

Dan Cohen's Digital Humanities Blog » Blog Archive » Search Engine Optimization For Smarties

A Google search for “Sputnik” gives you an authoritative site from NASA in the top ten search results, but also a web page from the skydiver and ballroom-dancing enthusiast Michael Wright. This wildly democratic mix of sources perennially leads some educators to wring their hands about the state of knowledge, as yet another op-ed piece in the New York Times does today^[1] (“Searching for Dummies” by Edward Tenner). It’s a strange moment for the Times to publish this kind of lament; it seems like an op-ed left over from 1997, and as I’ve previously written in this space (and elsewhere with Roy Rosenzweig)^[2], contrary to Tenner’s one example of searching in vain for “World History,” online historical information is actually getting better, not worse (especially if you assess the web as a whole rather than complain about a few top search results). Anyway, Tenner does make one very good point: “More owners of free high-quality content should learn the tradecraft of tweaking their sites to improve search engine rankings.” This “tradecraft” is generally called “search engine optimization,” and I’ve long thought I should let those in academia (and other creators of reliable, noncommercial digital resources) in on the not-so-secret ways you can move your website higher up in the Google rankings (as well as in the rankings of other search engines).

1. Start with an appropriate domain name. Ideally, your domain should contain the top keywords you expect people searching for your topic to type into Google. At CHNM we love the name “Echo” for our history of science website^[3], but we probably should have made the URL `historyofscience.gmu.edu` rather than `echo.gmu.edu`. Professors like to name digital projects something esoteric or poetic, preferably in Greek or Latin. That’s fine. But make the URL something more meaningful (and yes, more prosaic, if necessary) for search engines. If you read Google’s Web Search API documentation, you’ll realize that their spider can

actually parse domain names for keywords, even if you run these words together.

2. If you've already launched your website, don't change its address if it already has a lot of links to it. "Inbound" links are the currency of Google rankings. (You can check on how many links there are to your site by typing "link:[your domain name here]" into Google.) We can't change Echo's address now, because it's already got hundreds of links to it, and those links count for a lot. (Despite the poetic name, we're in the top ten for "history of science.") There are some fancy ways to "redirect" sites from an old domain to a new one, but it's tricky.

3. Get as many links to your site as you can from high-quality, established, prominent websites. Here's where academics and those working in museums and libraries are at an advantage. You probably already have access to some very high-ranking, respected sites. Work at the Smithsonian or the Library of Congress? Want an extremely high-ranking website on any topic? Simply link to the new website (appropriately named, of course) from the home page of your main site (the home page is generally the best page to get a link from). Wait a month or two and you're done, because www.si.edu and www.loc.gov wield enormous power in Google's mathematical ranking system. A related point is...

4. Ask other sites to link to your site using the keywords you want. If you have a site on the Civil War, a bad link is one that says, "Like the Civil War? Check out this [site](#)^[4]." A helpful link is one that says, "This is a great site on the [Civil War](#)^[5]." If you use the [Google Sitemap service](#)^[6], it will tell you what the most popular keywords are in links to your site.

5. Include keywords in file names and directory names across your site, and don't skimp on the letters. This point is similar to #1, only for subtopics and pages on your site. Have a bibliography of Civil War books? Name the file "civilwarbibliography.html" rather than

just “biblio.html” or some nonsense letters or numbers.

6. Speaking of nonsense letters and numbers, if your site is database-driven, recast ungainly numbers and letters in the URL (known in geek-speak as the “query string”), e.g., change www.yoursite.org/archive.cfm?author=x15y&text=15325662&lng=eng to www.yoursite.org/archive/rousseau/emile/english_translation.html. Have someone who knows how to do “URL rewriting” change those URLs to readable strings (if you use the Apache web server software, as 70% of sites do, the software that does this is called “mod_rewrite”; it still keeps those numbers and letters in memory, but doesn’t let the human or machine audiences^[7] see them).

7. Be very careful about hiring someone to optimize your site, and don’t do anything shifty like putting white text with your keywords on a white background. Read Google’s warning about search engine optimization^[8] and shady methods and their propensity to ban sites for subterfuge.

8. Don’t bother with metatags. Google and other search engines don’t care about these old, hidden HTML tags that were supposed to tell search engines what a web page was about.

9. Be patient. For most sites, it’s a slow rise to the top, accumulating links, awareness in the real world and on the web, etc. Moreover, there is definitely a first-mover advantage—being highly ranked creates a virtuous circle, because by being in the top ten, other sites link to your site because they find it more easily than others. Thus Michael Wright’s page on Sputnik, which is nine years old, remains stubbornly in the top ten. But one of the advantages a lot of academic and nonprofit sites have over the Michael Wrights of the world is that we’re part of institutions that are in it for the long run (and don’t have ballroom dancing classes). I’m more sanguine than Edward Teller that in near future, great sites, many of them from academia, will rise to the top, and be found by all of those Google-centric students the educators worry about.

But these sites (and their producers) could use a little push. Hope this helps.

(You might also want to read the chapter Roy and I wrote on building an audience for your website^[9] in *Digital History*, especially the section that includes a discussion of how Google works^[10], as well as another section of the book on “Site Structure and Good URLs.”^[11])

This entry was posted on Sunday, March 26th, 2006 at 10:00 pm and is filed under Audience^[12], Google^[13], Search^[14], Tutorials^[15]. You can follow any responses to this entry through the RSS 2.0^[16] feed. You can leave a response^[17], or trackback^[18] from your own site.

References

1. ^ yet another op-ed piece in the New York Times does today (nytimes.com)
2. ^ I've previously written in this space (and elsewhere with Roy Rosenzweig) (www.dancohen.org)
3. ^ our history of science website (echo.gmu.edu)
4. ^ site (www.dancohen.org)
5. ^ Civil War (www.dancohen.org)
6. ^ Google Sitemap service (www.google.com)
7. ^ machine audiences (www.dancohen.org)
8. ^ Google's warning about search engine optimization (www.google.com)
9. ^ building an audience for your website (chnm.gmu.edu)
10. ^ the section that includes a discussion of how Google works (chnm.gmu.edu)
11. ^ “Site Structure and Good URLs.” (chnm.gmu.edu)
12. ^ View all posts in Audience (www.dancohen.org)
13. ^ View all posts in Google (www.dancohen.org)
14. ^ View all posts in Search (www.dancohen.org)
15. ^ View all posts in Tutorials (www.dancohen.org)

16. [^ RSS 2.0](#) (www.dancohen.org)
17. [^ leave a response](#) (www.dancohen.org)
18. [^ trackback](#) (www.dancohen.org)

Excerpted from *Dan Cohen's Digital Humanities Blog » Blog Archive » Search Engine Optimization for Smarties*

<http://www.dancohen.org/2006/03/26/search-engine-optimization-for-smarties/>

READABILITY — An Arc90 Laboratory Experiment

<http://lab.arc90.com/experiments/readability>