Lecture #9 – Marketing the Website

Search Engine Google, Alta Vista, Hotbot	**Subject Directory** Yahoo, About.com, AOL
Both programs that enables the user to search Internet sites	
User types keywords – list of websites is returned where the keywords were found ("BEST/MOST APPROPRIATE" pages)	User chooses a choose subject of interest and follows one category, subcategory... until you reach the website
	Not automated; real people decide how to organize it

*85% of people find sites thru Search engines

*A new search engine that's getting a good share of the searching market: The Colbert Report

Search Engines vs. Directories vs. Metacrawlers: How info is organized?

Search Engine **Meta-search engine**	**Subject Directory** **Meta-crawler**
Uses software called WebSpiders, Webcrawlers	Webmaster submits website to Subject Directory
Gathers information: goes from web page to web page via links, looking at all the words on the page, building an index (database):	Human editors review web pages and decide if worthy then rank and organize them categorized lists
Index contains list of : • alphabetical list of words (keywords) it finds • where within the page the word was • the links (URL to the page) where it found the words (keywords	
Search engines searches the database of information	

Search Engine Ranking: meta tags/purpose

How does the search engine decide which pages to return to the searcher?

- Search engine uses index to decide which pages have the given keywords
- Every engine uses slightly different algorithms to decide the order of displaying the returned pages
- Google uses the "PageRank" algorithm as ONE of the factors to decide what order to present the pages it found to you

Google PageRank algorithm

- Algorithm gives each webpage returned from the key word search a weight between 0-1
- The higher the weight given to the page, the more likely it is that this page will be displayed first to you

How PageRank works

- PageRank evaluates two things:
  1. How many links there are to a web page from other pages
  2. What is the quality of the linking sites → what is more important
- Page Rank does not take into account the content of the page
- Page Rank ranks webpages NOT web sites

- Each inbound link is important in the overall total except for banned site, they don't count

SEE LECTURE 9 SLIDES # 22-23 FOR HOW TO CALCULATE PAGE RANKS!

SEO: Search Engine Optimization

- 3 Components you should look at when you are trying to make your site more visible are:
    1. Text
    2. Link
    3. Popularity

1.Text : Keywords

- Users will type a set of words or phrases into the search engine box to find what they want
- STEP 1: IDENTIFY THE CRUCIAL KEYWORDS of your webpage
- STEP 2: PUT THE KEYWORDS IN THE BEST LOCATIONS
- Title-Tag (Property Title) is considered the MOST important text by all search engines; **Title tag** is what is shown as the link for all returned results
- Body of the Webpage
    - *Keyword Prominence* - keywords found near the top of the page make the page more relevant and thus give those pages a higher ranking
    - *Keyword Density* also gives a higher ranking

2.Link Component

- Make sure that your links are easy to navigate, no broken links so web spiders can search through your content easily and index ALL your content
- **NOTE: It is hard for spiders to crawl through:**
    - Image maps
    - JavaScript
    - Database Driven Web Pages (ones with symbols in the URL)
    - Links in Flash documents
    - Poorly written html

3.Popularity

- Broken into 2 parts:
    - Link Popularity
      →Quantity: Getting many links to your site
      →Qualitynk: Popular websites that link to you
    - Click-through popularity:
      Measures:
      →How many times your site is clicked on
      →How often a user returns to your site
      →How long a user stays at your site

Usage Statistics- what do some of the key terms mean

- Analyze your stats from your ISP provider: Web servers keep logs of all visitor activity on your webpage
- Usage report includes: total hits, total files, total pages, total visits
- Hits: # of graphics on page + page itself
- Pages: # of pages looked at per visit

World Wide Web History

1990: The first tool for searching the Internet was called Archie (short for “archives”)

- the world's first internet search engine
- written by Alan Emtage at McGill University in Montreal
- It downloaded directory listings of all files located on public anonymous FTP servers; creating a searchable database of filenames

1991: "Gopher" was created

- by Mark McCahill of the University of Minnesota
- Gopher is a distributed document (shared by computers) search and retrieval network protocol designed for the Internet
- became obselete

1991: World Wide Web is developed at CERN

- Developed by Tim Berners Lee in Geneva Switzerland
- Problem: data was difficult to access and exchange for research
- All papers/files were now all stored in the same standardized manner → html files

1994: Yahoo is developed

- by Jerry Yang and David Filo
- Started as a Search Engine and Web Directory

1998: Google is developed

- By Larry Page and Sergey Brin

Review

- PageRank algorithm →the algorithm Google uses in their search engine
- Most popular search directory website → Yahoo
- Most popular search engine website → Google
- 3 components to get page to higher ranking → Text, Link, Popularity
- Who started Yahoo, who started Google? → Yahoo – David Filo and Jerry Yang , Google - Larry Page and Sergey Brin
- Search tool earlier that Yahoo and Google → Gopher, Archie