



# Search Engine Optimization & Google Search Appliance at Texas A&M University

# Overview

- About myself
- Server name: search.tamu.edu
- License covers all sites on \*.tamu.edu
- Powered by the actual Google search algorithms
- More than 572,000 pages indexed by default
- More than 477,000 searches each month
- Crawls hundreds of thousands of pages a day.
- Recrawls each site frequently, depending on Google algorithms.
- More flexible and accurate than other search engines: try "bookstore" or "alumni"

# How to Use the Google Search Appliance



- Any webmaster can add a search box to:
  - Search your site
  - Related sites
  - Department or College
  - University
  - the Web
  - Can limit by date or numerical range, meta-data or file type.
- To prevent indexing of content, disallow user agent “tamu-googlebot” in your robots.txt.
- To update or remove sensitive content quickly, email the URL to webmaster@tamu.edu.

# Search Engine Optimization Basics



- With the Google Search Appliance, visitors can find your pages more easily – as long as you’ve practiced good search engine optimization.
  - Unique, descriptive titles
  - Unique meta-descriptions & keywords
  - Simple directory structure – one URL per page
  - Descriptive file names and URLs
  - Easy, hierarchical navigation, mostly text, not images
  - Descriptive, concise anchor text for links
  - Headings as your site outline
  - Brief, descriptive file names and alt text for images
  - Focused content that’s worth linking to

# Crawl Options

- Crawl URLs: Crawl Only, Do Not Crawl
- Freshness Tuning:
  - Crawl frequently
  - Infrequently
  - Now
- Duplicate Hosts: careercenter=jobs=careers
- Host Load Schedule:
  - 4 concurrent connections by default

# Customization

- KeyMatches
  - Like Google Adwords, keywords lets you promote specific web pages on your site.
  - Makes sure the most important results are on top on our searches
  - Proper search optimization is still better
- Related Queries
  - “You could also try...”
  - Encourages searchers to learn/use your chosen keywords

# Advanced Customization

- Collections: customized site search or across sites
  - Especially useful for interdisciplinary sites
- **Front Ends**: customizable for departments or divisions
  - **Output Format**: modify look, feel, format, and results
    - Page Layout Helper: custom logo, header, footer, options
    - **XSLT Stylesheet Editor**: complete customization
    - Add KeyMatches and Related Queries for your site

# Status and Reports

- Crawl Diagnostics: Successful, Errors, Excluded
  - Not found, conversion error, authentication failed, forbidden, etc.
  - Useful for creating sitemaps, finding “phantom 404’s”, inbound links
- Content Statistics: by Mime Type
- Serving Status: searches per minute
- Search Reports: most popular search terms in date range
- Search Logs: same format as server logs

# Resources

From the TAMU Webmasters Blog

<http://webmaster.tamu.edu>:

- [Search Engine Optimization according to Google](#)
- [Adding your own site search box](#)
- [Our Google is not like their Google](#)
- [Tweaking the search engine](#)
- [Ah, the luxury of our own search engine](#)

From the Marketing and Communications site:

- [Creating Search Forms](#)

# Questions? Sure!

- Michael McGinnis
  - [webmaster@tamu.edu](mailto:webmaster@tamu.edu)
  - [mdmcginnis@tamu.edu](mailto:mdmcginnis@tamu.edu)
  - <http://webmaster.tamu.edu>
  - <http://twitter.com/tamuwww>
  - <http://twitter.com/mdmcginn>