

# Lecture 26 – Driving Traffic to Websites

December 6, 2016

# Search Engine vs. Directory

*Note: Engines “won” ...they became more popular than directories*

- Search engine
  - An online tool that helps users of the Internet find sites and information
  - Most search engines use automated programs (spiders, robots) to look for relevant information to index in their database(s)
  - Examples: [www.google.com](http://www.google.com), [www.yahoo.com](http://www.yahoo.com), [www.bing.com](http://www.bing.com)
- Search directory
  - A catalog of websites organized by category to allow users to easily browse for the information they need
  - Search directories are organized by real people who discover new sites and relevant information by exploring the Internet themselves and by reviewing submitted sites
  - Example: [www.dmoz.org](http://www.dmoz.org); previously [www.yahoo.com](http://www.yahoo.com)
    - Historically, DMOZ provided database access to search engines like Google, Yahoo, others, to enhance their results (not so much anymore)

# History

- **Yahoo!** founded: January 1994
  - Started as a directory (manual)
- **AltaVista** founded: December 1995
  - The first search engine (automated)
- **Google** founded: September 1998
  - Started as a search engine
- **Bing**, formally Live Search, formally MSN Search founded: 1998
  - Originally used AltaVista data under their own brand

*Note: Most services, both directories and engines, provide access to each other's databases; when using any particular service it's not clear which company actually indexed (obtained) the results you're looking at*

# Definitions

- Search robots
  - A.k.a. bots, wanderers, spiders, and crawlers,
  - Tools many web search engines use to build their databases.
  - Most robots work like web browsers, going to webpages and following their links, except they don't require user interaction
- Search Engine Results Pages (SERP)
  - The page you see *after* you perform a search
  - A ranked listing of results (web pages) based on the keywords you provided as part of your search
  - Ranking is performed by an algorithm (usually proprietary); the higher the ranking (the position in the list) the higher the relevance to your search term(s)
- “Organic” results vs. Paid Placements
  - In addition to ranked listings from the algorithm (“organic” results), paid placements may appear on a SERP
  - The ranking of the paid placements is also determined by an algorithm which primarily counts payment as a ranking factor

# More Definitions (Paid)

- Impression
  - Every time a user views a SERP that displays your ad, it is counted as an impression
- Page view
  - Every time a user clicks your ad, it is counted as an page view
- Pay per click (PPC) and Cost per (thousand) Impressions (CPM)
  - A.k.a. Cost per click (CPC) and Cost per impression (CPI)
  - An advertising model used to direct traffic to websites, in which advertisers (website owners) pay a publisher (typically a search engine company) when the ad is viewed (an impression) or the ad is clicked (a page view)
  - Impressions → Clicks → Page Views → Conversions (sales)
- Clickthrough rate (CTR)
  - The number of ad clicks divided by the number of impressions, in a percentage
  - $CTR = (Clicks / Impressions) * 100\%$   
For example, if you received 7 clicks out of 1000 impressions, your page CTR would be 0.7%

# Web Development Activities for Search

- SEO - Search Engine Optimization
  - Getting the website to appear well in the “organic” SERPs
- SEM - Search Engine Marketing
  - Paying to get placement (advertise) on SERPs
  - Using media (any) to get people to visit a website (e.g. QR codes)
- SMO - Social Media Optimization
  - The use of social media outlets to generate publicity with the intent of driving website traffic

# Search Engine Optimization

On-page-factors and Off-page-factors

ON-THE-PAGE FACTORS		
These elements are in the direct control of the publisher		
CONTENT		
Cq	QUALITY	Are pages well written & have substantial quality content?
Cr	RESEARCH	Have you researched the keywords people may use to find your content?
Cw	WORDS	Do pages use words & phrases you hope they'll be found for?
Cf	FRESH	Are pages fresh & about "hot" topics?
Cv	VERTICAL	Do you have image, local, news, video or other vertical content?
Ca	ANSWERS	Is your content turned into direct answers within search results?
Vt	THIN	Is content "thin" or "shallow" & lacking substance?
ARCHITECTURE		
Ac	CRAWL	Can search engines easily "crawl" pages on site?
Ad	DUPLICATE	Does site manage duplicate content issues well?
Am	MOBILE	Does your site work well for mobile devices & make use of app indexing?
As	SPEED	Does site load quickly?
Au	URLS	Do URLs contain meaningful keywords to page topics?
Ah	HTTPS	Does site use HTTPS to provide secure connection for visitors?
Vc	CLOAKING	Do you show search engines different pages than humans?
HTML		
Ht	TITLES	Do HTML title tags contain keywords relevant to page topics?
Hd	DESCRIPTION	Do meta description tags describe what pages are about?
Hs	STRUCTURE	Do pages use structured data to enhance listings?
Hh	HEADERS	Do headlines & subheads use header tags with relevant keywords?
Vs	STUFFING	Do you excessively use words you want pages to be found for?
Vh	HIDDEN	Do colors or design "hide" words you want pages to be found for?

# THE PERIODIC TABLE OF SEO SUCCESS FACTORS

Search engine optimization — SEO — seems like alchemy to the uninitiated. But there's a science to it. Below are some important "ranking factors" and best practices that can lead to success with both search engines and searchers.

## ON-THE-PAGE SEO

CONTENT	ARCHITECTURE	HTML	TRUST	LINKS	PERSONAL	SOCIAL
<b>Cq</b> +3 Quality	<b>Ac</b> +3 Crawl	<b>Ht</b> +3 Titles	<b>Ta</b> +3 Authority	<b>Lq</b> +3 Quality	<b>Pc</b> +3 Country	<b>Sr</b> +2 Reputation
<b>Cr</b> +3 Research	<b>Ad</b> +2 Duplicate	<b>Hd</b> +2 Description	<b>Te</b> +2 Engage	<b>Lt</b> +2 Text	<b>Pl</b> +3 Locality	<b>Ss</b> +1 Shares
<b>Cw</b> +2 Words	<b>Am</b> +2 Mobile	<b>Hs</b> +2 Structure	<b>Th</b> +2 History	<b>Ln</b> +1 Numbers	<b>Ph</b> +3 History	
<b>Cf</b> +2 Fresh	<b>As</b> +1 Speed	<b>Hh</b> +1 Headers	<b>Ti</b> +1 Identity	<b>Vp</b> -3 Paid	<b>Ps</b> +2 Social	
<b>Cv</b> +2 Vertical	<b>Au</b> +1 URLs	<b>Vs</b> -2 Stuffing	<b>Vd</b> -1 Piracy	<b>Vi</b> -3 Spam		
<b>Ca</b> +1 Answers	<b>Ah</b> +1 HTTPS	<b>Vh</b> -1 Hidden	<b>Va</b> -1 Ads			
<b>Vt</b> -2 Thin	<b>Vc</b> -3 Cloaking					

## FACTORS WORK TOGETHER

All factors on the table are important, but those marked 3 carry more weight than 1 or 2. No single factor guarantees top rankings or success, but having several favorable ones increases the odds. Negative "violation" factors shown in red harm your chances.

OFF-THE-PAGE FACTORS		
Elements influenced by readers, visitors & other publishers		
TRUST		
Ta	AUTHORITY	Do links, shares & other factors make site a trusted authority?
Te	ENGAGE	Do visitors spend time reading or "bounce" away quickly?
Th	HISTORY	Has site or its domain been around a long time, operating in same way?
Ti	IDENTITY	Does site use means to verify its identity & that of authors?
Vd	PIRACY	Has site been flagged for hosting pirated content?
Va	ADS	Is your content ad-heavy, especially "above-the-fold"?
LINKS		
Lq	QUALITY	Are links from trusted, quality or respected web sites?
Lt	TEXT	Do links pointing at pages use words you hope they'll be found for?
Ln	NUMBER	Do many links point at your web pages?
Vp	PAID	Have you purchased links in hopes of better rankings?
Vi	SPAM	Have you created links by spamming blogs, forums or other places?
PERSONAL		
Pc	COUNTRY	What country is someone located in?
Pl	LOCALITY	What city or local area is someone located in?
Ph	HISTORY	Has someone regularly visited your site or socially favored it?
Ps	SOCIAL	Has someone or their friends socially favored the site?
SOCIAL		
Sr	REPUTATION	Do those respected on social networks share your content?
Ss	SHARES	Do many share your content on social networks?

Note:  
White Hat  
(in green)  
vs.  
Black Hat  
(in red)



# Content

On-the-page SEO

# On-the-page: content quality



- Are pages well written and have substantial quality content?
- The main reason search engines want to rank websites highly: *good content*
- E.g.: a search for “octopus facts” → National Geographic website

# On-the-page: content research



- Have you researched the keywords people may use to find your content?
- Understand the target audience – whatever words they use, even if incorrect (including misspelled), must be understood
- E.g. “Eastman Gelatine” (everyone searches for “kodak gelatin”)
- E.g. “web designer rochester ny” (even though Professor Kostin is a web “developer”)

# On-the-page: content words



- Do pages use words and phrases you hope they'll be found for?
- Helps: liberal use of keywords in titles, headings a paragraphs
- E.g. "...Kostin is a **web developer** who has been **designing websites**..."

# On-the-page: content freshness



- Are pages fresh, and are they about hot topics?
- Content that changes from time to time suggests it is more relevant
- News websites, bulletin boards, blogs rank better than static sites

# On-the-page: content vertical



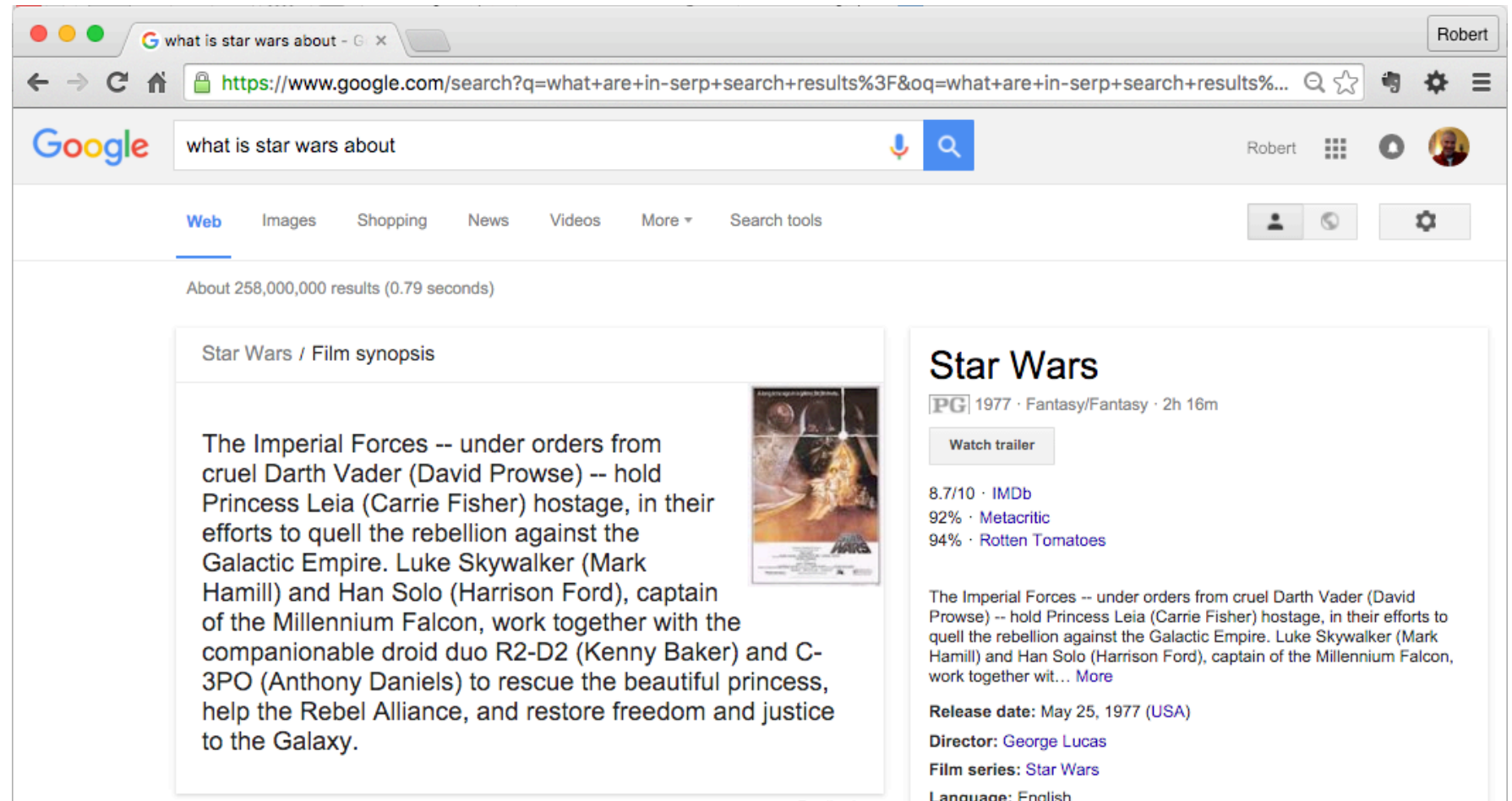
“**Vertical** content, as distinct from a general content, focuses on a specific segment. They are also called specialty or topical search engines. The **vertical content** area may be based on topicality, media type, or genre of **content**.”

- Do you have image, local, news video or other “vertical” content?

# On-the-page: content answers



- Is your content turned into direct answers within search results?



On-the-page: content thinness/shallowness



- Is content thin or shallow – lacking substance?
- Websites are penalized if they have very little content
- Noted: Are sites with strong imagery but few words penalized?



# Architecture

On-the-page SEO

# On-the-page: architecture, crawling



- Can search engines easily find and crawl pages on the site?
- XML Sitemaps
- [www.xml-sitemaps.com](http://www.xml-sitemaps.com)

# On-the-page: architecture, duplicate management



- Does site manage duplicate content issues well?
- Reduce how often blocks of content are replicated across multiple pages in one website
- Mostly an issue for very large, content management system (CMS) driven websites

# On-the-page: architecture, mobile



- Does your site work well for mobile visitors, on smartphones and tablets?

# On-the-page: architecture, speed



- Does the site load quickly?
- What is the “weight” of all the HTML, CSS, JavaScript and images?  
...and web fonts, and more?

## On-the-page: architecture, URLs



- Are URLs short and do they contain meaningful keywords to page topics?

## On-the-page: architecture, HTTPS



- Does the site use HTTPS to provide secure connections for visitors?

On-the-page: architecture, cloaking



- Do you show search engines different pages than humans?

# HTML

On-the-page SEO

# On-the-page: HTML title tags



- Do HTML title tags contain keywords relevant to the page topic?
- E.g. `<title>Octopus - Wikipedia, the free encyclopedia</title>`
- Note: specificity goes left-to-right



# On-the-page: HTML description tags



- Do meta description tags describe what the page is about?
- Whatever is in the meta description, is added to the SERP
- Better than leaving it to chance (the search engine will pick the first plain text it finds)

```
<meta charset="utf-8">  
<title>Robert Kostin Web Developer | Web Designer | Rochester, NY - New York</title>  
<meta name="description" content="Robert Kostin is a Master Web Developer in the Upstate New York area.">
```

SERP

**Robert Kostin Web Developer | Web Designer | Rochester, ...**  
[www.rkostin.com/](http://www.rkostin.com/) ▼  
Robert Kostin is a Master Web Developer in the Upstate New York area.



# On-the-page: HTML structure/outline



- Are pages structured with heading tags to describe the content in a logical manner?
- Well structured documents are indexed better than loose text
- Note: don't use an H2 without an H1 above it; don't use an H3 without an H2 above it; et cetera
- <http://gsnedders.html5.org/outliner>

# On-the-page: HTML heading tags



- Do headlines and subheadings use header tags with relevant keywords?
- Headings: `<h1>`, `<h2>`, `<h3>`
- Words are highly indexed; must be relevant
- “Lead-in” text is not relevant; e.g.:  
`<h1>Octopi in the Wild</h1>`  
`<h2>It’s really interesting!</h2>` ← *this isn’t helping*  
`<p>The story of octopi begins with...`

# On-the-page: HTML stuffing



- Do you excessively use words you want pages to be found for?
- Gratuitous use of keywords in HTML is penalized
- I.e. using keywords in a way that doesn't enhance the content

# On-the-page: HTML hiding



- Do colors or design styles hide words you want pages to be found for?
- Tricks to stuff HTML with keywords are penalized
- E.g.  

```
<div style="background-color: #333333; color: #333333;">  
    keyword, keyword, keyword, keyword, keyword, keyword  
</div>
```

# Trust

Off-the-page SEO

# Off-the-page: trust authority



- Do links, shares and other factors make site a trusted authority?

# Off-the-page: trust engagement



- Do visitors spend time reading or “bounce” away quickly?  
(high bounce rate = bad)

# Off-the-page: trust history



- Has site or its domain been around a long time, operating in the same way?

# Off-the-page: trust identify



- Does site use means to verify its identify and that of authors?



Off-the-page: trust, pirated



- Has site been flagged for hosting pirated content?

Off-the-page: trust, ads



- Is your content ad-heavy, especially above-the-fold?

# Links

Off-the-page SEO

## Off-the-page: link quality



- Are links from trusted, quality or respected web sites?

## Off-the-page: link text



- Do links point at your webpages use words you hope they'll be found for?

## Off-the-page: link number



- Do many links point at your web pages?

## Off-the-page: paid links



- Have you purchased links in hopes of better rankings?

## Off-the-page: link spam



- Have you created many links by spamming blogs, forums or other places?

# Personal

Off-the-page SEO

## Off-the-page: personal country



- What country is someone located in?

## Off-the-page: personal locality



- What city or local area is someone located in?

## Off-the-page: personal history



- Has someone regularly visited your site or socially favored it?

# Off-the-page: personal social



- Have your friends socially favored the site?

# Social

Off-the-page SEO



# Off-the-page: social reputation



- Do those respected on social networks share your content?

# Off-the-page: social shares



- Do many share your content on social networks?

# Google Tools - Online

Webmaster tools, Analytics

# Create a Google Analytics Account

- Create a Google Analytics account: <http://www.google.com/analytics>
- Add an Account to Google Analytics:  
<https://support.google.com/analytics/answer/1009694?hl=en>
  - Use any website you own (and can edit code)
  - Project 2 would work

# Sign-in to Google Webmasters Tools

- <https://www.google.com/webmasters>
- Click the “ADD A SITE” button

## Generate an XML Sitemap

- [www.xml-sitemaps.com](http://www.xml-sitemaps.com)