The (Dirty) Secrets of Search

Presented by

Steve Fenick

New Horizons
Computer Learning Center

Richmond, VA

FORUM2013

> 242.6 million

S Internet Users

FORUM2013

http://visual.ly/yoda-london-your-guide-seo

> 3.7 Trillion

Webpages on the Internet

FORUM2013

http://visual.ly/yoda-london-your-guide-seo

> 4 Billion

Google

Searches Daily

FORUM2013

http://visual.ly/yoda-london-your-guide-seo

> 121 Million

per Hour

FORUM2013

http://visual.ly/yoda-london-your-guide-seo

> 2 Million

per Minute

FORUM2013

http://visual.ly/yoda-london-your-guide-seo

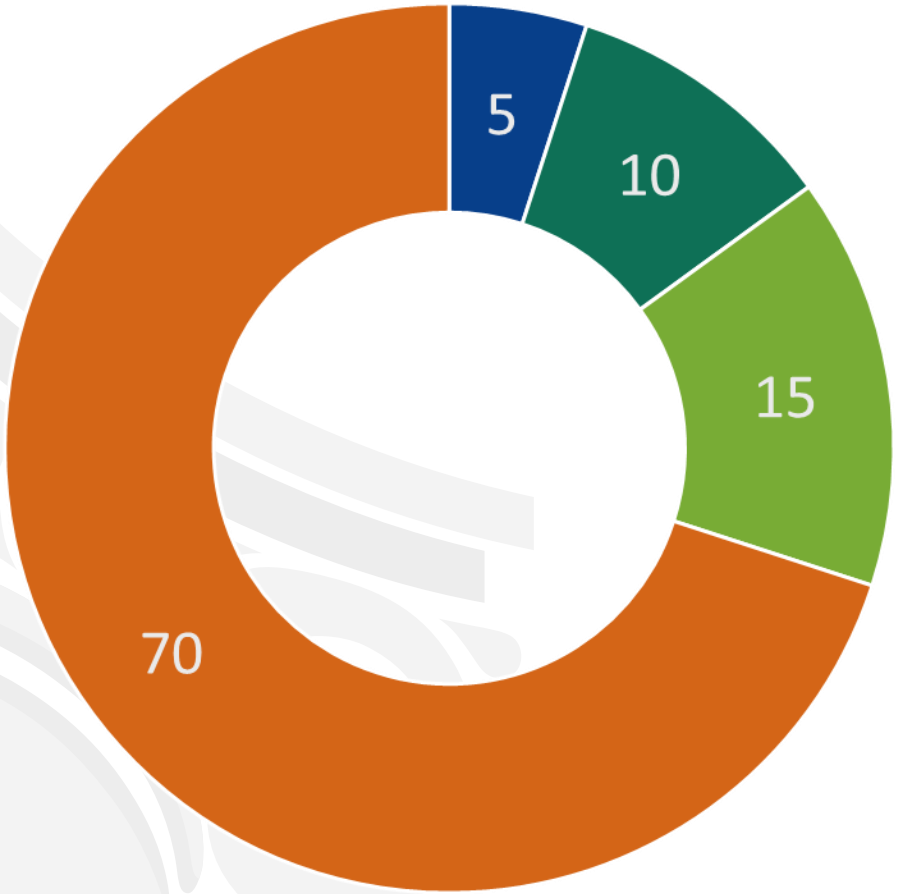
> 34,000

per Second

FORUM2013

http://visual.ly/yoda-london-your-guide-seo

Search is the #1 Driver of Traffic to Sites



- Other
- Bing
- Yahoo
- Google

FORUM2013

11/6/2013 http://visual.ly/yoda-london-your-guide-seo

Search is used

300%

more than

FORUM2013

http://visual.ly/yoda-london-your-guide-seo

93%
of Online
experience
begins with search.

75%
of Users never
scroll passed the
first page.

23%
search the
second
page.

40%
of Users click the
first ranking
site.

70%
of search links
Users click are
organic.

86%
of traffic is
generated
through search.

FORUM2013

http://visual.ly/yoda-london-your-guide-seo

Topics

1. The Science of Search
2. Bias
3. Manipulation
4. Research

FORUM2013

The Science of Search

- Web Crawler
- Index
- Meta Tags & Keywords
- Results

Web Crawler

- Search engines use special software robots: Spiders
- Beginning on a popular site, spiders crawl through the web building lists of words found on sites or pages.
- Spiders are looking for two things:
    1. The words within the page
    2. Where the words were found

FORUM2013

http://computer.howstuffworks.com/internet/basics/search-engine.htm

Meta Tags

allow the owner of a page to specify key words and concepts under which the page will be indexed.

Dangers

Meta tagging popular topics that have nothing to do with page content.

Protection

Spiders will correlate meta tags with page content, rejecting meta tags that don't match the page.

FORUM2013

http://computer.howstuffworks.com/internet/basics/search-engine.htm

Index

Key components involved in making data accessible:

1. The **information stored with the data**
2. The **method by which the information is indexed**

Hash Tables

A formula is applied to attach a numerical value to each word. The formula is designed to evenly distribute the entries across a predetermined number of divisions. This numerical distribution is different from the distribution of words across the alphabet, and that is the key to a hash table's effectiveness.

FORUM2013

http://computer.howstuffworks.com/internet/basics/search-engine.htm

Building a Search

- Boolean Operators
  AND, OR, NOT, FOLLOWED BY, NEAR, “ ”
- Concept-Based Searching
- Natural-Language Queries

FORUM2013

http://computer.howstuffworks.com/internet/basics/search-engine.htm

Bias

- Search Engine Optimization
- Personalization
- The Filter Bubble

The Filter Bubble

Search Engines Collect and Personalize for

1. Search History
2. Signed-Out Web History
3. Signed-In Web History

Google uses 'signals' including

- Search history
- Location
- Active browser
- Computer being used
- Language configured

FORUM2013

http://computer.howstuffworks.com/search-engine-optimization1.htm

Pop the Bubble

- Make yourself difficult to stereotype.
- Erase your browser history and cookies regularly.
- Use an incognito window for exploring content.
- Use Twitter instead of Facebook for news.

FORUM2013

http://computer.howstuffworks.com/search-engine-optimization1.htm

Search Engine Optimization

Search engine optimization (SEO) is the process of affecting the visibility of a website or a web page in a search engine's "natural" or un-paid ("organic") search results

FORUM2013

http://computer.howstuffworks.com/search-engine-optimization1.htm

White Hat Techniques

- Key Words
  - Page Titles
  - Section Headers
- Linking to increase Page Rank
  - Link Exchanges

FORUM2013

http://computer.howstuffworks.com/search-engine-optimization1.htm

Manipulations

- Black Hat Optimization
- Link Schemes (JC Penney Scam)

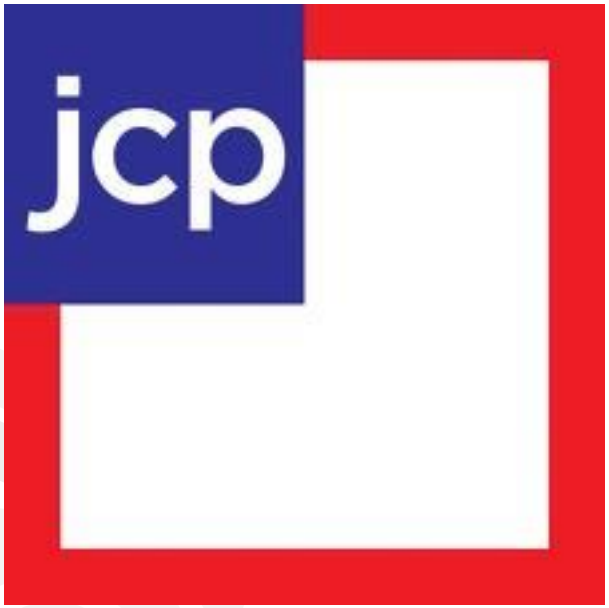
Black Hat Techniques

- Keyword Stuffing
- Page Stuffing
- Selling / Farming Links

FORUM2013

http://computer.howstuffworks.com/search-engine-optimization1.htm

Link Schemes



http://computer.howstuffworks.com/search-engine-optimization1.htm

Think like a Researcher

FORUM2013

Hard or Soft Research?

- Hard Research
  - Scientific or objective
  - Expectations of data and/or proof
  - Most reliable
- Soft Research
  - Subjective, cultural or opinion-based

http://computer.howstuffworks.com/internet/basics/search-engine.htm

Choose Suitable Online Authorities

- Government Publications
  Google's **Uncle Sam** search
- Government Authorities
  NHTSA
- Non-Government Websites NOT influenced by advertising or sponsorship
  Consumer Watch

FORUM2013

http://computer.howstuffworks.com/internet/basics/search-engine.htm

Use Different Search Engines & Keywords

- Start with broad initial researching
  - Internet Public Library
  - DuckDuckGo
  - Clusty
  - Wikipedia
  - Mahalo
- Visible Web Searches
  - Google
  - Ask.com

FORUM2013

http://computer.howstuffworks.com/internet/basics/search-engine.htm

Filter & Validate the Content

- Consider the Source and Date of Publication
- Be Suspicious of
  - Amateurish Presentation
  - Scientific or medical pages that display scientific or medical advertising.
  - Any ranting, overstating, overly-positive, or overly-negative commentary.
  - Commercial consumer websites can be good resources, but be skeptical of every comment you read.

FORUM2013

http://computer.howstuffworks.com/internet/basics/search-engine.htm

Choose a Research-Friendly Web Browser

- Multiple tab pages open simultaneously.
- Bookmarks/favorites that are fast and easy to manage.
- Page history that is easy to recall.
- Loads pages quickly for your computer's memory size.

2013's Best research browsers are Chrome and Firefox

FORUM2013

http://computer.howstuffworks.com/internet/basics/search-engine.htm