take your web site to the top

how to boost your search rankings

brian gorbett
architect – microsoft
briangorbett.com
twitter.com/briangorbett
Agenda

- Search Engine Optimization Basics
  - How does search engines work
  - What Impacts SEO
  - Best Practices

- IIS SEO Toolkit
  - Site Analysis
    - Key Features
    - Demo
  - Robots
  - Sitemap and Sitemap Index Editor
Search engine optimization (SEO)
  ◦ is the process of improving the volume or quality of traffic to a web site from search engines via "natural" ("organic" or "algorithmic") search results. (Wikipedia.org)

Why should you care?
  ◦ Different studies have concluded that in average, 50%-80% of web site traffic originates from a Search engine
How does Search Engines work?

1. Finds a Reference to your Site
   ◦ Usually through a Referral in another Web Site or through Submitting your URL to their SE

2. Requests the /Robots.txt (Robots Exclusion Protocol or REP)
   ◦ List of Files/Directories that it should not visit
   ◦ Discovers SiteMaps on it. (sitemap.org)

3. Process the URL
   ◦ If allowed by Robots
     • Downloads the Resource
     • Analyzes the Content
     • Discovers links on it... and continues to 3.
What Impacts SEO

There are several factors that impact SEO:

- **Links**
  - External and Internal
  - The more the better (quality matters)

- **Navigation/Organization**
  - Clear/Consistent navigation into and around your site
  - Valid Markup
  - Does not require authentication, cookies, sessions, scripts, objects, referral, etc.

- **Content**
  - Proper page form and html usage
  - Title, Keywords, Headings, Link Text, Bolded Text, etc
  - Lost content (404)

- **Content Relevance**
  - Is your Web Site considered an authority in the community?
  - How fresh is the content?
10 KB of HTML

Example of bad SEO:

http://www.microsoft.com/surface

What the browser gets
Best Practices

- Provide as much information as possible using Text
  - Use Title, Heading, Description, Keywords, Bolded Text, alt text, `<noscript>`, etc
  - Use proper HTML semantics to highlight importance
- Maximize the time spent by the crawlers in your site
  - Use Robots.txt to filter non-relevant areas
  - Use Sitemaps to guide crawler to difficult places
  - Clear descriptive internal links to content
  - Avoid duplicate content
- Optimal & Simple Navigation
  - Ensure JavaScript is not required to access pages
  - If using images/RIA make sure to provide textual representation
- Beware of...
  - Malware, Hidden Text, Too many arguments in URLs, Login pages, unnecessary redirects,
Sites can get large and complex so tracking...
- Is the content and structure of my site optimized for a search engine to crawl it?
- How do I ensure that all external links on my site are valid?
- If I change something on my site – what will it break?
- Are all pages on my site easily discoverable?
- Do all pages on my site load correctly?
- Is robots.txt too restrictive or too permissive?
- Is sitemap.xml up to date with valid URL’s?
- What are the most difficult to access pages?
- Has anybody inserted a reference to a bad web site?
- What has changed since last time I analyzed the site?
- Etc...
IIS SEO Toolkit

- Improve the volume and quality of traffic to your web site from search engines
- Improve the content on your web-site for human users
- Help you analyze and fix common SEO problems in your Web Site and look at it from a Search Engine perspective
- Run it at development, staging, production, any time updates are made, etc.
- Consists of Three Main Features
  - Site Analysis
  - Robots Editor
  - Sitemap and Sitemap Index Editor
Key Features

- Integrated with IIS 7
  - Built on IIS 7 Admin Tool Framework
  - Requires Windows Vista, Windows 7, Windows Server 2008/R2

- Analyze local and remote Web sites
  - Run in your own development or staging environments
  - Run against any Web Server and any framework
  - Keep history and track changes over time, compare side-by-side older versions of your Site

- Over 45 best-practices included
  - Built-in Reports, also include performance analysis too

- Powerful Query Engine
  - Create your own queries
  - Save/Open existing queries to reanalyze after making site changes

- Extensible
  - Add your own parsing logic
  - Add your own violation detection logic
  - Extend the User Interface with your own functionality
Robots Editor

- Simple tool to generate Robots.txt files
  - Easily author Robots file for your Web site
  - Support for:
    - Physical locations: files, folder, etc
    - Virtual locations: URLs discovered by Site Analysis report
  - Test the rules using Site Analysis
Sitemap and Sitemap Index Editor

- Simple tool to generate Sitemap and Sitemap Index files
  - Support for:
    - Physical locations: files, folder, etc
    - Virtual locations: URLs discovered by Site Analysis report
    - Automatically fills additional data like Last Modified date based on content
  - Test the Sitemap files using Site Analysis
Site Analysis Report

Start Date: 7/27/2009 1:43:33 PM
End Date: 7/27/2009 1:59:45 PM
Downloaded Items: 1082
Number of Items: 33105

Violations Summary

<table>
<thead>
<tr>
<th>Violation Title</th>
<th>Violation Level</th>
<th>Violation Category</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>The &lt;img&gt; tag does not have an ALT attribute.</td>
<td>Warning</td>
<td>SEO</td>
<td>2,071</td>
</tr>
<tr>
<td>The page contains unnecessary redirects.</td>
<td>Information</td>
<td>SEO</td>
<td>1,523</td>
</tr>
<tr>
<td>The link text is not relevant.</td>
<td>Warning</td>
<td>SEO</td>
<td>778</td>
</tr>
<tr>
<td>The description is missing.</td>
<td>Warning</td>
<td>SEO</td>
<td>406</td>
</tr>
<tr>
<td>The page was excluded by a nofollow attribute.</td>
<td>Information</td>
<td>SEO</td>
<td>406</td>
</tr>
<tr>
<td>The &lt;h1&gt; tag is missing.</td>
<td>Warning</td>
<td>SEO</td>
<td>186</td>
</tr>
<tr>
<td>The page contains a large amount of script code.</td>
<td>Warning</td>
<td>Performance</td>
<td>55</td>
</tr>
<tr>
<td>The title is too long.</td>
<td>Warning</td>
<td>SEO</td>
<td>42</td>
</tr>
<tr>
<td>The page contains invalid markup.</td>
<td>Warning</td>
<td>Content</td>
<td>32</td>
</tr>
<tr>
<td>The title is empty.</td>
<td>Error</td>
<td>SEO</td>
<td>26</td>
</tr>
<tr>
<td>The page contains multiple &lt;h1&gt; tags.</td>
<td>Warning</td>
<td>Standards</td>
<td>15</td>
</tr>
<tr>
<td>The page contains a large number of Cascading Style Sheets.</td>
<td>Warning</td>
<td>Performance</td>
<td>5</td>
</tr>
<tr>
<td>The content type for the resource is incorrect.</td>
<td>Warning</td>
<td>Content</td>
<td>3</td>
</tr>
<tr>
<td>An unexpected error has occurred.</td>
<td>Error</td>
<td>Content</td>
<td>2</td>
</tr>
<tr>
<td>The page contains multiple canonical formats.</td>
<td>Error</td>
<td>SEO</td>
<td>2</td>
</tr>
</tbody>
</table>
There are three main extensibility points provided by Microsoft.Web.Management SEO.Client.dll:

- **Crawler Module**
  - Provide custom parsing logic for new content types and extend the set of violations with your own.

- **Site Analyzer Extension**
  - Extend the Site Analysis User Interface with your own set of tasks.

- **Sitemap Extension**
  - Extend the Sitemaps User Interface with your own set of tasks.
Summary

- IIS SEO Toolkit
  - Automatically analyze your Web site for SEO Best Practices
  - Run it against any application running in any server built in any framework
  - Free Download: http://www.iis.net/expand/SEOToolkit
  - Other IIS Extensions:
    - http://www.iis.net/expand
  - SEO Posts: