Libraries in Search Engines
Search Engine Optimization (SEO) for Libraries

Florida Library Webinars
Novare Library Services
April 2, 2014

Patrick OBrien+
Semantic Web Research Director
Montana State University Library

Jason Clark
Head of Library Informatics & Computing
Montana State University Library
@jaclark #libseo
A Research Team Effort

Scott Young
Digital Initiatives Librarian at Montana State University

Jason Clark
Head of Library Informatics & Computing at Montana State University Library

Patrick OBrien
Semantic Web Research Director at Montana State University Library

Kenning Arlitsch
Dean of the Library at Montana State University. His book, co-authored with Patrick OBrien is titled Improving the Visibility and Use of Digital Repositories through SEO, and was published in February 2013.
Today's Agenda

● What is Search Engine Optimization (SEO)?
● Why do SEO?
● Black Hat vs. White Hat SEO
● Recommendations
● Key SEO techniques
● Resources
● Questions
What is SEO for Libraries?

The act of creating indexable content for commercial search engines (SE)...
...to increase access, visibility and use of library content and resources ...
...by patrons that value it
Why do SEO?

- ~200 million people in the US use the Internet\(^1\)
  - 81% of U.S. adults
  - 95% of U.S. teens
- The US submits 19+ Billion queries per month to commercial SE\(^2\)
- Google owns 67% of these queries

---


"Search engines continue to dominate, topping the list of electronic sources most used to find online content (93%), followed closely by Wikipedia (88%). The key difference in usage between search engines and Wikipedia is the frequency - 75% of students who use search engines do so daily, compared to 20% of those who use Wikipedia."


http://www.oclc.org/content/dam/oclc/reports/2010perceptions/2010perceptions_all_singlepage.pdf
http://www.oclc.org/content/dam/oclc/reports/2010perceptions/collegestudents.pdf
"Discovery happens elsewhere."

- Lorcan Dempsey (in 2007)

Vice President and Chief Strategist
Online Computer Library Center (OCLC)
http://orweblog.oclc.org/archives/001430.html
A MILLION people with SEO in their job title? That fact blows me away. Great post. I'm still amazed that people are trying to game the algorithm and out-smart Google.
A WONDERFUL PAIN DESTROYING COMPOUND

THE STRONGEST AND BEST LINIMENT KNOWN FOR THE CURE OF ALL PAIN AND LAMENESS.

USED EXTERNAL ONLY FOR:

RHEUMATISM
NEURALGIA
SCIATICA
LAME BACK
LUMBAGO
CONTRACTED MUSCLES
TOOTHACHE
SPRAINS
SWELLING

Etc.

CLARK STANLEY
Providence, R.I.

A CURES FROST BITES
CHILL BLAINS
BRUISES
SORE THROAT
BITES OF ANIMALS
INSECTS AND
REPTILES.

IS GOOD FOR EVERYTHING.

A LINIMENT SHOULD BE GOOD FOR.

It Gives Immediate Relief.

Price: 50 Cts.

http://www.flickr.com/photos/22245459@N06/2386354917
https://en.wikipedia.org/wiki/Clark_Stanley
Black Hat SEO = Banned

- Using deceptive redirects
  - "Cloaking"
  - "Shadow" domains
- Multiple "Doorway Pages" with poor content optimized for different keywords
- "Link Schemes" trying to game the SE
- "Hidden" text or links
- Anything "sneaky" or untrustworthy
White Hat SEO

● SE approved "Webmaster Guidelines"
● Focus on improving user experience
  ○ Unique & accurate content
  ○ User relevant vocabulary & keywords
  ○ Intuitive site architecture
● Build Stability & Trust
Recommended Library Process

1. Institutionalize SEO
   ○ Accurate Measurement Tools
   ○ Strategic Plan
2. Traditional SEO
   ○ Get Indexed
   ○ Get Visible
3. Semantic SEO
   ○ Get Relevant
Google – Baseline May 15 2013

Google search results for "montana state university library"
Setup Measurement & Evaluation

1. Associate a single Google Account with your Institution
   (e.g., msuseo@gmail.com)

2. Establish baseline metrics
   a. Google Analytics (gAnalytics)
   b. Google Webmaster Tools (gMaster)

3. Activate Google Services
   a. Google Places for Business (gPlaces)
   b. Google+ (G+)
   c. Google AdWords (gAdWords)
Make Improvements

1. Address gWebmaster "errors"
   a. Ensure content is indexable
   b. Reduce Index noise
2. Claim, consolidate and standardize all listings
   a. gPlaces
   b. Google Maps (gMaps)
   c. G+
3. Improve On-Page Optimization
Creating Indexable Content

● User intuitive hierarchy & site architecture
  ○ Short URLs containing keywords
  ○ Simple directory structure
  ○ Incorporate "Breadcrumbs"
  ○ Single URL to each page
  ○ Easy navigation between homepage and any webpage, and back

“Every page should be reachable from at least one static text link.”
https://support.google.com/webmasters/answer/35769
Creating Indexable Content

- Link a human readable HTML sitemap from homepage
- Implement machine-readable XML file(s) to provide SE content inclusion direction
  - "sitemap index"
  - "sitemap" file(s)

“Google adheres to Sitemap Protocol 0.9 as defined by sitemaps.org.”
https://support.google.com/webmasters/answer/156184
Reduce “Noise” in the Index

Implement robots.txt
  ● text file at top level / root of web site
  ● Provides SE content exclusion rules

Disallow: /staff/
#allow primary mobile page
Allow: /finding-aids/m/index.php
Allow: /finding-aids/m/$
Reduce “Noise” in the Index

- Eliminate duplicate content
- Dynamic pages, search result pages
- Ensure every page is unique

Techniques:

1. robots.txt directives
2. link relations and index directive markup
Reduce “Noise” in the Index

Link relations and index directive markup

- HTML page <head> markup that provides canonicalization and index exclusion rules

```html
<link rel="canonical" href="http://arc.lib.montana.edu/finding-aids/item/23" />

<a rel="nofollow" href="http://www.lib.montana.edu/login.html">User Login</a>

<meta name="robots" content="nofollow,noindex" />
```
Claim, consolidate and standardize listings

1. Search gMaps and G+ for all variations
   a. Name of library
   b. Name of library building
   c. All library phone numbers

2. Login gPlaces with Institutional Google Account

3. Search gPlaces & claim ownership of all variations from step 1
Claim, consolidate and standardize listings

4. Identify gPlaces variation with G+ "pre-approved custom URL"

5. Consolidate under gPlaces variation with G+ "pre-approved custom URL"

6. Complete and standardize consolidated gPlaces and G+ profiles
   a. Name
   b. Address
   c. Phone Number(s)
Improve On-Page Optimization

1. Keyword phrase analysis
   a. Patron relevant
   b. Low to Moderate competition
   c. High volume

2. Write / rewrite Web page
   a. HTML <title>
   b. <description> metatag

3. Ensure
   a. all links contain descriptive "Anchor Text"
   b. Use <H1>, <H2> HTML tags to "outline" the page around the topic (i.e. <title>)
   c. Text is natural, easy-to-read and relevant to the page outline
Keyword Phrase Analysis

1. Use the vocabulary of your patrons

2. Which keyword phrase?
   a. institutional repository
   b. academic research
   c. research papers
   d. open access library
Keyword Phrase Analysis

- Use Google AdWords

<table>
<thead>
<tr>
<th>Search terms</th>
<th>Avg. monthly searches</th>
<th>Competition</th>
<th>Suggested bid</th>
</tr>
</thead>
<tbody>
<tr>
<td>research papers</td>
<td>3,600</td>
<td>High</td>
<td>$3.53</td>
</tr>
<tr>
<td>academic research</td>
<td>880</td>
<td>Low</td>
<td>$1.40</td>
</tr>
<tr>
<td>institutional repository</td>
<td>210</td>
<td>Low</td>
<td>$0.00</td>
</tr>
<tr>
<td>open access library</td>
<td>110</td>
<td>Low</td>
<td>$0.00</td>
</tr>
</tbody>
</table>
Keyword Phrase Analysis

- Check Singular vs Plural

<table>
<thead>
<tr>
<th>Search terms</th>
<th>Avg. monthly searches</th>
<th>Competition</th>
<th>Suggested bid</th>
</tr>
</thead>
<tbody>
<tr>
<td>research paper</td>
<td>~14,800</td>
<td>Medium</td>
<td>$3.13</td>
</tr>
<tr>
<td>research papers</td>
<td>~3,600</td>
<td>High</td>
<td>$3.53</td>
</tr>
</tbody>
</table>
Writing Titles and Descriptions

Writing web copy for the robot and the human

<title></title>

<meta name="description" />
Writing Titles and Descriptions

<title>
"Keyword Phrase Topic - Context"

Example:

“Staff Directory - Contact Information - Montana State University (MSU) Library”
Writing Titles and Descriptions

<description>
a short, declarative sentence that incorporates a reversal of the page's <title> keyword phrase topic and context

Example:
"Montana State University (MSU) Library staff directory and employee contact information that includes library departments, roles, job titles, and phone numbers."
Links and Anchors

A link is the connection between one piece of hypertext and another.
Introducing Semantic Markup

- HTML5 semantic tags and microdata that helps classify page types and types of content on the page

“If Google understands the content on your pages, we can create rich snippets—detailed information intended to help users with specific queries.”

https://support.google.com/webmasters/answer/99170?hl=en&ref_topic=1088472
Item: Women posing on college locomotive engine

Title: Women posing on college locomotive engine
Creator: unknown
Date: unknown
Description: Women posing on a college locomotive engine.
Notes:
Physical Description: Photo print - Black and White
Subjects: Locomotive
Keywords: locomotive
Photograph ID: parc-000432
Item: Women posing on college locomotive engine

Title: Women posing on college locomotive engine
Creator: unknown
Date: unknown
Description: Women posing on a college locomotive engine.
Notes:
Physical Description: Photo print - Black and White
Subjects: Locomotive
Keywords: locomotive
Photograph ID: parc-000432
<h2 class="mainHeading">Item: <strong>Women posing on college locomotive engine</strong></h2>
<ul class="item">
<li>
<ul class="metadata">
<li class="object">
<a title="zoom & scan view - Women posing on college locomotive engine" href="/msu-photos/viewer.php?id=432" alt="Women posing on college locomotive engine"><img src="/msu-photos/objects/parc-000432.jpg" alt="Women posing on college locomotive engine" /></a>
</li>
<li class="describe">
<p><strong>Title:</strong> Women posing on college locomotive engine</p>
<p><strong>Creator:</strong> unknown</p>
<p><strong>Date:</strong> 1931-05-15</p>
<p><strong>Description:</strong> Women posing on a college locomotive engine.</p>
<p><strong>Physical Description:</strong> Photo print - Black and White</p>
<p><strong>Keywords:</strong> locomotive</p>
<p><strong>Photograph ID:</strong> parc-000432</p>
</li>
</ul>
</li>
</ul>
Item: <strong><span property="name">Women posing on college locomotive engine</span></strong>

<p><strong>Title:</strong> <span property="name">Women posing on college locomotive engine</span></p>

<p><strong>Creator:</strong> <span property="author">unknown</span></p>

<p><strong>Date:</strong> <span property="dateCreated">1931-05-15</span></p>

<p><strong>Description:</strong> <span property="description">Women posing on a college locomotive engine.</span></p>

<p><strong>Physical Description:</strong> <span property="genre">Photo print - Black and White</span></p>

<p><strong>Keywords:</strong> <span property="keywords">locomotive</span></p>

<p><strong>Photograph ID:</strong> parc-000432</p>
Schema.org

A lingua franca for web classification and markup

Google, Bing, Yahoo, Yandex
SEO in Libraries

Questions?