publications:
  - However in terms of undergraduate research the quality of papers produced at either category should be sufficient (check with your course instructor)
  - Which ones are at least reasonably good ones? Start with the instructor's publication list and look at the journals and conferences in which this person has published.
- Overall the quality of the information is generally good
- Since it is written by 'experts' for 'experts' it may be difficult for the typical student to read and understand

## Journals And Conference Proceedings (2)

- Examples (in case you're curious):
  - <a href="http://www.sigchi.org/conferences/">http://www.sigchi.org/conferences/</a> "User-friendly technology"
  - <a href="http://cscw.acm.org/2015/">http://cscw.acm.org/2015/</a> "Technology to support groups"

#### **Books**

- Books: First look at the intended audience
  - –General population books: with no background knowledge in the domain?
    - e.g., "A brief history of time" by Stephen Hawking
    - Probably too rudimentary for use in post-secondary courses.
  - -Students: high school or post-secondary
    - May be at an appropriate level of detail (for the latter category) if you just want general knowledge of a topic (e.g., instead of taking a course)
    - Probably not specialized enough if you need to write a paper for a course
    - e.g., you are taking CPSC 203 there is a brief hardware introduction but the introduction is probably not sufficient if you need to write a paper on the specific hardware needed for high-end gaming computers

#### Online: Websites

- Even more than the other sources, websites vary greatly in quality
  - The publishers of conference may host websites describing the latest research conducted by the experts
  - On the other hand: anyone can make their own website about a particular topic
- Some things to keep in mind when evaluating the quality of a website:
  - –Who is the author:
    - Widely regarded expert in the field
    - Industry expert (which field, is it similar to the one described in the website)
    - Does the person have any apparent conflicts or agendas e.g., a CEO of a computer manufacturing company writing a review of laptops (sometimes affiliations are not so obvious)
    - Are points backed by facts? Are those facts citing reputable sources?
    - Some present opinions as facts.

#### Online: Websites (2)

- -Who is the intended audience of the website:
  - Expert researchers?
  - General population?
  - Combination? E.g., <u>www.mayoclinic.org</u>, <u>http://www.hopkinsmedicine.org</u>
- If the content is produced by a university or reputable research group then it's probably correct
  - -E.g., Mayo clinic, Johns Hopkins
  - But may or may not be targeted specifically to a 'general' audience and be of insufficient depth for your research
- In general use websites with a strong degree of a caution.

### Wikipedia

- "Wikipedia is a collaboratively edited, multilingual, free Internet encyclopedia..." – from www.wikipedia.org
- "Many eyes view it"
- Q: Is this a good quality source of information when researching a topic?

## Wikipedia (2)

- Many 'eyes' view but are those the 'right eyes'
  - -Some topics require expert knowledge e.g., 'Rocket science'
- There is no guarantee that the experts in the area will be viewing, let alone actively editing Wikipedia documents
- Furthermore the document that you view one day may not be the same one that you view on another day
  - –Actively study the history of changes of a document and compare differences?
- Subject to bias extreme cases of "vandalism" has lead to "lock downs"
  - Yes files can be secured from further edits but consider what's the point of a wiki (anyone on the web can edit)

#### Using Wikipedia

- Similar to regular paper encyclopedia or perhaps a newspaper article it's a distillation of different sources of information.
  - -Generally it is not accepted as a direct primary source (don't directly cite a Wikipedia article in your research paper).
- It may be a starting point to find good sources of information that can directly be cited.
- Example: "Human-Computer interaction"

Footnotes [edit]

1. ^ Card, Stuart K.; Thomas P. Moran; Allen Newell (July 1980). "The keystroke-level model for user performance time with interactive systems".

Communications of the ACM 23 (7): 396–410. doi:10.1145/358886.358895.g/.

2. ^ Carlisle, James H. (June 1976). "Evaluating the Impact of office automation on top management communication". Proceedings of the June 7–10, 1976, National Computer Conference and Exposition. pp. 611–616.

doi:10.1145/1499799.1499885.g/. "Use of 'human-computer interaction' appears in references"

3. ^ a \* De Hewett: Beacker, Card, Carey, Gasen, Mantei, Perlman, Strong, Verplank \*\*ACM\*\* SIGCHI Curricula for Human-Computer Interaction' g/. ACM\*\*

Wikipedia cites a good research conference (usability "user friendly" technology)

# Using The University Library



Services provided for students:

SIGCHI. Retrieved 15 July 2014. -

- -From <a href="http://library.ucalgary.ca/student-support">http://library.ucalgary.ca/student-support</a>)
  - "As a student attending the University of Calgary, you have access to the books, journals and online resources of a major research institution. Your subject librarians are available for research assistance as well.
  - Your University of Calgary <u>ONEcard</u> is your library card, and gives students the following privileges:
  - Borrowing books with a loan period of two weeks for undergraduate students and a term loan for graduate students
  - Interlibrary loan
  - Off-campus access to Library Research Databases"

## The University Library: Online Resources

- The publications of many journals and conference proceedings can be accessed "for free" (licensed) through the university portal:
  - Requires login with the UC login credentials.
  - -Combine it with an online web search:
- Use regular search engines to find the relevant paper.
  - E.g., "Direct Stimulation of Angiotensin II Type 2 Receptor Initiated After Stroke Ameliorates Ischemic Brain Damage" – found via Google scholar
- Then access the content of the paper through the university portal



#### Authentication Is Require To Access The Article



## Online Web Search Without The Library

- The general Internet user does not have a license to access many good online resources.
- After finding the publisher of the article: Journal of American Medical Association



### After This Section You Should Now Know

- How do search engines build up a database of websites using spiders
- Some factors that can determine the ranking of search results

### After This Section You Should Now Know (2)

- Techniques for more effective web searches:
  - -Searching for exact phrases: using quotes
  - -How to include "stop words" in searches: quotes and the plus operator +
  - -How to search for synonyms: ~
  - -How to use the wildcard: \*
  - -Benefits of the wildcard vs. relying on "suggested searches"
  - –Searching ranges: ..
  - -Searching for alternates: OR
  - -Excluding words and phrases: (subtraction operator)
  - -Conducting site specific searches:
    - Finding content within the webpage currently viewed: 'find in page'
    - Searching for content within a particular website: site

#### After This Section You Should Now Know (3)

- Sources for researching information (journals/conferences, books, websites)
  - -Strengths and weaknesses
  - -How to evaluated their value
- How a website such as Wikipedia can and should not be used when conducting research
- How the university library can complement a web search (licensed access to online resources)