Structuring Your Site for Better SEO

March 1, 2017
Hello!

Dawn Pointer McCleskey, MLS

Ashley Wichman
Learning Goals

- Websites need to be machine-friendly
- Add meta elements and markup to your page templates
- Test your site’s SEO, and iterate
Our favorite SEO sites

webmasters.googleblog.com
developers.google.com/search/docs/guides/

moz.com/blog
searchenginejournal.com
searchengineeland.com
searchenginewatch.com
yoast.com
Crawl

“Can search engines easily get around your site?”
Crawl

- Commercial services crawl to populate their indexes
- They display results based on their crawling
- Bad or incomplete crawling can be influenced by...
Robots.txt

- A text file read by search engine spiders
  - Placed on the root of your domain:
    https://www.agency.gov/robots.txt

- Tells the spiders where they can and can’t go on your site

- Also should tell spiders where to find the official list of your site’s URLs (aka your sitemap)
Robots.txt

User-Agent: *
Disallow: /tagged/
Disallow: /post/
Sitemap: https://search.digitalgov.gov/sitemap.xml
Sitemap

- A file that lists the URLs of your site
  - Also placed on the root of your domain: https://www.agency.gov/sitemap.xml

- Tells crawlers about your site, its content and its organization

- Maximum file size: 50 MB
This XML file does not appear to have any style information associated with it. The document tree is shown below.

```xml
<url>
  <loc>https://search.digitalgov.gov/</loc>
  <changefreq>daily</changefreq>
  <priority>1</priority>
</url>

<url>
  <loc>https://search.digitalgov.gov/manual/training.html</loc>
  <lastmod>2017-02-23</lastmod>
  <changefreq>monthly</changefreq>
</url>

<url>
  <lastmod>2017-02-22</lastmod>
  <changefreq>monthly</changefreq>
</url>

<url>
  <loc>https://search.digitalgov.gov/releases/november-2016.html</loc>
  <lastmod>2017-02-22</lastmod>
  <changefreq>monthly</changefreq>
</url>
```
Fetch as Google

- “The Fetch as Google tool enables you to test how Google crawls or renders a URL on your site.

- You can use Fetch as Google to see whether Googlebot can access a page on your site, how it renders the page, and whether any page resources (such as images or scripts) are blocked to Googlebot.”
Fetch as Bingbot

- Same concept as Google’s :)

Fetch as Bingbot

- Same concept as Google’s :)
Duplicate

“Does your site manage duplicate content issues well?”
If a single link is being served multiple ways, search engines have to "guess" which version is the right one.

Multiple URLs for same content, depending on folder
- http://agency.gov/similar-content
- http://www.agency.gov/similar-content
- https://agency.gov/similar-content
- https://www.agency.gov/similar-content
- https://agency.gov/content/similar-content
- https://content.agency.gov/similar-content
Canonical URLs

- A rel=canonical link in the <head> tells search engines which version is the version of record

- Google’s rel=canonical on their page about rel=canonicals
301 Redirects

- Use 301 redirects when permanently moving content
- The redirect will tell search engines that a page has permanently moved to a new location
Internal Linking

- Be consistent when linking within your website
  
  - I.e., make sure all your page links are
    https://agency.gov/page/ or www.agency.gov/page/
    or /page/
Preferred Domain

- For Google, set your preferred domain (www or non-www)
  - Search Console -> Choose the site you want to set it for -> Gear Icon -> Site Settings -> Preferred Domain section

- Can use 301s to redirect traffic from your non-preferred domain
Mobile

“Does your site work well for mobile devices?”
Mobile

- In spring 2016, Google increased the impact of mobile-friendliness as a ranking signal

- Also:

<table>
<thead>
<tr>
<th>Devices</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>desktop</td>
<td>56.9%</td>
</tr>
<tr>
<td>mobile</td>
<td>37.2%</td>
</tr>
<tr>
<td>tablet</td>
<td>5.9%</td>
</tr>
</tbody>
</table>

Federal websites!
Mobile-friendliness Testing
Speed

“Does your site load quickly?”
Speed

- Users will abandon your site if it does not load fast enough

- Impact on relevancy ranking not as great as other factors (Google noted this back in 2010) but it is critical to overall UX
URLs

“Do URLs contain meaningful keywords to page topics?”
URLs

● Use simple, human-friendly URLs

● Use punctuation
  ○ “The URL
    
    http://www.example.com/green-dress.html
    
    is much more useful to us than
    
    http://www.example.com/greendress.html.
  ○ We recommend that you use hyphens (-) instead of underscores (\_\_)”
HTTPS

“Does your site use HTTPS to provide a secure connection for visitors?”
HTTPS

- Since 2014, Google has been advocating “HTTPS Everywhere”
- It is a lightweight ranking signal
- Also, we make it easy with DigitalGov Search!
HTTPS Testing

Secure HTTP (HTTPS)
Last updated February 24, 2017

This data measures whether federal domains support the HTTPS protocol (e.g., //www) and the strength of that support. HTTPS provides a secure connection across the Internet between websites and their visitors, and is becoming the norm for e-commerce and public web services. As part of this shift, the U.S. federal government is in the process of transitioning entirely to HTTPS.

Note that HTTPS generally does not affect whether a website is vulnerable to hacking. For more information on what HTTPS does (and doesn’t) do, visit the HTTPS FAQ.

HTTPS and TLS data was last collected through scans of the public Internet on February 24, 2017.

By Agency

Show 10 | 25 | 50 | 100 entries  Download CSV

- Domain
- Uses HTTPS
- Strict Transport Security (HSTS)
- Preloaded (recommended)
- SSL Labs Grade

HTTPS enforced. Consider preloading this domain to enforce HTTPS across the entire zone.

Known public subdomains:
100% of sites https://www.whitehouse.gov enforce HTTPS.
100% of public sites known to the Digital Analytics Program enforce HTTPS.

For more details, read our methods page or download subdomain data for this agency.

Showing 1 to 1 of 1 entries (filtered from 1,118 total entries)
Titles

“Do HTML title tags contain keywords relevant to page topics?”
Key Attributes of Excellent Titles

- **Unique to the page**
  - Only include site or agency name following a pipe
  - `<title>API documentation for Commerce.gov | Department of Commerce</title>`

- **Descriptive and short** (<70 characters)

- **Relevant to queries**
Description

“Do meta descriptions describe what pages are about?”
<meta name="Description" content="Description text goes here.">

- Unique to page
- ~155 characters including spaces
- Keywords at beginning
Headers

“Do headlines & subheads use header tags with keywords?”
Headers

- Use `<h#>` tags!
- Page title should be in the `<h1>`
- Tags needs to be in order
  - `<h1>`
  - `<h2>`
  - `<h3>`
  - etc.
<meta> tags Google understands

Positive commands:

<title>
<meta name="description" content="blah blah"/>
<meta name="google-site-verification" content="...">
<meta http-equiv="Content-Type" content="...; charset=..."/>
<meta charset="..."/>

Negative commands:

<meta name="robots" content="various_commands_see_help_page">
<meta name="googlebot" content="various_commands_see_help_page">
<meta name="google" content="nositelinkssearchbox"/>
<meta name="google" content="notranslate"/>
Structure

“Do pages use structured data to enhance listings?”
Structure

- Informs rich snippets and “cards” in search results
- Allows computers to interpret what the text means
  - “Semantic markup”
Say what?

Entities & Attributes

Nouns & Adjectives

What is it, and what’s it like?
Apple

Color: red
Variety: Gala
Background: white
Quantity: 1
Structured markup options

Schema.org
- itemprop

Open Graph Protocol
- og:

Dublin Core
- dc:

...and more!
Google prefers schema.org

Recommended entity type: CreativeWork

Subtypes:
- articles
  - news
  - reports
- blog posts
- datasets
- maps
- webpages
- entire websites
Google prefers schema.org

Other useful entity types

- **Organization**
  - GovernmentOrganization
  - NOTE! Google’s org cards are managed through Google My Business

- **ContactPoint**
  - PostalAddress
Google prefers schema.org (2)

**JSON-LD**
- `<script> in the <head>`

**RDFa**
- inline properties in the `<body>`

**Microdata**
- inline properties in the `<body>`
<script type="application/ld+json">
{
"@context": "http://schema.org",
"@type": "BreadcrumbList",
"itemListElement": [{
"@type": "ListItem",
"position": 1,
"item": {
"@id": "https://example.com/books",
"name": "Books",
"image": "http://example.com/images/icon-book.png"
}
}, {
"@type": "ListItem",
"position": 2,
"item": {
"@id": "https://example.com/books/authors",
"name": "Authors",
"image": "http://example.com/images/icon-author.png"
}
}, {
"@type": "ListItem",
"position": 3,
"item": {
"@id": "https://example.com/books/authors/annleckie",
"name": "Ann Leckie",
"image": "http://example.com/images/cover-ancillary-justice.png"
}
}, {
"@type": "ListItem",
"position": 4,
"item": {
"@id": "https://example.com/books/authors/ancillaryjustice",
"name": "Ancillary Justice",
"image": "http://example.com/images/cover-ancillary-justice.png"
}
}
]
</script>
Microdata
You don’t have to do it all!

Essential schema.org properties

Author
dateModified
datePublished
Headline
Name
Url
Description

Breadcrumb (WebPage)

http://www.w3resource.com/schema.org/introduction.php
Special stuff for your homepage

Control your site name

```html
<head itemtype="http://schema.org/WebSite">
  <title itemprop="name">Your WebSite Name</title>
  <link rel="canonical" href="https://example.com/" itemprop="url">
</head>
```

Enable site search within Google results

```html
<div itemtype="http://schema.org/WebSite">
  <meta itemprop="url" content="https://www.example.com/">
  <form itemprop="potentialAction" itemtype="http://schema.org/SearchAction">
    <meta itemprop="target" content="https://query.example.com/search?q={search_term_string}"/>
    <input itemprop="query-input" type="text" name="search_term_string" required/>
    <input type="submit"/>
  </form>
</div>
```
Don’ts

Stuffing
- Meta keywords
- Meta description

Hidden
- Text on page

Cloaking
- Showing different pages to engines than to humans
Learning Goals

- Websites need to be machine-friendly
- Add meta elements and markup to your page templates
- Test your site’s SEO, and iterate
Upcoming Sessions

- DigitalGov Search Basics:
  - Thursday, March 30, 2017, 2:00 PM - 3:00 PM EST
  - Held Bimonthly

- April - Indexing your content directly with DigitalGov Search
- June - Data driven search design
- October - TBD

- Video recordings available on our website
Thank you!